PERTANIKA JOURNAL OF SOCIAL SCIENCES & HUMANITIES

About the Journal

Overview

Pertanika Journal of Social Sciences & Humanities is the official journal of Universiti Putra Malaysia. It is an open-access online scientific journal. It publishes original scientific outputs. It neither accepts nor commissions third party content.

Recognised internationally as the leading peer-reviewed interdisciplinary journal devoted to the publication of original papers, it serves as a forum for practical approaches to improve quality in issues pertaining to social and behavioural sciences as well as the humanities.

Pertanika Journal of Social Sciences & Humanities is a **quarterly** (*March*, *June*, *September*, *and December*) periodical that considers for publication original articles as per its scope. The journal publishes in **English** as well as in **Bahasa Malaysia** and it is open for submission by authors from all over the world.

The journal is available world-wide.

Aims and scope

Pertanika Journal of Social Sciences & Humanities aims to develop as a pioneer journal for the social sciences with a focus on emerging issues pertaining to the social and behavioural sciences as well as the humanities. Areas relevant to the scope of the journal include Social Sciences—architecture and habitat, consumer and family economics, consumer law, education, media and communication studies, political sciences and public policy, population studies, psychology, sociology, and tourism; Humanities—arts and culture, dance, language and linguistics, literature, music, philosophy, and sports.

History

Pertanika Journal of Social Sciences & Humanities was founded in 1993 and focuses on research in social and behavioural sciences as well as the humanities and its related fields.

Vision

To publish journal of international repute.

Mission

Our goal is to bring the highest quality research to the widest possible audience.

Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions can expect to receive a decision within 120 days. The elapsed time from submission to publication for the articles averages 180 days. We are working towards decreasing the processing time with the help of our editors and the reviewers.

Abstracting and indexing of Pertanika

Pertanika Journal of Social Sciences & Humanities is now over 27 years old; this accumulated knowledge and experience has resulted the journal being abstracted and indexed in SCOPUS (Elsevier), Clarivate Web of Science (ESCI), EBSCO, DOAJ, Agricola, ASEAN CITATION INDEX, ISC, Microsoft Academic, Google Scholar, and MyCite.

Citing journal articles

The abbreviation for Pertanika Journal of Social Sciences & Humanities is *Pertanika J. Soc. Sci. & Hum.*

Publication policy

Pertanika policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in proceedings.

Code of ethics

The *Pertanika* journals and Universiti Putra Malaysia take seriously the responsibility of all of its journal publications to reflect the highest publication ethics. Thus, all journals and journal editors are expected to abide by the journal's codes of ethics. Refer to *Pertanika*'s **Code of Ethics** for full details, or visit the journal's web link at: http://www.pertanika.upm.edu.my/code_of_ethics.php

Originality

The author must ensure that when a manuscript is submitted to *Pertanika*, the manuscript must be an original work. The author should check the manuscript for any possible plagiarism using any program such as Turn-It-In or any other software before submitting the manuscripts to the *Pertanika* Editorial Office, Journal Division.

All submitted manuscripts must be in the journal's acceptable **similarity index range**:
≤ 20% – PASS; > 20% – REJECT.

International Standard Serial Number (ISSN)

An ISSN is an 8-digit code used to identify periodicals such as journals of all kinds and on all media–print and electronic.

Pertanika Journal of Social Sciences and Humanities: e-ISSN 2231-8534 (Online).

Lag time

A decision on acceptance or rejection of a manuscript is reached in 120 days (average). The elapsed time from submission to publication for the articles averages 180 days.

Authorship

Authors are not permitted to add or remove any names from the authorship provided at the time of initial submission without the consent of the journal's Chief Executive Editor.

Manuscript preparation

Most scientific papers are prepared according to a format called IMRAD. The term represents the first letters of the words *Introduction*, *Materials and Methods*, *Results*, *And Discussion*. IMRAD is simply a more 'defined' version of the "IBC" (*Introduction*, *Body*, *Conclusion*) format used for all academic writing. IMRAD indicates a pattern or format rather than a complete list of headings or components of research papers; the missing parts of a paper are: *Title*, *Authors*, *Keywords*, *Abstract*, *Conclusions*, *References*, and *Acknowledgement*. Additionally, some papers include *Appendices*. For manuscripts in **Bahasa Malaysia**, the title, abstract and keywords should be written in both **English** and **Bahasa Malaysia**.

The *Introduction* explains the scope and objective of the study in the light of current knowledge on the subject; the *Materials and Methods* describes how the study was conducted; the *Results* section reports what was found in the study; and the *Discussion* section explains meaning and significance of the results and provides suggestions for future directions of research. The manuscript must be prepared according to the journal's **Instruction to Authors** (http://www.pertanika.upm.edu.my/Resources/regular_issues/Regular_Issues_Instructions_to_Authors.pdf).

Editorial process

Authors who complete any submission are notified with an acknowledgement containing a manuscript ID on receipt of a manuscript, and upon the editorial decision regarding publication.

Pertanika follows a **double-blind peer -review process**. Manuscripts deemed suitable for publication are sent to reviewers. Authors are encouraged to suggest names of at least 3 potential reviewers at the time of submission of their manuscripts to *Pertanika*, but the editors will make the final selection and are not, however, bound by these suggestions.

Notification of the editorial decision is usually provided within 120 days from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

As articles are double-blind reviewed, material that may identify authorship of the paper should be placed only on page 2 as described in the first-4-page format in Pertanika's Instruction to Authors (http://www.pertanika.upm.edu.my/Resources/regular_issues/Regular_Issues_Instructions_to_Authors.pdf).

The journal's peer review

In the peer review process, 2 or 3 referees independently evaluate the scientific quality of the submitted manuscripts. At least 2 referee reports are required to help make a decision.

Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

Operating and review process

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are 7 steps to the editorial review process:

1. The journal's Chief Executive Editor and the Editor-in-Chief examine the paper to determine whether it is relevance to journal needs in terms of novelty, impact, design, procedure, language as well as presentation and allow it to proceed to the reviewing process. If not appropriate, the manuscript is rejected outright and the author is informed.

2. The Chief Executive Editor sends the article-identifying information having been removed, to 2 to 3 reviewers. They are specialists in the subject matter of the article. The Chief Executive Editor requests that they complete the review within 3 weeks.

   Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions for strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the research field.

3. The Editor-in-Chief examines the review reports and decides whether to accept or reject the manuscript, invite the authors to revise and resubmit the manuscript, or seek additional review reports. In rare instances, the manuscript is accepted with almost no revision. Almost

without exception, reviewers' comments (to the authors) are forwarded to the authors. If a revision is indicated, the editor provides guidelines to the authors for attending to the reviewers' suggestions and perhaps additional advice about revising the manuscript.

4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors return a revised version of the paper to the Chief Executive Editor along with specific information describing how they have addressed the concerns of the reviewers and the editor, usually in a tabular form. The authors may also submit a rebuttal if there is a need especially when the authors disagree with certain comments provided by reviewers.

5. The Chief Executive Editor sends the revised manuscript out for re-review. Typically, at least 1 of the original reviewers will be asked to examine the article.

6. When the reviewers have completed their work, the Editor-in-Chief examines their comments and decides whether the manuscript is ready to be published, needs another round of revisions, or should be rejected. If the decision is to accept, the Chief Executive Editor is notified.

7. The Chief Executive Editor reserves the final right to accept or reject any material for publication, if the processing of a particular manuscript is deemed not to be in compliance with the S.O.P. of *Pertanika*. An acceptance notification is sent to all the authors.

9. The editorial office ensures that the manuscript adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any minor queries by the editorial office. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, **only essential changes are accepted**. Finally, the manuscript appears in the pages of the journal and is posted on-line.

ABSTRACTING AND INDEXING OF PERTANIKA JOURNALS

The journal is indexed in Clarivate Web of Science (JCR-ESCI), SCOPUS (Elsevier), BIOSIS, Google Scholar, MyCite, ISC. In addition, Pertanika JSSH is recipient of "CREAM" Award conferred by Ministry of Higher Education (MoHE), Malaysia.

Pertanika Journal of Social Sciences & Humanities
Vol. 32 (3) Sep. 2024

Contents

Foreword

Welcome to the third issue of 2024 for the Pertanika Journal of Social Sciences and Humanities (PJSSH)!

PJSSH is an open-access journal for studies in the field of Social Sciences and Humianities published by Universiti Putra Malaysia Press. It is independently owned and managed by the university for the benefit of the world-wide science community.

This issue contains 20 articles; four review articles; two case studies; and the rest are regular articles. The authors of these articles come from different countries namely China, Indonesia, Kingdom of Saudi Arabia, Malaysia, Pilippines and Thailand.

A selected article from the scope of psychology studies, titled "The Impact of Antisocial Media towards Aggressive Behavior among Online Gamers in Malaysia," examines the predictors of aggressive behaviors among online gamers using a purposive and snowball sampling technique on 384 online gamers. This article concludes that aggressive behaviors have a positive relation with antisocial media. This study proves that high levels of moral identity, self-esteem, and a low level of antisocial media exposure are important predictors in decreasing aggressive behavior among local online gamers. Details of this study are available on page 1027.

A study entitled, "The Digital Dynamics of Political Engagement among Filipino Youth: Examining Participation in Social Media Platforms" discussed how youth participation has been pivotal in shaping the political landscape of the Philippines. This study employed the Social Media Political Participation Scale to analyze the online political behaviors of Filipino Generation Z members. It compared their engagement across several online platforms, including Facebook, Twitter, Instagram, YouTube, and TikTok. The findings highlight the complex nature of digital political engagement and its impact on political involvement, revealing that Instagram and TikTok are the primary platforms driving political participation among respondents. Details of this study are available on page 1051.

An investigation on incorporating moral education into college English courses to support the holistic development of students was conducted by Lina Wang and colleagues. This study employed a qualitative case study approach to investigate the changes in pedagogical content knowledge (PCK) among college English teachers during the transition from college English (CE) to moral education integrated college English (MECE) over three years. The findings indicate that during the transition from CE to MECE, changes in pedagogical content knowledge (PCK)

were influenced by several key factors: engaging in objective-based study with reflection, participating in training sessions and workshops, interacting within effective online and face-to-face communities, and adopting a co-learning approach with students. The detailed information of this article is presented on page 1201.

We anticipate that you will find the evidence presented in this issue to be intriguing, thought-provoking and useful in reaching new milestones in your own research. Please recommend the journal to your colleagues and students to make this endeavour meaningful.

All the papers published in this edition underwent Pertanika's stringent peer-review process involving a minimum of two reviewers comprising internal as well as external referees. This was to ensure that the quality of the papers justified the high ranking of the journal, which is renowned as a heavily-cited journal not only by authors and researchers in Malaysia but by those in other countries around the world as well.

We would also like to express our gratitude to all the contributors, namely the authors, reviewers, Editor-in-Chief and Editorial Board Members of PJSSH, who have made this issue possible.

PJSSH is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

Chief Executive Editor
Mohd Sapuan Salit
executive_editor.pertanika@upm.edu.my

Malaysia Modern Rural Ranking Index: Validated Dimensions, Criteria Groups, and Performance Criteria

Mohd Fadzil Abdul Rashid[1]*, Kushairi Rashid[1], Muhamad Azrul Azwan Azman[1], Aizazi Lutfi Ahmad[1] and Hafriz Suhairi Mohd Rejab[2]

[1]Department of Built Environment Studies and Technology, College of Built Environment, Universiti Teknologi MARA, Perak Branch, 32610 Seri Iskandar, Perak, Malaysia
[2]IZM Consultant, Taman TTDI Jaya, 40150 Shah Alam, Selangor, Malaysia

ABSTRACT

In Malaysia, rural areas are essential economic focal points for improving rural livelihoods and reducing the life quality gaps between urban and rural areas. To that extent, rural assets and resources, such as agricultural, entrepreneurial, business, and tourism, along with rural infrastructure and service centres, become important elements to further explore and assess for synergising rural change towards a modern approach. To our knowledge, no attempt has been made to assess rural development progress based on the transition towards a modern rural approach. Inspired by this, the study aims to offer a Malaysia modern rural ranking index, the Malaysia Assessment Measure for Modern Rural Development (MAMRD)—a missing link approach for assessing rural development progress towards a modern rural system. The MAMRD was validated based on focus group discussions (FGDs) on three case studies and finally constructed on the three dimension-objectives measure, comprising 13 criteria groups and 141 criteria, to rank a village in a MAMRD star rating index. Using the MAMRD for the rural assessment is reliable for assessing the rural performance towards the future niches of rural development in Malaysia—rural resilient-liveability-and-smart.

Keywords: MAMRD, modern rural, resilient-liveability-and-smart, rural infrastructure, technology practices

INTRODUCTION

In Malaysia, rural areas serve as vital guardians of food security, economic resources, and assets while also being wonderful places to work, live, invest,

and visit—a legacy for future generations. Unfortunately, rural assets and resources, such as agriculture, entrepreneurship, business, and tourism, still require diversification (Rural Development Policy 2030 or *Dasar Pembangunan Luar Bandar 2030* [*DPLB*], 2018; Rashid et al., 2021). Without this diversification, there will be a shortfall in bridging the gap between rural and urban living standards. This alarming issue has resulted in youth out-migration from rural to urban areas as they seek more significant opportunities and desire to overcome the limitations of financial resources (Kusumo et al., 2023). Consequently, this trend has led to a dearth of rural human resources and low productivity (*DPLB*, 2018; Rashid et al., 2021), demanding immediate attention and action. In addition, it is crucial to establish a link between the national development policies and strategies while fostering understanding among rural stakeholders to effectively translate rural development agendas into actionable initiatives.

In line with the government's directive aspirations, the Sustainable Development Goals (SDGs), particularly Goal 9: Industry, Innovation, and Infrastructure and Goal 10: Reduced Inequalities (United Nations, 2015), along with the advent of new technologies in rural practices both nationally and globally, highlight four crucial factors that underscore the urgency of assessing rural development progress, as summarised in Figure 1.

These outlined factors call for a notable contribution to existing government policies, such as the National Rural Physical Planning Policy 2030 or *Dasar Perancangan Fizikal (DPF) Desa Negara* 2030 (2017), particularly during the implementation stage. Therefore, adopting a place- or strength-based strategy is crucial, enabling all rural settlements to achieve remarkable progress and increase sustainability and prosperity. Inspired by this, the current paper aims to offer a Malaysia modern rural ranking index, the Malaysia Assessment Measure for Modern Rural Development (MAMRD)—a missing link approach for assessing rural development progress towards a modern rural ecosystem.

LITERATURE REVIEW

A MAMRD is an extension of a Framework of Modern Rural Development (FMRD),

Figure 1. Factors alarming to the need for rural development measure (*Source*: Authors' work, 2023)

effectively translating the government's directive aspirations into a quantifiable measure of rural progress. Rashid et al. (2021) have developed the FMRD as a new approach to synergise rural transformation and solutions. The FMRD is structured as an objective mechanism for assessing rural transition towards a modern approach, functioning as a translation of the essential criteria required for modern rural development. It is grounded on three dimensions of objective measures: (1) rural economic boosters and catalyst infrastructures, representing the concept of resilience; (2) rural characteristics and social well-being infrastructures, representing the concept of liveability; and (3) smart and green technology practices, representing the concept of smartness. It relies on the adequacy of rural infrastructures, social facilities, services, and technology practices, particularly in agricultural, entrepreneurship, business, and tourism developments. Then, it is strengthened by incorporating the best practices or ideas from a smart village approach into the existing Malaysia rural development approach, which emphasises resilience and liveability, as outlined in *DPF Desa Negara 2030* (2017) and *DPLB 2030* (2018).

In Malaysia, the concept of 'modern' or 'rural modernisation' has long been incorporated into rural development agendas, starting as early as 1971 with the National Economic Policy (NEP, 1971–1990). This policy aims to modernise rural areas by introducing advanced techniques in agriculture, improving irrigation and

drainage systems, establishing credit and marketing institutions, developing rural infrastructure, enhancing transportation and communication networks, and introducing new secondary schools in rural regions (Bruton, 2007). Building upon this, the *Falsafah dan Strategi Baharu Pembangunan Luar Bandar* (A New Philosophy of Rural Development) in 1994 focused on empowering people (human development) and initiating the *Gerakan Desa Wawasan* (Visionary Capability Movement) programme. This programme aimed to uplift villages, making them more advanced, attractive, and beneficial. The key approach was to raise awareness and change the attitudes of rural residents towards rural transformation. In 2010, the Rural Development Master Plan was introduced as a strategic action plan blueprint to promote rural development. The plan centred around three main pillars: (1) fostering a sustainable and advanced economy, (2) creating a prosperous society, and (3) ensuring environmental sustainability (*Ministry of Rural and Regional Development*, 2010). These efforts are further emphasised in the current policies of *DPF Desa Negara 2030* (2017) and *DPLB 2030* (2018), which highlight specific niches and discoveries for rural advancement.

In Malaysia's context of rural development, one significant strategy initiated by the government, namely a village action plan strategy launched in 2009, marks a new movement of rural planning in Malaysia (Ngah et al., 2010). It is a rural development strategy also

under the Visionary Capability Movement Programme to strengthen the capacity building of rural leaderships, such as the Village Community Management Council or *Majlis Pengurusan Komuniti Kampung* (*MPKK*) and rural people in planning and executing the village activities and projects. Ngah et al. (2010) demonstrate their research findings on the strategy implemented in the initial 17 villages throughout Malaysia. They found that village people have the capability to participate in planning and implementing development projects in their villages according to their needs and aspirations. In this manner, the knowledge and experience available from the various backgrounds of village people could be easily transferred into the plan-making and implementation processes if adequately organised and encouraged. They also found that the village people were ready and able to identify a range of development priorities, covering physical, economic, social and institutional dimensions. This participatory empowerment that has already been nurtured in rural people can potentially be embedded in the new environment of the rural digital era to well-integrated the government's directive aspirations and rural people's needs and capabilities to synergise rural transformation and solutions in a quantifiable way.

Therefore, the formulation of the FMRD by Rashid et al. (2021) is timely for incorporating technology disruption and practices (transpired from a smart village approach) into rural activities at the community level, including the Internet of Things (IoT), robotics, and big data analytics. It offers solutions to everyday problems in farming, SMEs, energy usage and healthcare in rural areas (Alabdali et al., 2023).

All processes underwent three major stages. Stage 1 involved reviewing contemporary rural development concepts. Stage 2 involved analysing the content of *DPF Desa Negara 2030* (2017) (Thrust 2 to Thrust 5) and *DPLB 2030* (2018), extracting the key performance criteria. Lastly, Stage 3 involved a single-round expert view survey using a structured assessment form, which included a rating priority exercise to validate and assess the importance and relevance of dimensions, group criteria, and performance criteria that reflect modern rural development outcomes. As a result, the FMRD, tailored to the national development agendas and smart technology practices, would enhance rural opportunities for change and align with future government aspirations (Figure 2). Rashid et al. (2021) provide detailed information about the formulation process, criteria groups, and performance criteria of the FMRD.

In continuation (extension) of the FMRD, the current paper attempts to provide an overview of the completed work on this rural performance measure, the so-called Malaysia Assessment Measure for Modern Rural Development (MAMRD). The MAMRD is a new measurement tool that assigns a rating (ranging from zero to six stars) to assess a village's progress according to the MAMRD criteria. This measurement tool takes
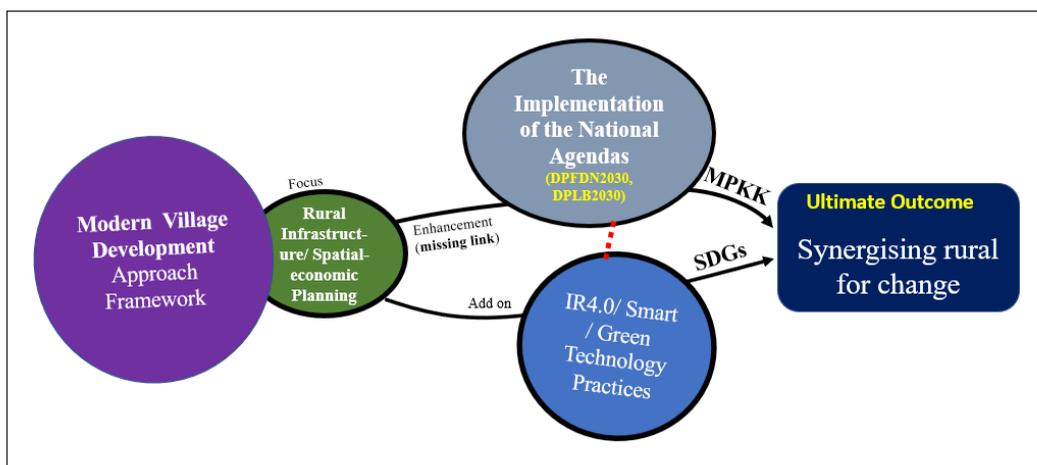
Figure 2. The FMRD serves as a missing link and catalyst for rural change in Malaysia (*Source:* Authors' work, 2023)

into account the aspirations of the rural communities, particularly the youth, amidst the challenges posed by globalisation. To the best of our knowledge, this study represents a pioneering rural measurement index in Malaysia, highlighting its novelty. Others, such as Rashid et al. (2019), work in different dimensions, concentrating more on the rural revitalisation framework based on five economic performance factors: economic, human, social, cultural and environmental. By formulating a rating tool index for rural development progress incorporating technology practices for rural livelihoods and solutions, this study emphasises the importance of adopting a modern approach to rural development. Technology applications have become vital to narrowing the gap between rural and urban living standards and driving rural development to be smarter, more advanced and more efficient. It is worth noting that it aligns with the government effort, where the Malaysian government

initiated the National Fourth Industrial Revolution (4IR) in 2021 (Economic Planning Unit, 2021a) as the guiding principle, positioning Malaysia to stay ahead of the 4IR technologies curve and optimise its associated benefits.

METHODOLOGY

Overall, this research uses a qualitative approach involving literature review and contents analysis, experts' views, and FGD sessions. However, a quantitative method has also been used in the data analysis, particularly in a decision-making problem regarding MAMRD's dimensions and criteria groups. Figure 3 depicts the five stages of the research process involved in developing MAMRD.

There are two main processes (Stages 4 and 5) in formulating a MAMRD measure: (a) weighting the dimensions and criteria groups and (b) the implementation of MAMRD through FGDs.
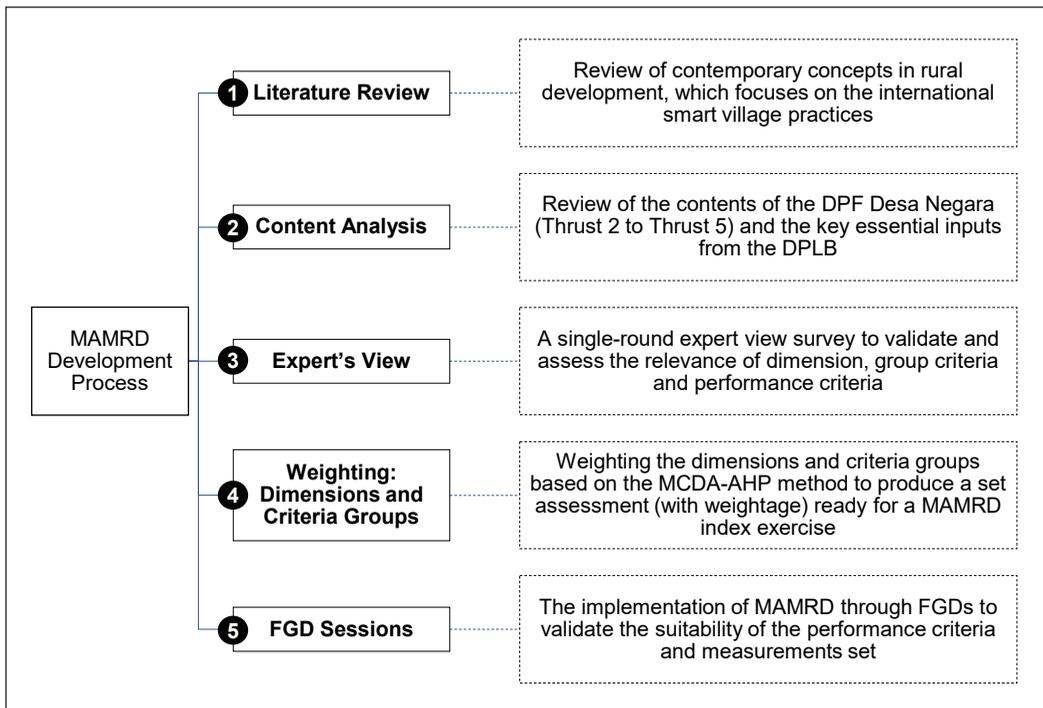
Figure 3. Overall research process: The five stages involved in MAMRD development (*Source*: Authors' work, 2023)

Note. Stages 1–3 can be further referred to Rashid et al. (2021)

Weighting the Dimensions and Criteria Groups

According to the experts' views (Rashid et al., 2021), the MAMRD should emphasise rural economic boosters and catalyst infrastructures (D1) more than rural characters and social well-being infrastructures (D2) and smart and green technology practices (D3). Thus, the three dimensions hold varying degrees of importance in the overall assessment, similar to the criteria groups that possess different extents of influence. This situation poses a decision-making challenge for the MAMRD measure, necessitating further refinement to enhance precision. Consequently, the study employed Multi-Criteria Decision Analysis

(MCDA), which was based on Saaty's (1980) Pairwise Comparison Matrix (PCM) in the context of the Analytic Hierarchy Process (AHP). This procedure aims to enhance the ranking method (Saaty, 1980; Saaty & Kearns, 1985) to ensure that each dimension and criterion group contributes appropriately to the measurement results. This technique is a relatively popular and practical approach in decision-making processes related to urban development strategies (Malczewski, 1999; Rashid et al., 2023). As such, Table 1 shows the dimension and criteria group weighting results. D1 was weighted at 54%, D2 was weighted at 30%, and finally, D3 was set at 16%. This weighting process was carried out

Table 1

Weightage of each MAMRD's dimension and criteria group based on PCM

Dimensions	Criteria Groups
D1: Rural Economic Boosters and Catalyst Infrastructures *Weight = 0.54*	CG1-D1: Economic and Rural Services Centre (Town) *Weight = 0.50*
	CG2-D1: Rural Growth Centre (RGC) *Weight = 0.25*
	CG3-D1: Rural economic cluster (agricultural, entrepreneurial, and tourism) *Weight = 0.25*
	Consistency Ratio (RC) = 0.000
D2: Rural Characters and Social Well-Being Infrastructures *Weight = 0.30*	CG1-D2: Rural spatial characters and heritage *Weight = 0.064*
	CG2-D2: Transportation networks of rural-town-city, and rural accessibility *Weight = 0.275*
	CG3-D2: Efficient infrastructure *Weight = 0.402*
	CG4-D2: Internal village amenities *Weight = 0.098*
	CG5-D2: Rural governance (MPKK) and database *Weight = 0.161*
	Consistency Ratio (RC) = 0.024
D3: Smart and Green Technology Practices *Weight = 0.16*	CG1-D3: Rural agricultural, infrastructures, technologies, and innovations *Weight = 0.40*
	CG2-D3: Rural entrepreneurial technologies and innovations *Weight = 0.29*
	CG3-D3: Rural marketing and e-commerce *Weight = 0.13*
	CG4-D3: Village smart and green technology practices *Weight = 0.12*
	CG5-D3: Community-IoT-based smart technology practices *Weight = 0.06*
Consistency Ratio (RC) = 0.008	***Consistency Ratio (RC) = 0.023***

Source: Authors' work (2023)

in accordance with the Consistency Ratios (CR) being below the threshold of 0.10, ensuring the validity of the results.

Within the dimensions, the criteria groups were compared. In D1, for example, the criteria group with the highest weight is the economic and rural services centre (CG1-D1), with a weight of 0.50, followed by CG2-D1 and CG3-D1, both carrying a weight of 0.25. Further details can be seen in Table 1. The individual performance criteria, however, were not assigned weights because they were given scores of 1 (available) or 0 (not available) within their respective criteria groups. This approach allows each village or case study to be evaluated according to its performance within the criteria. It is important to point out that the

overall scores and rankings are contingent upon data availability, thus accommodating flexibility in the face of data limitations.

The Implementation of MAMRD Through FGDs

The application of MAMRD in various village categories serves three main purposes. Firstly, it aims to validate the suitability of the performance criteria and measurements set. Secondly, it seeks to determine the score (index) for each participating village. Finally, it aims to identify issues and challenges while finding solutions, enabling lessons to be learned for further improvements in rural development strategies. However, this paper focuses on the first result, clarifying and validating the 139 performance criteria outlined in the FMRD (Rashid et al., 2021). The MAMRD intends to inform rural development actors, including the funding agencies, on the strengths and weaknesses of each participating village to become a resilient, liveable, and smart village. Therefore, the MAMRD is based on the active participation of rural actors in fostering smart and sustainable rural livelihoods. It acknowledges their ability to form an aspiration (specific goal or desired state) and adapt to new solutions, providing creative solutions, innovation, and diversity to their communities. This participatory approach is critical for a village to be considered modern, embodying the principles of being resilient, liveable, and smart.

As mentioned earlier, the availability of each criterion in the case study will be given one mark each. The total mark obtained was then weighted based on the dimensions and criteria groups (Table 1). Dimension 1 carries a weight of 54% of the total marks, Dimension 2 is weighted at 30%, and Dimension 3 is at 16% of the marks. All marks are summed up to calculate each village's final score or overall index under study. Based on the obtained final score, the village is assigned a star rating reflecting its MAMRD index performance. The star ratings range from 0 to six stars, corresponding to a score of 0 to 100 (Table 2).

As presented in Table 2, the minimum score required for each village to be eligible for a star is 15 points. The higher the star rating, the higher the score required, with a maximum score of 90 or more needed to attain the six stars. As previously stated, the results of the MAMRD star rating will be presented in other publications.

This study was conducted in two main sessions of FGD with three different village categories: (1) aqua-tourism and smart community-based village (Kampung Padang Rumbia, Pekan, Pahang), (2) fishing-based

Table 2
Star rating to pursue in the MAMRD index

Star Rating	*MAMRD Rating Range (%)
☆☆☆☆☆☆	>90 - 100
☆☆☆☆☆	>75 - <90
☆☆☆☆	>60 - <75
☆☆☆	>45 - <60
☆☆	>30 - <45
☆	>15 - <30
	>0 - <15

Note. *This star rating is adopted from the Malaysia Research Assessment Instrument (MyRA), an established rating index in Malaysia
Source: Shamsir (2021)

village (Kampung Tepi Sungai, Sungai Muda, Kedah), and (3) agriculture-based village (Kampung Alur Gandak, Alor Setar, Kedah). The selection of these three villages was based on their profiles and specialised economic activity, which have the potential to embrace smart technologies. It helps validate the MAMRD's performance criteria and adds other essential criteria to suit their rural activities or solutions to problems. For the three case studies, the representatives of the *MPKK*, rural development agencies, and local institutions (ranging from 7 to 15 participants) were initially identified and invited through official letters, e-mails or WhatsApp prior to the FGD sessions.

In the FGD sessions, they were led by two researchers, one chairing the discussion and another taking notes. In the initial session, the MAMRD and the objectives of the FGD were explained. The researchers also provided details on the actions required from the participants. For the second session, the researchers assisted the participants in completing the MAMRD criteria checklist, a structured assessment form through FGD. The participants' inputs were crucial in validating and informing the availability of each criterion within their respective villages. Rashid (2020) goes into detail on this.

RESULTS AND DISCUSSION

Formulation of a Modern Rural Model

A Modern Rural Model—a translation from the FMRD—is the foundation for the MAMRD. Its goal is to synergise for rural change by improving local asset-based economies, infrastructures, social facilities, and smart technology practices. As a directive vision, the model of modern rural development is constructed based on a three dimension-objectives measure that translates the smart village and the resilient and liveable rural concepts into a new approach to modern rural development (see Figure 4).

Resilient criteria (D1) highlight the readiness and preparedness of rural areas and their communities in various aspects, such as physical, social, and economic, to effectively handle and recover from unforeseen scenarios like disasters. Liveable criteria (D2) focus on preserving rural characters, attractiveness, comfort, adequate infrastructure and support, and economic opportunities as desired by all people, including entrepreneurs, investors, and urban residents. On the other hand, smart criteria (D3) emphasise rural creative and innovative technology practices (intelligent rural supports), enabling solutions that enhance livelihoods. These three dimensions are crucial pillars in promoting and materialising the modern rural development approach, as depicted by their functions in Figure 5. The model is also built to cater to the four characteristics of rural areas as proposed in the *DPF Desa Negara 2030* (2017): Urban-rural (UR), rural-urban (RU), rural (R5), and mainly rural (R6).

The modern rural model plays a crucial role in synergising and boosting rural areas, and it is based on eight core principles that serve as the cornerstone for constructing the key assessment criteria and the factors contributing to the MAMRD's success as a
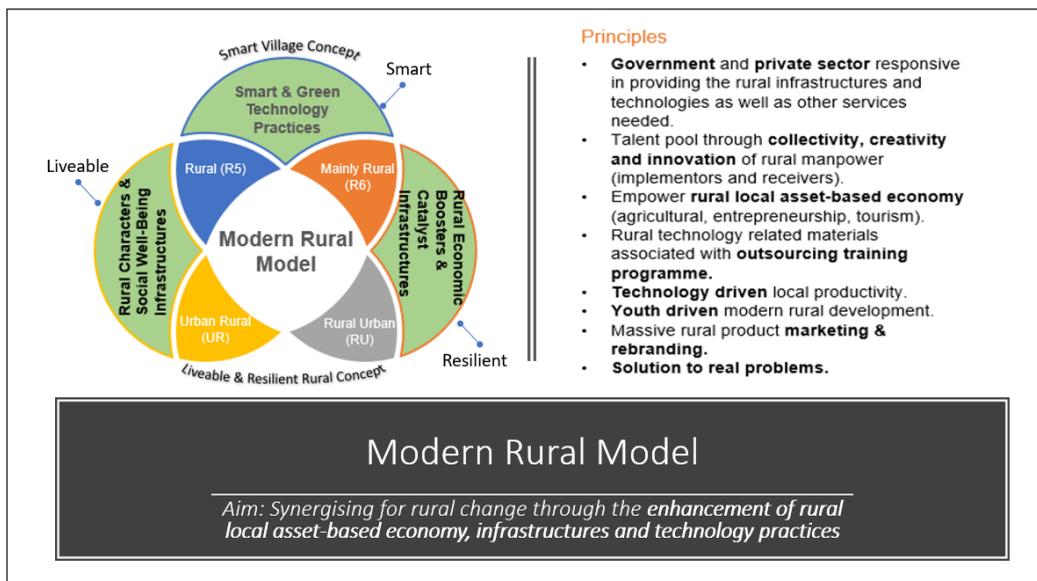
Figure 4. Modern Rural Model (*Source*: Authors' work, 2023)
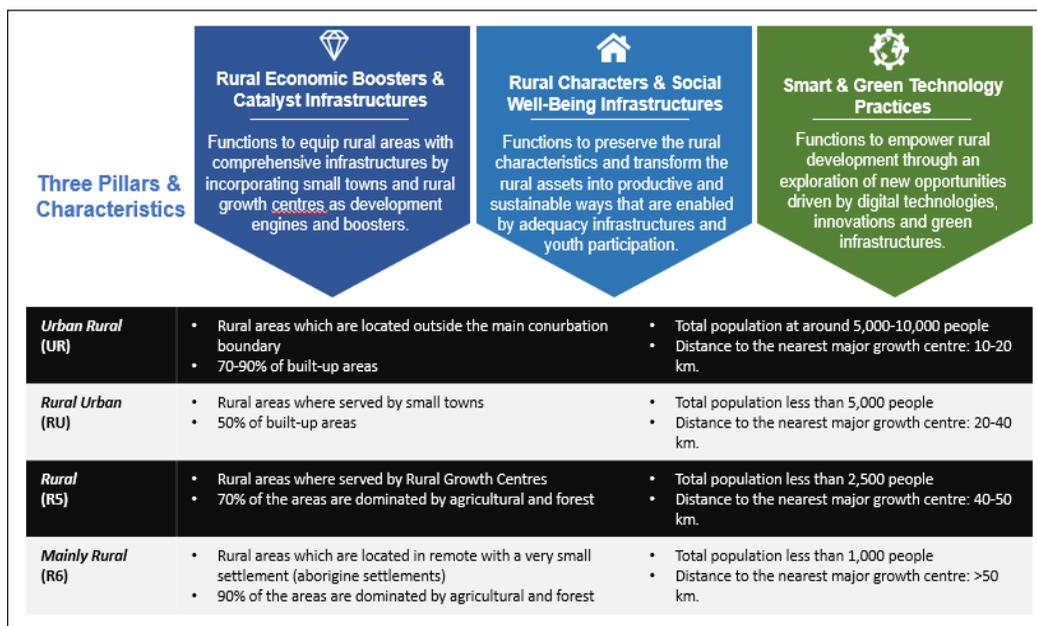
Figure 5. Functions of the three dimension-objectives measure and characteristics of rural areas for modern rural development approach (*Source*: Authors' work, 2023)

whole. It emphasises the government and the private sector to function proactively and responsively, especially with regard to providing social facilities, infrastructures, and the latest needed technologies required by rural actors, such as farmers and entrepreneurs), consequently providing higher standards of services. Moreover,

the model aims to provide a better quality of life for rural communities and fulfilling their needs. The efforts from the government and private sectors are not enough to boost the synergy and transformation in rural areas. It demands concerted energy from the community as the driven actors in exploring and utilising rural assets that can enhance their value and economic chain. In other words, they are the key players in the rural progress towards modern rural development. It includes the community leaders and local champions who are regarded as catalysts for planning, monitoring, and empowering the rural community's living conditions (Rami et al., 2021). Technology and courses related to technology, as well as their practices, must be parallel or suited to the local context requirements (Zavratnik et al., 2018). It ensures that technology is effectively used to build a knowledge-based society, consequently driving local productivity and transformation. Training courses on drone handling skills for agricultural purposes is one of the many examples of how technology can be used and benefited. Using drones in agriculture is essential today because the technology is more cost-effective in terms of time, human resources, and finances. It would reduce the need for human resources in tasks such as fertiliser and pesticide operations, where human resources become an obstacle in rural areas (Rashid et al., 2019). Using drones is also more effective because they can cover a vast agricultural area and be introduced concisely. This situation can attract financiers, investors, or private

sectors to contribute to the sustainability of rural areas through a concerted idea, energy, and capital to develop the rural area (*DPLB*, 2018) as these financiers, investors, or private sectors are profit-oriented.

The younger generation is an asset for rural areas, and harnessing their potential optimally and wisely is essential. Some are educated, motivated, and have a strong interest and energy to contribute to the rural transformation process. Furthermore, the youth are more adept at using technology, particularly for gaining advantages in various job sectors, like agriculture, than the older generation (Marescotti et al., 2021). Given their familiarity with technology, it is vital to introduce them to skills focused on using technology in agriculture, fisheries, and businesses, ultimately leading to increased local productivity. Higher productivity, in turn, increases the marketing potential of rural products and widens the opportunities for rebranding, ultimately translating into higher incomes.

MAMRD's Performance Criteria: Validated and Finalised Version

After analysing and interpreting the MAMRD survey form for the three case studies, it was determined that only 13 criteria are needed for further refinement, with one criterion being removed and two new criteria added to comply with the local context or practices in the case studies. Finally, the MAMRD will be assessed based on a three-dimensional objective measure comprising 13 criteria groups and 141 performance criteria (Table 3). It differs

Table 3

The finalised MAMRD's performance criteria for assessing modern rural index

Dimensions (With Weightage)	Criteria Groups (With Weightage)	Criteria	
D1: Rural Economic Boosters and Catalyst Infrastructures (W = 0.54)	CG1-D1: Economic and Rural Services Centre (Town) **(W = 0.50)**	**Economic Development**	
		1	Mini market
		2	Retail: Food and beverage
		3	Retail: Home appliances
		4	Retail: Vehicle equipment
		5	Retail: Agricultural equipment
		6	Souvenir shop
		7	Market/stall/bazaar
		8	Farmers market/night market/day market
		9	Restaurants/other food outlets*
		10	Food court
		11	Small-medium business
		12	Agricultural product collection centre
		13	Petrol station
		14	Insurance agent/company*
		15	Hotel/budget hotel/guest house
		Infrastructure Facilities	
		16	Road network
		17	Power and water supply
		18	Telecommunication and ICT/Internet Service Provider (ISP)*
		19	Bus station/terminal
		20	Bus stop
		21	Railway station
		22	Ferry/boat terminal
		Service Centre	
		23	Secondary school
		24	Primary school
		25	Kindergarten
		26	Mosque
		27	*Surau*
		28	Church
		29	Hindu temple
		30	Buddhist temple
		31	Cemetery
		32	Health clinic
		33	Rural clinic
		34	Police station
		35	Fire station

Table 3 *(continue)*

Dimensions (With Weightage)	Criteria Groups (With Weightage)	Criteria	
		36	Multipurpose hall
		37	Public hall
		38	Community working hall (*Balai raya*)
		39	Rural library
		40	Local park
		41	Neighbourhood park
		42	Playground
		43	Bank
		44	Registered bank agent
		45	Mini Rural Trade Centre (RTC)
		Human Development	
		46	Local centre for business and consultation services
		47	Entrepreneurship skills training centre
		48	Community Rehabilitation Programme (CRP)
	CG2-D1: Rural Growth Centre (RGC) **(W = 0.25)**	**Economic Development**	
		1	Agricultural (including fishery) product collection centre*
		2	Small-scale retail
		3	Shop that supplies modern agriculture (including fishing and livestock) equipment and technology (including technical services)*
		4	A workshop that provides maintenance and repair services for agricultural equipment
		5	Hardware shop
		Infrastructure Facilities	
		6	Road network
		7	Power and water supply
		8	Telecommunication, high-speed broadband, and other ICT services
		9	Public transport terminal
		Service Centre	
		10	Community and recreational facilities
		11	Mobile Community Transformation Centre (CTC)
		12	Registered bank agent
		Human Development	
		13	Community Rehabilitation Programme (CRP)
		14	Elderly activity centre
		15	Youth and innovation centre (multipurpose skills training centre, including for disadvantaged individuals and single mothers)*

Table 3 *(continue)*

Dimensions (With Weightage)	Criteria Groups (With Weightage)	Criteria	
	CG3-D1: Rural economic cluster (agricultural, entrepreneurship, and tourism) (W = 0.25)	1	Tourist information unit*
		2	Homestays operated by the community through MPKK
		3	Cheap accommodation/ budget motel
		4	Traditional and casual food premise concept
		5	Permanent Food Production Farm (TKPM)
		6	Rural trade and retail
		7	Eco-friendly farm-based rural tourism attraction**
		8	High-speed broadband facilities (dedicated space) for retailers and purchasers*
D2: Rural Characters and Social Well-Being Infrastructures (W = 0.30)	CG1-D2: Rural spatial characters and heritage (W = 0.064)	1	Rural boundary (delineation of village assets)*
		2	Rural landmark (gateway, statue, and welcoming signage)
		3	Excellent rural asset development award
		4	Agricultural areas maintained as a rural buffer zone*
		5	New development of low-density housing (detached) suits with rural characters and B40
		6	Adaptive reuse or restoration of old house
		7	Preservation of traditional Malay houses (or maintain the traditional archi-style)
		8	Individual registration as National Heritage Living Person (WAKOH)
	CG2-D2: Transportation networks of rural-town-city, and rural accessibility (W = 0.275)	1	The bus stop for stage buses (located within 500 meters of the village)
		2	Shuttle train station
		3	Water transport jetty
		4	Rural paratransit stops (minibus/van)
		5	MyCar, Grab, and other e-hailing service providers
		6	Paved main entrance/access
		7	Paved rural internal road
	CG3-D2: Efficient infrastructure (W = 0.402)	1	Continuous, adequate, and clean water supply*
		2	Extensive power supply
		3	1Malaysia Internet Centre (PI1M)
		4	High-speed broadband
		5	Fibre optic (fixed bandwidth) coverage
		6	Sanitary landfill
		7	Recycling centres are operated either by the government or in partnership with the local community
		8	Septic tank system

Table 3 *(continue)*

Dimensions (With Weightage)	Criteria Groups (With Weightage)		Criteria
	CG4-D2: Internal village amenities **(W = 0.098)**	1	Mobile facilities (clinic and library)
		2	Community hall/rural community centre
		3	*Surau*
		5	Football field/public field/recreational park/ sports facility separate from school*
		6	Healthcare centres (for the elderly, disabled people, and neglected mothers)
		7	Temporary shelter/transit service for disasters with a dedicated command centre separate from schools
	CG5-D2: Rural governance (MPKK) and database **(W = 0.161)**	1	MPKK working room
		2	Rural community co-operative centre
		3	Rural village database managed by a dedicated or a paid staff
D3: Smart and Green Technology Practices (W = 0.16)	CG1-D3: Rural agricultural, infrastructures, technologies, and innovations **(W = 0.40)**	1	Tractor
		2	Plough
		3	Harvesting machine
		4	Micro-watershed management
		5	Farmers Information System (FIS)/Fisheries Information System/drone technology*
		6	Drone or UAV technology (crop monitoring and pest control)
		7	Smart database for agriculture through sensors and satellite data
		8	Smart weather and irrigation system
		9	Vertical farming
		10	Vinyl greenhouse agriculture
		11	Smart dairy through smart devices (livestock)
		12	Production of high-demand agricultural products (kenaf, vanilla, basmati rice, *musang king*, stingless bee/*lebah kelulut*, and *burung walit*)
	CG2-D3: Rural entrepreneurship, technologies, and innovations **(W = 0.29)**	1	Agro-industry basic facilities (such as incubator centres for up to district scale)
		2	Community biogas plant for entrepreneurship activities
		3	Market analysis tools/software
		4	Village community radio
		5	Telecommunication and video conferencing
		6	ICT-related materials and outsourcing training
		7	Mentor-mentee training programme or rural icon in business

Table 3 *(continue)*

Dimensions (With Weightage)	Criteria Groups (With Weightage)		Criteria
	CG3-D3: Rural marketing and e-commerce **(W = 0.13)**	1	Fresh fruit stall (GBBS)
		2	Agrobazaar
		3	KShoppe
		4	Training centre and e-commerce services (equipped with high-speed broadband)
	CG4-D3: Village smart and green technology practices **(W = 0.12)**	1	Rainwater harvesting
		2	Renewable energy (through solar rooftop PV, solar micro grid, micro-hydroelectric, and solar farming)
		3	Generate energy through biogas digestion
		4	Micro-hydroelectric power for multipurpose uses**
		5	Solar cookers
		6	LEDs
		7	Low-energy motors
		8	Flood risk alarming through a smartphone
		9	Biochar for transforming garden waste into organic fertilisers—waste-to-wealth
	CG5-D3: Community-IoT-based smart technology practices **(W = 0.06)**	1	Smart healthcare facilities/healthcare mobile apps
		2	Waste monitoring and management system through wireless sensors monitor
		3	Smart education (through videos, smart classroom, and fun-toy library)
		4	CCTV cameras/smart surveillance system
		5	Goods and services delivery system via mobile apps

Note. *The criteria which have gone through the refinement; and **the additional new criteria of the MAMRD (*Source*: Authors' work, 2023)

from the FMRD by adding only two criteria to the original 139 criteria.

The MAMRD's performance criteria have compounded all the essential criteria of *DPF Desa Negara 2030* (2017) and *DPLB 2030* (2018) while integrating them with smart technologies and ICT applications to modernise infrastructures and village/community practices within rural areas. This kind of approach reflects a holistic approach to addressing rural problems, harnessing the potential of new technology practices (Morris et al., 2022; Zavratnik et al., 2018), and catching up to rural economies and overall productivity (United Nations, 2021). However, the successful application largely depends on technological diffusion, socio-cultural aspects, and the readiness of rural actors, particularly their educational

levels and skills (Curry et al., 2021; Morris et al., 2022; Salemink et al., 2017). Developing telecommunications and information infrastructures (ICT and digital infrastructures), especially high-speed broadband, also plays a crucial role in expanding coverage and capacity within and around rural areas. Improving the quality of these infrastructures and adopting new technologies in agriculture, business, and other rural sectors contribute to creating an attractive and smart environment. Moreover, it empowers individuals, especially the younger generation, to actively participate in modern-smart rural life and adapt to new solutions, fostering creative solutions, innovation, and diversity within their villages and communities.

The 141 MAMRD criteria have been considered in shaping the future niches and the pathways of rural development in Malaysia, as well as globally. They can be summarised but not limited to:

- Future rural development patterns are becoming more complex, necessitating a more strategic vertical and horizontal development approach.
- The involvement and participation of all parties, including rural actors, private sectors, and government interventions, are crucial for promoting proactive and well-balanced dimensions of rural development.
- The integration of bottom-up and community engagement (involving youths, entrepreneurs, pensioners, and private sectors) alongside top-down approaches is essential to effectively implement current policies and fulfil the aspirations and capabilities of the rural community.
- To account for the characteristics and uniqueness of the rural community, "there is no one-size-fit-all solution."
- The rural community needs to undergo a paradigm shift, moving away from overdependence on the government and instead fostering their creativity, innovation, and proactive approach to developing their villages.
- Rural growth centres (RGC)/ rural community centres (RCC) become the foundation for the development and liveability of rural areas by providing goods and services needed by the community, especially daily and weekly goods.
- The adoption and application of technology and ICT (IoT) in rural activities can potentially elevate rural areas in terms of information, knowledge, business, marketing, and accessibility. Moreover, it enhances resilience in the face of risks, such as the ongoing COVID-19 endemic, which continues to impact Malaysia and the global community. It is one of the strategic principles resetting rural development for the 21st century outlined by the United Nations (2021).

Therefore, the MAMRD plays a timely role in empowering rural areas for transformation, focusing on infrastructure, services, and the rural community. It is worth noting that existing literature highlights the need for higher adoption levels of advanced technologies, which progress slowly among small-scale farmers in developing countries (Curry et al., 2021). Hence, this paper contributes significantly to promoting the widespread application of digital and ICT technologies to realise the modern rural approach in Malaysia.

Dissemination of the MAMRD Exercise

As mentioned earlier, the model of modern rural is derived from four of the six typologies (characteristics) outlined in the *DPF Desa Negara 2030* (2017), namely UR, RU, R5, and R6. The remaining two typologies, U1 and U2, are exempted as their primary land use is no longer agriculture, and their communities primarily consist of service sector workers. Therefore, implementing the MAMRD is suggested for the four mentioned typologies (see Figure 4), which are based on agriculture, rural assets, or rural resources comprising entrepreneurship, SMEs, tourism, and other sectors.

The third case study, Kampung Alor Gandak, is categorised as U1 as the community is situated within the perimeter of Alor Setar city. Nevertheless, it is still selected as a case study because its land use is still dominated by agriculture, and most residents are farmers. This situation indicates that the MAMRD is suitable for all categories of villages. However, the main focus remains on the villages in the four categories stated earlier. The Indigenous people villages located on the outskirts are also deemed suitable for implementing the MAMRD because they are categorised as sustainable, given equipped facilities and the community's adherence to conservative or environmental-friendly practices. However, it should be noted that Indigenous villages in remote areas face challenges in meeting all 141 criteria listed by the MAMRD. In such cases, certain criteria, particularly those related to D2 and D3, can be considered for measurement to attain the MAMRD index while exempting D1.

Based on these considerations, the implementation of the MAMRD is summarised as the following village characteristics:

- All villages from the four typologies in the model (Figure 4) are eligible for assessment using the MAMRD.
- Villages from typologies U1 and U2 are also suitable for evaluation if agriculture remains the dominant land use.
- Although Indigenous villages are included in the MAMRD assessment, the focus is not on their index or star rating. Instead, the evaluation focuses on measuring the level of technology practices and the adequacy of provided facilities that align with the aspirations of modern rural characteristics. In this regard, it is crucial to establish and enforce

the RGC/RCC for the sake of the community's welfare, services, amenities, and prosperity.

CONCLUSION

This research mainly focuses on formulating the MAMRD as an innovative approach to undertaking existing government policies such as the *DPF Desa Negara 2030* (2017) and *DPLB 2030* (2018). The MAMRD offers a measurement tool for rural stakeholders to monitor and realise the aspirations of rural development agendas. It would provide additional impacts and efforts in transforming rural change towards a modern approach. It is important to highlight that this research has discovered a new approach to rural development by synthesising and adopting the concepts of resilience, liveability, and smartness, reflecting Malaysia's modern rural development context. The MAMRD deserves endorsement as it pioneers the integration of innovative modern technology practices in rural areas. Furthermore, applying weights to the dimensions and criteria groups ensures transparency and facilitates meaningful comparisons across regions, states, districts, and other levels.

The modern-smart approach in agriculture has been one of the main elements of the Twelfth Malaysia Plan (Economic Planning Unit, 2021b). This approach includes using innovative technologies and information and knowledge-based technology in rural development. Technologies such as sensor networks, mobile applications, and other advancements contribute tremendously to the transition towards a knowledge-driven society where constraints of time and place are no longer significant. As a result, new models for rural businesses, agriculture, and industries will emerge to replace the existing processes with new roles, mechanisms, and technologies. Indeed, the utilisation of these technologies will not only increase productivity and decrease the dependency of the workforce but also attract younger generations and investors to participate in rural activities and actively contribute to their enhancement.

The MAMRD, while primarily focused on providing infrastructures, facilities, services, and rural technology practices, also enforces investment in rural communities. It emphasises the opportunities available in the social, economic, and environmental dimensions, as well as the development of smart individuals to synergise and boost the rural change that will lead to modern rural outcomes. Therefore, using the MAMRD for rural assessments is a reliable approach to optimise rural performance towards future niches and pave the way for future advancements in rural development in Malaysia. As mentioned earlier, the best practice to obtain data for the MAMRD assessment is through an FGD, and the MPKK, in particular, are the key informants to provide the needed information. More importantly, they will receive first-hand output (direct feedback) regarding the advancement of their village towards a modern approach, as indicated by a rating star.

Implications for Practices

This study has achieved the aim pertaining to the validated dimensions, criteria groups, and 141 MAMRD's performance criteria for the rural development measurement index towards a modern approach. The MAMRD measure offers a significant star rating index that enables the rural stakeholders to monitor rural development transformation by relying on technology practices and enhanced solutions. More importantly, it has been tailored to the national rural development policies, such as the *DPF Desa Negara 2030* (2017) and *DPLB 2030* (2018), so it has its impacts on rural development in Malaysia. It is important to emphasise that the policies above have progressed to the revision stage after a duration exceeding five years in the implementation phases. Consequently, the MAMRD's dimension, specifically D3 (Smart and Green Technology Practices), is poised to provide pertinent guidelines. This dimension focuses on enhancing smart village practices in Malaysia, and it is evident that neither or less of the policies were designed considering this aspect. Malaysia is igniting its path towards increased digitalisation and embracing the Fourth Industrial Revolution (4IR) in 2021, two to three years after the policies were implemented.

So, the MAMRD comes timely and impacts future rural transformation in Malaysia, either at top-down (policymakers) or bottom-up (rural actors) levels. Anticipated outcomes include enhanced benefits for rural communities, such as connectedness, increased economic chances, increased readiness for a disaster, and improved healthcare facilities by implementing the MAMRD index measurement.

Limitations and Recommendations

The above summary findings and implications are subjected to the methods and assumptions applied by the study. Therefore, the generalisation of the study findings in different contexts or cases should be cautiously considered, considering the three limitations. First, the MAMRD was developed using content analyses and expert opinion surveys. For the expert opinion, five local/international experts on rural development were invited to validate the criteria used in the MAMRD in only a single-round survey. During the survey, criteria validation was made using Likert scale measurement. Despite that, experts were encouraged to suggest new criteria and confirm the dimensions and the criteria groups. Although the number of experts and the round involved in the validation process is relatively limited, the study further validated the criteria via the case studies approach – the participants' (rural actors) opinion in the FGD sessions for further validation. Therefore, a further round of opinion surveys is not necessary. Future research may benefit from an enlarged number of experts in the expert opinion survey and an additional round of the expert survey.

Second, the important findings are based on the three village categories: aqua-tourism and smart community-based village, fishing-based village, and agriculture-based village.

The selection of the village categories was made considering agricultural activities that are mainly practised in Malaysia's rural areas. Considering all types of villages in all grids as stipulated in the *DPF Desa Negara 2030* (2017) may yield a multi-dimensional insight into the rural activities and various ground conditions and benefit future researchers.

The third limitation is that the study only focused on rural infrastructures, facilities, services and technology practices as the key assessment criteria for facilitating modern rural development. Instead, the *DPF Desa Negara 2030* (2017), the *DPLB 2030* (2018), and other policies have comprised the strategies for rural development, including infrastructure, social, economy, environment, land use, and many more. Including these elements in future research may yield a better comprehension of the issues and challenges faced by agricultural-entrepreneurial-tourism-based rural residents.

It is necessary to note that all the above limitations do not compromise the quality of the study outcomes since it was conducted with vigorous methodology or processes and valid and reliable data. Moreover, potential areas for future research could be explored as a limitation of this research. As suggested in the following:

- Participatory-based research will implement the technology practices proposed in the MAMRD to achieve a society 5.0 @ rural in Malaysia. It involves rural actors in decision-making, adapts technologies to rural contexts, and

aims for societal advancement through technological innovation and community participation.

- Profiling the knowledge-based society and technology practices among rural communities in Malaysia. This research involves an inventory of rural champions and technology practices. It will explore how they engage with technology to promote socio-economic development and enhance their way of life and prosperity.

- Rural actors' behaviours framework for synergising rural transformation towards a modern approach. This research involves developing a framework of rural actors' behaviours to enhance rural transformation, aligning with the MAMRD's dimensions and index measurement.

In addition to that, it recommended a more frequent site visit to explore multiple ways of looking at the issues, challenges, and ideas in rural areas extensively, including understanding the forces and changes of the past, recording narratives from vastly experienced local residents and entrepreneurs to calibrate the reality of their lives, and gaining real insights into how things progress in rural areas.

ACKNOWLEDGEMENTS

It is initially part of a project funded by the National Real Property Research Coordinator (NAPREC), National Institute of Valuation (INSPEN), Valuation & Property Services Department (JPPH), Ministry of Finance, Malaysia.

REFERENCES

Alabdali, S. A., Pileggi, S. F., & Cetindamar, D. (2023). Influential factors, enablers, and barriers to adopting smart technology in rural regions: A literature review. *Sustainability*, *15*(10), Article 7908. https://doi.org/10.3390/su15107908

Bruton, M. J. (2007). *Malaysia: The planning of a nation*. Persada.

Curry, G. N., Nake, S., Koczberski, G., Oswald, M., Rafflegeau, S., Lummani, J., Peter, E., & Nailina, R. (2021). Disruptive innovation in agriculture: Socio-cultural factors in technology adoption in the developing world. *Journal of Rural Studies*, *88*, 422–431. https://doi.org/10.1016/j.jrurstud.2021.07.022

Dasar Pembangunan Luar Bandar. (2018). *Dasar Pembangunan Luar Bandar: Luar bandar sejahtera* [Rural Development Policy: Prosperous rural areas] https://www.rurallink.gov.my/dasar-pembangunan-luar-bandar-dplb/

Dasar Perancangan Fizikal. (2017). *Dasar Perancangan Fizikal (DPF) Desa Negara 2030* [Rural Development Policy 2030]. https://myplan.planmalaysia.gov.my/portal-main/publication-details?id=2

Economic Planning Unit. (2021a). *National Fourth Industrial Revolution (4IR) policy*. Prime Minister's Department, Malaysia. https://www.ekonomi.gov.my/sites/default/files/2021-07/National-4IR-Policy.pdf

Economic Planning Unit. (2021b). *Twelfth Malaysia Plan 2021-2025: A prosperous, inclusive, sustainable Malaysia.* Prime Minister's Department, Malaysia. https://rmke12.ekonomi.gov.my/file/download/2021092855_executive_summary.pdf?path=fileUpload/2021/09/2021092855_executive_summary.pdf&name=Executive%20Summary.pdf

Kusumo, R. A. B., Kurnia, G., Setiawan, I., & Tirtosudarmo, R. (2023). Migration and farmer household livelihood strategies: Factors influencing the decision to migrate. *Pertanika Journal of Social Sciences & Humanities*, *31*(1), 57-79. https://doi.org/10.47836/pjssh.31.1.04

Malczewski, J. (1999). *GIS and multi-criteria decision analysis*. Wiley.

Marescotti, M. E., Demartini, E., Filippini, R., & Gaviglio, A. (2021). Smart farming in mountain areas: Investigating livestock farmers' technophobia and technophilia and their perception of innovation. *Journal of Rural Studies, 86*, 463–472. https://doi.org/10.1016/j.jrurstud.2021.07.015

Ministry of Rural and Regional Planning Development. (2010). *Pelan Induk Pembangunan Luar Bandar [Rural Development Master Plan] [Powerpoint Slides].* SlideShare. https://www.slideshare.net/slideshow/pelan-induk-pembangunan-luar-bandar-kklw/11504488

Morris, J., Morris, W., & Bowen, R. (2022). Implications of the digital divide on rural SME resilience. *Journal of Rural Studies*, *89*, 369–377. https://doi.org/10.1016/j.jrurstud.2022.01.005

Ngah, I., Preston, D., & Asman, N. (2010). *Current planning priorities in rural villages in Malaysia learning from the new Malaysian Village Action Plans* [Paper presentation]. The ISDA 2010, Montpellier, France. https://hal.science/hal-00525961

Rami, A. A. M, Aziz, M. F. A., Nasharudin, N. A. M., & Rasdi, R. M. (2021). Characteristics and success factors of rural community leadership in Malaysia: A focus group analysis. *Pertanika Journal of Social Sciences & Humanities*,

29(3), 1591–1609. https://doi.org/10.47836/pjssh.29.3.07

Rashid, M. F., Ngah, I., & Misnan, S. H. (2019). Revitalizing rural areas in Malaysia: A framework for economic revitalization. *IOP Conference Series: Earth and Environmental Science, 385*, Article 012004. https://doi.org/10.1088/1755-1315/385/1/012004

Rashid, M. F. A. (2020). *FGD and its application in a modern rural index measurement (MAMRD)* [PowerPoint Slides]. ResearchGate. https://www.researchgate.net/publication/346667201_FGD_and_its_application_in_a_modern_rural_index_measurement_MAMRD

Rashid, M. F. A., Mokhtar, S., Kamaruddin, S. M., Kamar, M. A. A., Abdullah, S., & Ali, M. A. F. (2023). A GIS-based Sustainability and Highest-Best Use (SHBU) for Felda lands development planning decision-making. *Planning Malaysia*, *21*(1), 444-458. https://doi.org/10.21837/pm.v21i25.1249

Rashid, M. F. A., Muhamad, A. K., Rashid, K., Ahmad, A.L., & Azman, M. A. A. (2021). Formulation of a Malaysia modern rural development framework: Synergising rural for change. *Planning Malaysia*, *19*(16). 14–26. https://doi.org/10.21837/pm.v21i25.1249

Saaty, T. L. (1980). *The analytic hierarchy processes*. McGraw-Hill.

Saaty, T. L., & Kearns, K. P. (1985). *Analytical planning: The organization of systems*. Pergamon Press.

Salemink, K., Strijker, D., & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. *Journal of Rural Studies, 54,* 360-371. https://doi.org/10.1016/j.jrurstud.2015.09.001

Shamsir, M. S. (2021). *Malaysia Research Assessment (MyRA): Innovating the measurement of research performance* [PowerPoint Slides]. MyRA Workshop. https://www.iium.edu.my/media/78354/MyRA%20Engagement_V1.6.-2021.pdf

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development.* https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

United Nations. (2021). *World social report 2021: Reconsidering rural development*. The United Nations Department of Economic and Social Affairs (UN DESA). https://desapublications.un.org/publications/world-social-report-2021-reconsidering-rural-development

Zavratnik, V., Kos, A., & Duh, E. S. (2018). Smart villages: Comprehensive review of initiatives and practices. *Sustainability,10,* Article 2559. https://doi.org/10.3390/su10072559

Development of a Real-time Digital Learning Platform for Diagnosing Degrees of Mathematical Competency

Sineenart Phaengkham[1], Putcharee Junpeng[1]* and Keow Ngang Tang[2]

[1]*Faculty of Education, Khon Kaen University, 40002 Khon Kaen, Thailand*
[2]*Postgraduate Program in Education, Faculty of Business, Hospitality and Humanities, Nilai University, 71800 Nilai, Negeri Sembilan, Malaysia*

ABSTRACT

The research aims to develop and verify a real-time digital learning platform for diagnosing the degree of mathematical competency of seventh-grade students. A total of 1,559 students from four regions of Thailand and six experts participated in this research. The researchers employed a design-based approach and the Multidimensional Coefficient Multinomial Logit Model to evaluate the effectiveness of the real-time digital learning platform. The measurement tool consists of two aspects, namely mathematical measures and the construction of knowledge dimensions, with 58 items to simulate students' responses to the three respective content constituents: number and algebra, measurement and geometry, and statistics and probability. The findings showed that the multidimensional model offered a significantly better statistical fit in measurement, geometry, statistics, and probability, while numbers and algebra fit better in the unidimensional model. There were positive, strong, and significant correlations between the dimensions. The findings indicated that all items fit the OUTFIT MNSQ and INFIT MNSQ, which are between 0.75-1.33. The findings revealed that the internal structure when it came to evaluating the degrees of students' mathematical competency using the Wright map and the Multidimensional Test Response Model conformed with the quality of the digital learning platform in terms of its usefulness, accuracy, and feasibility. This real-time digital learning platform for diagnosing the degree of mathematical competency can be accessed from anywhere with an internet connection, making education more accessible to a wider audience, including individuals in remote areas or those with physical disabilities.

Keywords: Content constituent, construction of knowledge, degrees of mathematical competency, mathematical measures, real-time digital learning platform

INTRODUCTION

The rapid development of technology has indeed prompted the integration of various technological tools into teaching and learning processes. Digital diagnostic tools are crucial in assessing and enhancing the learning experience, particularly in the post COVID-19 pandemic (Buakhamphu, et al., 2024). According to Thisopha et al. (2023), digital diagnostic tools can provide real-time assessments of students' understanding of the material. In the post-COVID world, education should make free and open-source technologies available to teachers and students. Open educational resources and open-access digital tools must be supported. Education cannot thrive with ready-made content built outside the pedagogical space and human relationships between teachers and students (International Commission on the Futures of Education, 2020). The COVID-19 pandemic has transformed the education system and created enormous disruption (Maisa et al., 2021) that needs the development of a real-time digital learning platform to diagnose students' mathematical competency. In addition, Kesorn et al. (2020) emphasized the importance of mathematics in providing practical knowledge and stimulating students' learning. Kantahan et al. (2020) highlighted that an interactive and creative digital learning platform focused on individual needs and competencies is required in this new era. In this study, the theoretical framework applied a holistic approach that combines assessment design, data analysis, feedback mechanisms, and ethical considerations to ensure that the platform serves its educational purpose and provides reliable and valid measurements of student outcomes.

Previously, the mathematics teacher was the primary source of knowledge, and learning was restricted to the classroom. Even though the role of the mathematics teacher remains undiminished in the current educational setting, students now have several resources or a wealth of opportunities literally at their fingertips, allowing them to self-direct and take control of their learning and assessment processes (Harman, 2021). According to Harman, all digital platforms, with their various functions and features, can support any type of learning, namely, in the form of online learning, classroom learning, and blended learning, hence fostering overall academic growth in students. Harman (2021) also explained the advantages of using digital learning platforms: easy access to information, up-to-date content, advanced reporting, multimedia learning, new communication channels, 24-hour access to learning resources, and self-directed study.

Phaniew et al. (2021) defined mathematical competency as a student's ability to engage in exploration, critical thinking, and logical reasoning. Mathematical competency goes beyond rote memorization and emphasizes the importance of understanding and reasoning through mathematical concepts and problems. The essential constituents of mathematical competency are numbers and algebra, measurement and geometry,

and statistics and probability. Therefore, the awareness of these content constituents and how they interact make up the basic stages of precision, stages, and problem-solving play a vital role in developing students' mathematical competency (Yulian & Wahyudin, 2018). According to these researchers, the role of the mathematics teacher is to allow students to acquire significant and evocative mathematical concepts and procedures, using interactive mathematics classes that permit every student to engage in stimulating, thought-provoking practices when it comes to learning mathematics. Students need to develop several mathematical competencies to successfully learn mathematics. In addition, Milgram (2007) also viewed mathematical competence as having five dimensions: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition. These mathematical competencies should be developed in an integrated and balanced way for students to learn mathematics (Kilpatrick et al., 2001).

The definition of mathematical competence is further supported by Junpeng et al. (2018), who identified mathematical competencies as the procedure of employing mathematical understanding. Junpeng et al. (2018) revealed that numerous attempts to pilot and review students' answers with regard to open-response questions in order to assess their capabilities are important to diagnose their degrees of mathematical competency. Therefore, Junpeng et al. (2018) identified mathematical measures

(MME) as one of the mathematical competence dimensions. They state that MME is the ability to execute mathematical practices elastically, precisely, competently, and correctly. On the other hand, Biggs and Collis (1982) defined the construction of knowledge (CKL) as it appears in the SOLO taxonomy, as a mathematical competence dimension that recognizes, illustrates or clarifies the degree of knowledge on the part of the student to establish the degree of student understanding outcomes. In other words, the CKL dimension is characterized as understanding mathematical concepts to solve mathematical problems using the students' capacities for logical thinking and justification (Junpeng et al., 2018).

Kepner and Huinker (2012) revealed that mathematical practices, as expectations of student behaviors, depend on essential methods and skills found in mathematics education and have long-lasting consequences. Therefore, each assessment claim indicates what would be expected of a student with regard to demonstrating competencies across a range of content and practices (Kepner & Huiner, 2012). They further emphasized that the consortia's promise of assessing and reporting on multifaceted student performance entails assessment assignments and items that go beyond select-response items to construct-response items. Furthermore, Adom et al. (2020) stated that there is a constant need to measure the outcomes or the quality of responsiveness of the teaching and learning referred to as assessment in the teaching-learning environment. This essential

symbiotic process can occur before, during, and after the teaching process.

In addition, Fowler et al. (2019) suggested that mathematics teachers should structure learning activities in such a way as to not only demonstrate content but involve students in profound acquaintance with mathematical thinking by utilizing an appropriate pedagogy that intertwines the mathematical competencies, leading to the emergence of lifelong users of mathematics. They recommended the use of digital tools to support the presentation of mathematical competencies by delivering active learning, as a result of which students are able to resolve mathematical problems confidently, independently, and collaboratively, thereby developing their understanding of mathematical concepts and being able to explain their reasoning. Digital learning platforms allow students to experience mathematical concepts that are hard to exhibit in actual circumstances. Through research-proven methods and the skillful incorporation of digital sources, mathematics teachers can meaningfully discover students' interpretations of mathematical competencies (Fowler et al., 2019).

The Thailand Mathematics Curriculum is established across the interface of these content constituents and the four competency constituents. As noted above, the content constituents are number and algebra, measurement and geometry, and statistics and probability. In short, the content constituents depict what is to be learned and understood by students (Thailand Ministry of Education, 2008). On

the other hand, the competency constituents of mathematics include understanding, fluency, problem-solving, and reasoning. These constituents illustrate by what means content is investigated or established; that is, it relates to the reasoning and performing of mathematics. Therefore, they offer an eloquent foundation for the development of concepts in the learning of mathematics and have been merged into the content descriptions of the three content constituents. Based on this line of reasoning, the Thailand Ministry of Education has adopted this approach to certifying the mathematical competencies that can be developed throughout the curriculum and become progressively advanced over the years of schooling (Manmai et al., 2021).

The above literature review indicates the challenges researchers face when developing and verifying a measurement model and its interpretation process in every way possible. The motivation for developing a real-time digital learning platform for diagnosing degrees of mathematical competency stems from several critical gaps in theoretical understanding and practical applications in the current educational landscape. It is undeniable that the COVID-19 pandemic has given a major boost to online education. It leads to all schools having already adopted, or being in the process of adopting, a digital learning platform to support learning (Harman, 2021). Moreover, the development of a real-time digital learning platform for diagnosing degrees of mathematical competency is motivated by the need to address theoretical gaps related to

individualized learning, dynamic assessment, and cognitive load, as well as practical gaps concerning timely intervention, resource accessibility student engagement, and data-driven decision-making in education.

Maoto et al. (2018) emphasized that mathematics teachers are having difficulties assessing the degree of students' mathematical competency. They need information and skills about what needs to be evaluated in terms of students' mathematics competencies and how to perform such evaluations related to the intended goals of the task. As a result, this research would be timely in providing an exploration of real-time mathematical competency to aid mathematics teachers in easily diagnosing their students' mathematical competencies. Following this line of reasoning, the research aimed to develop and verify a real-time digital learning platform for diagnosing the degree of mathematical competency of seventh-grade students. In this paper, the researchers planned to investigate 1,559 seventh-grade students' answers in a predictable digital assessment process to fix the criteria for measuring the degree of mathematical competency in terms of the three content constituents: number and algebra, measurement and geometry, and statistics and probability. It was followed by developing a mathematical competency measurement model using the Rasch model as a Multidimensional Random Coefficient Multinomial Logit Model (MRCMLM) to analyze the situation. Finally, the researchers evaluated the overall quality of the real-time digital

learning platform in terms of diagnosing the degree of mathematical competency on the part of seventh-grade students by using a measurement model involving three aspects: usefulness, feasibility, and accuracy, using heuristic evaluation.

MATERIALS AND METHODS

The researchers employed construct modeling to introduce a real-time digital learning platform in which the three mathematics content constituents were designed diagnostic tasks to measure the degree of mathematical competency on the part of seventh-grade students (Wilson, 2005). Moreover, a design-based research method was adopted from Adams (2005), and the MRCMLM (Adams et al., 1997) was applied to verify the quality of the real-time digital learning platform. Item Response Theory (IRT) assesses the relationship between a student's mathematical ability or trait and their responses to a set of items or questions. The primary idea behind IRT is to model the probability of a particular response to an item as a function of the student's latent trait and the item's characteristics (Linacre, 2005).

The key concepts in IRT include item parameters, latent traits, and test information. Each test item has its own set of parameters, such as difficulty and discrimination. Difficulty indicates how hard the item is, while discrimination measures the item's ability to differentiate. The unobservable trait in this study is the mathematical competency that the test aims to measure. IRT models estimate a student's

position on this latent trait. The amount of information a test provides about a student's ability level. IRT helps design tests that maximize information across a range of ability levels (Linacre, 2005).

Population and Sample

The research population consisted of all the seventh-grade secondary school students from the four regions of Thailand: northeast, central, north, and south. A total of 1,559 students were chosen as test-takers using stratified random sampling. It allowed the researchers to obtain a sample population including smart, moderate, and weak students in such a way as to best represent the entire population being studied (Hayes, 2024). According to Wright and Stone (1979), a minimum sample size of not less than 500 individuals is needed to obtain sufficient quality data for analysis when using the multidimensional test response theory. Table 1 demonstrates the distribution of the test-takers who participated in the three assessment tools in terms of the three content constituents.

Table 1
Distribution of the test-takers in the three content constituents

Region	School	Number of test-takers		
		Number and Algebra	**Measurement and Geometry**	**Statistics and probability**
Northeast	1	55	49	47
	2	72	71	72
	3	62	63	58
	4	2	3	3
	5	7	6	6
	6	64	66	65
	7	9	7	7
	8	-	2	2
Central	9	52	54	48
	10	51	46	50
North	11	4	2	2
	12	12	11	8
	13	3	3	4
	14	11	12	9
	15	1	1	2
	16	5	5	5
	17	5	5	5
	18	2	3	2
	19	7	5	4
	20	4	4	4
South	21	42	40	38
	22	12	14	13
	23	-	56	53
Total		524	528	507

Source: Authors' work

Research Procedure

Researchers utilized the construct modeling approach Wilson (2005) proposed to design the mathematical competency measurement model. In addition, design-based research was employed as the research design (Vongvanich, 2020) by adapting four building blocks as the sections of the measurement scheme (Wilson & Sloane, 2000). The four building blocks were (1) a progressive standpoint of students' understanding, (2) the association between teaching and measurement, (3) the teacher administering classroom teaching and measurement, and (4) the creation of high-quality evidence. On top of that, the researchers used software and analysis methods, including machine learning algorithms like regression, classification, clustering, and natural mathematical language processing techniques.

In the first phase, the researchers conducted numerous inclusive meetings with mathematics teachers and experts by referring to the basic education core curriculum 2008 (Ministry of Education, 2017) to create three assessment tools. These diagnostic tools were a multiple-choice test with four options in an online testing system entitled "e-MAT-Testing." The e-MAT-Testing approach assessed three main content constituents: (1) number and algebra (21 items), (2) measurement and geometry (20 items), and (3) statistics and probability (17 items). The internal structure of the e-MAT-Testing approach was then formed and evaluated using model fit and interdimensional correlation analysis.

The test results from the first phase were used in the second phase to create an innovative prototype, a real-time digital learning platform for diagnosing seventh-grade students' mathematical competency levels. In the third phase, the researchers utilized MRCMLM (Adams et al., 1997), which is a type of Rasch Model (Rasch, 1960), to generate quality evidence of the value of the real-time digital learning platform for diagnosing degrees of mathematical competency on the part of seventh-grade students.

The researchers analyzed the item fit of the mathematical competency measurement model. Then, they evaluated the degree of student mathematical competency using the Wright map and the Multidimensional Test Response Model. The Wright map was used to validate evidence of the quality of the progress map in each dimension with regard to each content constituent. As a result, the MRCMLM was used to correlate the modeling of items in multidimensional competencies, assess the student's competency parameters, and place the transition points for diagnosing the degree of the student's mathematical competency in each dimension. The ACER ConQuest Version 2.0 program (Wu et al., 2007) was applied to estimate the mathematical competency measurement model parameters.

In the final phase, the researchers revised the heuristic evaluation ideas of Stufflebeam (Kanjanawasi, 2011) to appraise the overall quality of the real-time digital learning platform for diagnosing the degree of mathematical competency of the seventh-grade students in three aspects, namely usefulness, feasibility, and accuracy. Six

experts carried out this process. A five-point Likert scale assessment document was employed to measure the experts' interpretation of the real-time digital learning platform, ranging from minimal to most applicable. Table 2 shows the five degrees of applicability, which were determined in accordance with the mean score range.

Table 2

Clarification of applicable degree according to the mean score range

Mean Score Range	Interpretation
1.00–1.49	Minimal applicable
1.50–2.49	Less applicable
2.50–3.49	Medium applicable
3.50–4.49	Very applicable
4.50–5.00	Most applicable

Source: Authors' work

FINDINGS AND DISCUSSION

The researchers adapted the mathematical competency measurement model in relation to the proposals recommended by past researchers that involve the MME dimension (Junpeng et al., 2018) and the CKL dimension (Biggs & Collis, 1982). For each degree of learning competency, an assessment process must be conducted correctly utilizing the most appropriate assessment tool for the three content constituents: number and algebra, measurement and geometry, and statistics and probability.

Creation of a Multidimensional Mathematical Competency Measurement Model

Researchers need to develop assessment tools to provide an overview that truly describes the degree of students' mathematical competency in terms of their abilities.

Internal Structure of the Assessment Tools

The researchers developed three assessment tools consisting of 21 items for the number and algebra content constituent, 20 items for the measurement and geometry content constituent, and 17 items for the statistics and probability constituent. The number and algebra assessment tool consisted of 11 items in the MME dimension and 10 in the CKL dimension. The measurement and geometry assessment tool comprises ten items in each MME and CKL dimension. The statistics and probability assessment tool included eight items in the MME dimension and nine in the CKL dimension. Table 3 shows the internal structure of the three assessment tools with regard to the digital learning platforms.

Comparing the Model Fit

The internal structure of the mathematical competency measurement model in terms of its accuracy with regard to the mathematical competency constructs was performed by comparing the model fit for unidimensional and multidimensional models. The unidimensional model refers to clustering all the items into one dimension. In contrast, the multidimensional model separates items into the MME and CKL dimensions. The findings indicated that the unidimensional model was found to fit significantly better statistically than the multidimensional model with regard to the

Table 3
Assessment tools of the three content constituents

Content constituent	MME (Item)	Total of MME	CKL (Item)	Total of CKL	Total item
Number and algebra	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11	11	12, 13, 14, 15, 16, 17, 18, 19, 20, 21	10	21
Measurement and geometry	1, 2, 3, 6, 15, 16, 17, 18, 19, 20	10	4, 5, 7, 8, 9, 10, 11, 12, 13, 14	10	20
Statistics and probability	3, 7, 8, 9, 10, 11, 12, 13	8	1, 2, 4, 5, 6, 14, 15, 16, 17	9	17

Source: Authors' work

number and algebra content constituent. The findings were shown in terms of the Likelihood Ratio Chi-Square G^2 ($\chi^2 = 3.83$, $df = 2$, $p = 0.01$) (Wilson & De Boeck, 2004) as well as the Akaike Information Criterion (AIC) (Yao & Schwarz, 2006) and the Bayesian Information Criterion (BIC) (Schwarz, 1978) which had a higher value in multidimensional concepts for measuring mathematical competencies in terms of the number and algebra content constituent.

However, the findings showed that the multidimensional model was a statistically better fit than the unidimensional model in terms of the measurement and geometry content constituent and the statistics and probability content constituent. It is reflected in the results with regard to the Likelihood Ratio Chi-Square G^2 ($\chi^2 = 10.72$, $df = 2$, $p = 0.01$) for the measurement and geometry content constituent, and the Likelihood Ratio Chi-Square G^2 ($\chi^2 = 7.32$, $df = 2$, $p = 0.01$) (Wilson & De Boeck, 2004). In addition, the AIC (Yao & Schwarz, 2006) and the BIC (Schwarz, 1978) had smaller values in terms of the multidimensional concepts for measuring mathematical competencies for both the measurement and geometry

content constituent as well as the statistics and probability content constituent. Table 4 demonstrates the comparative findings of the model fit for the three content constituents.

Interdimensional Correlation Analysis

The researchers continued by examining the correlation between the dimensions of the multidimensional mathematical competency measurement model. The findings indicated that there were positive, strong, and significant correlations between the MME and CKL dimensions, [$r(524) = .78$] for number and algebra content constituent, [$r(528) = .89$], the measurement and geometry content constituent, and the statistics and probability content constituent [$r = (507) = .89$] at a significant level of 0.05. Table 5 elucidates the findings of the interdimensional correlation analysis.

Item Fit Analysis of the Mathematical Competency Measurement Model

All the 58 items' difficulties in the mathematical competency measurement model were examined as presented in Tables 6, 7, and 8. The researchers referred to the criteria proposed by Adams and Khoo

Table 4
The comparison of model fit

Content constituent/ Model	Deviance	N of parameter	AIC	BIC
Number and algebra				
Unidimensional	13123.33	22	13167.33	13183.16
Multidimensional	13127.16	24	13175.16	13192.42
Likelihood Ratio Chi-Square G^2 ($\chi^2 = 3.83$, $df = 2$, $p = 0.01$)				
Measurement and geometry				
Unidimensional	12337.86	21	12379.86	12395.04
Multidimensional	12327.14	23	12373.14	12389.76
Likelihood Ratio Chi-Square G^2 ($\chi^2 = 10.72$, $df = 2$, $p = 0.01$)				
Statistics and probability				
Unidimensional	9560.79	18	9596.79	9609.48
Multidimensional	9553.44	20	9593.44	9607.54
Likelihood Ratio Chi-Square G^2 ($\chi^2 = 7.32$, $df = 2$, $p = 0.01$				

Source: Authors' work

Table 5
Interdimensional correlation analysis

Content constituent	Dimension	MME dimension	CKL dimension
Number and algebra	MME		0.50**
	CKL	0.78*	
	Variance	0.56	0.72
Measurement and geometry	MME		1.38**
	CKL	0.89*	
	Variance	1.37	1.76
Statistics and probability	MME		1.50**
	CKL	0.89*	
	Variance	1.97	1.50

Note. *Represents the correlation value **Represents the co-variances
Source: Authors' work

(1996) and Wilson et al. (2006) to verify the appropriateness of OUTFIT MNSQ and INFIT MNSQ values that have to be between 0.75–1.33. For the number and algebra content constituent, the findings showed that the item difficulties ranged from -0.674 logits (*SE* = 0.096) to +1.227 logits (*SE* = 0.106) and -0.187 logits (*SE* = 0.094) to 2.279 logits (*SE* = 0.142) in the MME dimension and the CKL dimension correspondingly. Moreover, the findings in terms of both OUTFIT MNSQ and the INFIT MNSQ values were between 0.88–1.96 and 0.89–1.22, fulfilling the criteria in terms of the suitability of the digital learning platform with regard to the number and algebra content constituent (Adam & Khoo, 1996; Wilson et al., 2006) as illustrated in

Table 6

Findings of item fit analysis for number and algebra constituent

Item	Difficulty	SE	OUTFIT		INFIT		Item score threshold			
			MNSQ	CI	MNSQ	CI	1	2	3	4
MME1	0.878	0.100	1.13	(0.88, 1.12)	1.09	(0.92, 1.08)			0.88	
MME2	0.536	0.095	1.07	(0.88, 1.12)	1.07	(0.93, 1.07)			0.54	
MME3	0.131	0.093	0.99	(0.88, 1.12)	0.99	(0.94, 1.06)		0.13		
MME4	0.760	0.098	1.00	(0.88, 1.12)	0.99	(0.92, 1.08)			0.76	
MME5	-0.664	0.096	0.95	(0.88, 1.12)	0.98	(0.93, 1.07)	-0.66			
MME6	1.227	0.106	1.08	(0.88, 1.12)	1.03	(0.90, 1.10)				1.23
MME7	1.216	0.106	1.04	(0.88, 1.12)	0.99	(0.90, 1.10)				1.22
MME8	1.031	0.102	0.96	(0.88, 1.12)	0.94	(0.91, 1.09)				1.03
MME9	0.666	0.097	0.95	(0.88, 1.12)	0.95	(0.93, 1.07)			0.66	
MME10	-0.282	0.093	0.97	(0.88, 1.12)	0.98	(0.94, 1.06)	-0.28			
MME11	-0.674	0.096	1.00	(0.88, 1.12)	1.01	(0.93, 1.07)	-0.67			
Transition Point							-0.54	0.13	0.71	1.16
CKL12	-0.187	0.094	0.96	(0.88, 1.12)	0.98	(0.94, 1.06)	-0.19			
CKL13	0.960	0.102	0.88	(0.88, 1.12)	0.90	(0.91, 1.09)			0.96	
CKL14	0.729	0.099	0.89	(0.88, 1.12)	0.90	(0.92, 1.08)			0.73	
CKL15	1.013	0.103	0.89	(0.88, 1.12)	0.89	(0.91, 1.09)			1.02	
CKL16	0.193	0.094	0.96	(0.88, 1.12)	0.97	(0.94, 1.06)		0.20		
CKL17	2.279	0.142	1.69	(0.88, 1.12)	1.22	(0.81, 1.19)				2.28
CKL18	1.066	0.104	1.28	(0.88, 1.12)	1.17	(0.90, 1.10)			1.06	
CKL19	1.327	0.109	1.30	(0.88, 1.12)	1.19	(0.89, 1.11)			1.33	
CKL20	0.220	0.094	0.88	(0.88, 1.12)	0.90	(0.94, 1.06)		0.22		
CKL21	0.798	0.100	0.91	(0.88, 1.12)	0.92	(0.92, 1.08)			0.80	
Transition Point							-0.19	0.21	0.98	2.28

Source: Authors' work

Figure 1. Therefore, it can be concluded that the number and algebra assessment tool is appropriate in terms of the data.

Next, the findings with regard to the measurement and geometry content constituent indicated that the item difficulties varied from -1.332 logits (*SE* = 0.109) to +0.935 logits (*SE* = 0.105) and -1.172 logits (*SE* = 0.108) to 0.870 logits (*SE* = 0.106) in the respective MME and CKL dimensions. Nevertheless, the findings in terms of the OUTFIT MNSQ and the INFIT MNSQ values were between 0.83–1.25 and 0.87–1.18. Consequently, they have achieved the criteria for determining the suitability of the digital learning platform in terms of the measurement and geometry content constituent (Adam & Khoo, 1996; Wilson et al., 2006), as demonstrated in Figure 2. Therefore, the researchers concluded that the measurement and geometry assessment tool is appropriate for the data.

The final findings of item fit relating to statistics and probability (17 items) showed
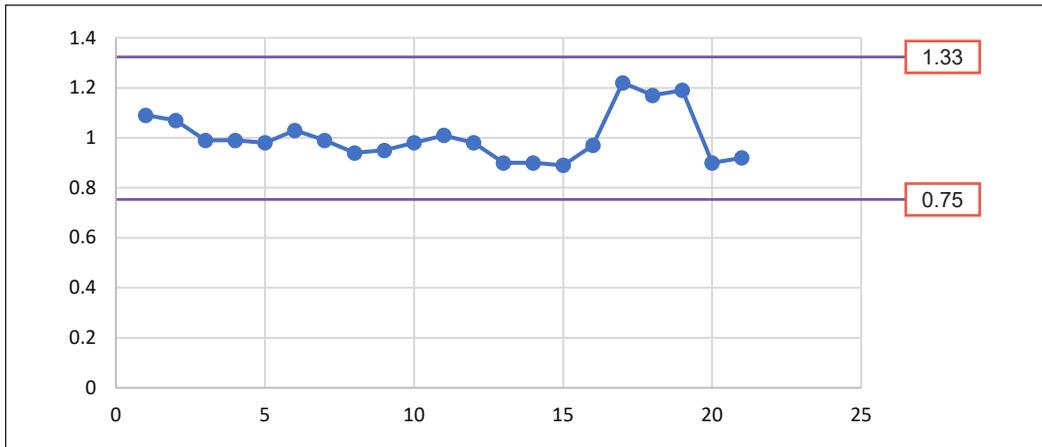
Figure 1. Item fit graph of INFIT MNSQ values of 21 items in number and algebra
Source: Authors' work

Table 7
Findings of item fit analysis for measurement and geometry constituent

Item	Difficulty	SE	OUTFIT		INFIT		Item score threshold			
			MNSQ	CI	MNSQ	CI	1	2	3	4
MME1	-0.173	0.098	0.86	(0.88, 1.12)	0.90	(0.93, 1.07)		-0.17		
MME2	-0.952	0.103	0.80	(0.88, 1.12)	0.91	(0.91, 1.09)	-0.95			
MME3	0.392	0.100	1.14	(0.88, 1.12)	1.03	(0.92, 1.08)			0.39	
MME6	0.452	0.100	0.87	(0.88, 1.12)	0.89	(0.92, 1.08)			0.45	
MME15	-0.067	0.098	1.16	(0.88, 1.12)	1.11	(0.92, 1.08)		-0.07		
MME16	0.935	0.105	1.05	(0.88, 1.12)	0.94	(0.90, 1.10)				0.94
MME17	0.719	0.103	1.31	(0.88, 1.12)	1.17	(0.91, 1.09)				0.72
MME18	-1.332	0.109	1.27	(0.88, 1.12)	1.10	(0.90, 1.10)	-1.34			
MME19	-0.703	0.101	0.90	(0.88, 1.12)	0.94	(0.92, 1.08)	-0.70			
MME20	0.166	0.099	1.02	(0.88, 1.12)	1.01	(0.92, 1.08)		0.16		
Transition Point							-1.00	-0.03	0.42	0.83
CKL4	-0.945	0.105	0.82	(0.88, 1.12)	0.94	(0.91, 1.09)	-0.95			
CKL5	0.175	0.101	0.95	(0.88, 1.12)	0.96	(0.92, 1.08)			0.17	
CKL7	0.605	0.104	1.13	(0.88, 1.12)	1.04	(0.91, 1.09)				0.60
CKL8	0.155	0.101	1.15	(0.88, 1.12)	1.12	(0.92, 1.08)			0.16	
CKL9	-1.080	0.107	1.13	(0.88, 1.12)	1.01	(0.90, 1.10)	-1.08			
CKL10	0.870	0.106	1.34	(0.88, 1.12)	1.24	(0.90, 1.10)				0.87
CKL11	-1.149	0.108	0.77	(0.88, 1.12)	0.91	(0.90, 1.10)	-1.15			
CKL12	-1.172	0.108	0.70	(0.88, 1.12)	0.85	(0.90, 1.10)	-1.17			
CKL13	0.802	0.106	1.04	(0.88, 1.12)	0.98	(0.90, 1.10)				0.80
CKL14	-0.508	0.102	0.82	(0.88, 1.12)	0.88	(0.92, 1.08)		-0.51		
Transition Point							-1.09	-0.51	0.17	0.76

Source: Authors' work
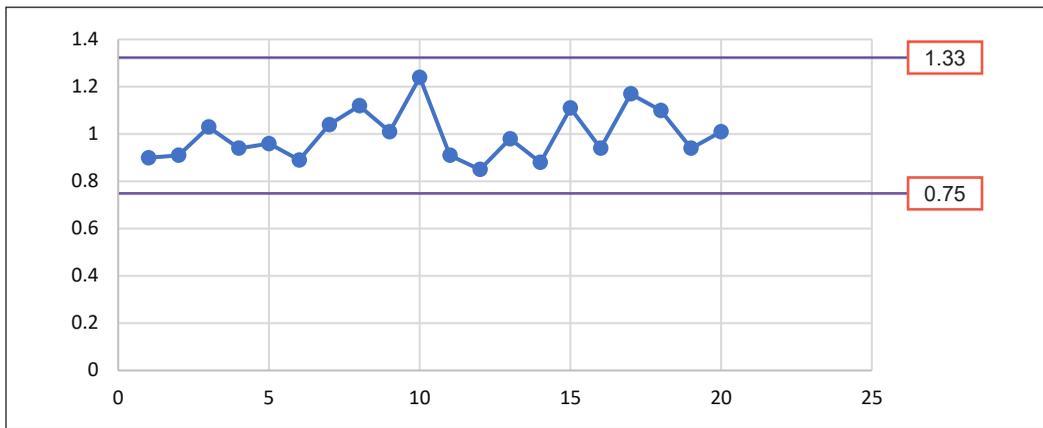
Figure 2. Item fit graph of INFIT MNSQ values of 20 items in measurement and geometry
Source: Authors' work

that the item difficulties fluctuated from –2.853 logits (*SE* = 0.158) to +0.608 logits (*SE* = 0.104) for the MME dimension and from -0.516 logits (*SE* = 0.103) to +3.158 logits (*SE* = 0.183) for the CKL dimension. In addition, the findings with regard to both OUTFIT MNSQ and INFIT MNSQ values were between 0.83–1.25 and 0.87–1.18 and have therefore achieved the criteria for determining the suitability of the digital learning platform in terms of the statistics

and probability content constituent (Adam & Khoo, 1996; Wilson et al., 2006) as verified in Figure 3. Therefore, it can be concluded that the statistics and probability assessment tool is in an acceptable range.

Evaluation of Degrees of Student Mathematical Competency Using Wright Map

The researchers continued by examining the structural validity of the construct map
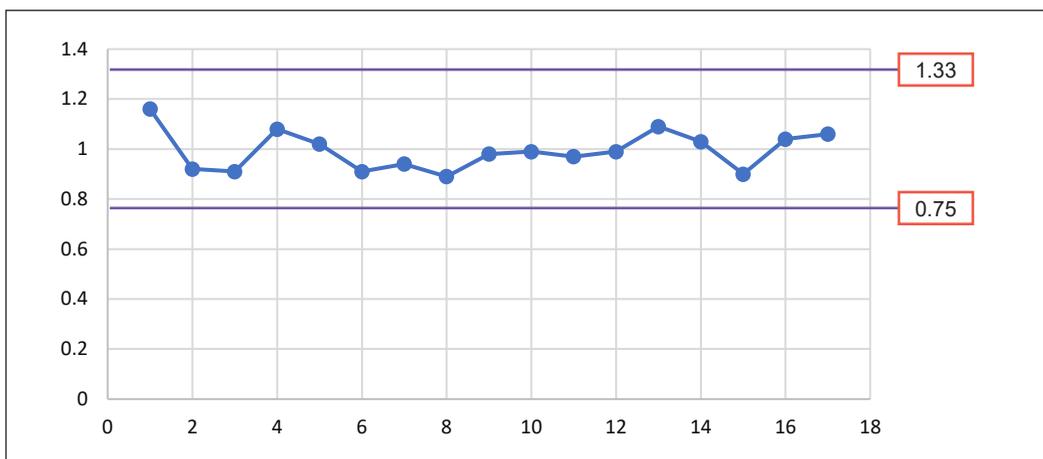
Figure 3. Item fit graph of INFIT MNSQ values of 17 items in statistics and probability
Source: Authors' work

Table 8
Findings of item fit analysis for statistics and probability constituent

Item	Difficulty	SE	OUTFIT		INFIT		Item score threshold			
			MNSQ	CI	MNSQ	CI	1	2	3	4
MME3	-1.196	0.109	0.79	(0.88, 1.12)	0.91	(0.90, 1.10)		-1.20		
MME7	-1.713	0.119	1.01	(0.88, 1.12)	0.94	(0.87, 1.13)		-1.71		
MME8	-2.853	0.158	0.71	(0.88, 1.12)	0.89	(0.78, 1.22)	-2.85			
MME9	-0.209	0.102	0.99	(0.88, 1.12)	0.98	(0.92, 1.08)			-0.21	
MME10	-0.406	0.102	1.02	(0.88, 1.12)	0.99	(0.92, 1.08)			-0.41	
MME11	-0.516	0.103	0.95	(0.88, 1.12)	0.97	(0.92, 1.08)			-0.52	
MME12	0.608	0.104	1.01	(0.88, 1.12)	0.99	(0.91, 1.09)				0.61
MME13	-0.146	0.102	1.12	(0.88, 1.12)	1.09	(0.92, 1.08)			-0.16	
Transition Point							-2.85	-1.46	-0.33	0.61
CKL1	0.105	0.102	1.21	(0.88, 1.12)	1.16	(0.92, 1.08)		0.11		
CKL2	-0.479	0.103	0.86	(0.88, 1.12)	0.92	(0.92, 1.08)	-0.49			
CKL4	3.158	0.183	2.33	(0.88, 1.12)	1.08	(0.74, 1.26)				3.16
CKL5	-0.516	0.103	1.00	(0.88, 1.12)	1.02	(0.92, 1.08)	-0.52			
CKL6	0.303	0.102	0.89	(0.88, 1.12)	0.91	(0.92, 1.08)		0.30		
CKL14	0.937	0.108	1.12	(0.88, 1.12)	1.03	(0.90, 1.10)			0.94	
CKL15	0.116	0.102	0.89	(0.88, 1.12)	0.90	(0.92, 1.08)		0.12		
CKL16	-0.132	0.102	1.07	(0.88, 1.12)	1.04	(0.92, 1.08)		-0.13		
CKL17	0.387	0.103	1.07	(0.88, 1.12)	1.06	(0.91, 1.09)		0.39		
Transition Point							-0.51	0.16	0.94	3.16

Source: Authors' work

and the criterion area of the Wright map to verify the transition points. The transition points were calculated from the mean of the item thresholds in each degree for each dimension, as demonstrated in Tables 6, 7, and 8.

Degrees of Student Mathematical Competency Using the Wright Map for the Number and Algebra Content Constituent

Table 6 shows the difficulty and threshold values of 21 items, 11 in the MME dimension and 10 in the CKL dimension. Each dimension has four degrees of mathematical competency. The four degrees of competency in terms of the MME dimension for the number and algebra content constituent are: Threshold 1 consists of items 5, 10, 11 with threshold values of -066, -0.28, and -0.67, respectively; Threshold 2 consists of item 3 only with a threshold value of 0.13; Threshold 3 consists of item 1, 2, 4, and 9 with threshold values of 0.88, 0.54, 0.76, 0.66, correspondingly, and Threshold 4 consists of item 6, 7, and 8 with threshold values 1.23, 1.22, 1.03, respectively. Conversely, the four degrees of competency of the CKL dimension for

the number and algebra content constituent are: Threshold 1 consists of item 12 with a threshold value of -0.19; Threshold 2 consists of items 16 and 20 with threshold values of 0.20 and 0.22; Threshold 3 consists of item 13, 14, 15, 18, 19, and 21 with threshold values 0.96, 0.73, 1.02, 1.06, 1.33, and 0.80, respectively, and Threshold 4 only has item 17 with a threshold value of 2.28.

Degrees of Student Mathematical Competency Using the Wright Map for the Measurement and Geometry Content Constituent

Table 7 shows the difficulty and threshold values of 20 items. Each dimension has ten items as well as four degrees of mathematical competency. The four degrees of competency of the MAP dimension for the measurement and geometry content constituent are: Threshold 1 consists of items 2, 18, 19 with threshold values of -095, -1.34, and -0.70, respectively; Threshold 2 consists of items 1, 15, and 20 with threshold values of -0.17, -0.07, and 0.16 respectively; Threshold 3 consists of items 3 and 6 with threshold values as 0.39 and 0.45, and Threshold 4 consists of items 16 and 17 with threshold values 0.94 and 0.72, respectively. In addition, the four degrees of competency of the CKL dimension for the measurement and geometry content constituent are: Threshold 1 consists of items 4, 9, 11, and 12 with threshold values of -0.95, -1.08, -1.15, and -1.17, respectively; Threshold 2 consists of item 14 only, with a threshold

value of -0.51; Threshold 3 consists of items 5 and 8 with threshold values 0.17 and 0.16, respectively, and Threshold 4 consists of items 7, 10, and 13 with threshold values of 0.60, 0.87, and 0.80, respectively.

Degrees of Student Mathematical Competency Using the Wright Map for the Statistics and Probability Content Constituent

Table 8 shows the difficulty and threshold values of 17 items, namely eight items in the MME dimension and nine in the CKL dimension. Each dimension has four degrees of mathematical competency. The four degrees of competency of the MME dimension for the statistics and probability content constituent are: Threshold 1 consists of item 8 only with a threshold value of -0.28; Threshold 2 consists of items 3 and 7 with threshold values of -1.20 and -1.71 respectively; Threshold 3 consists of item 9, 10, 11, and 13 with threshold values as -0.21, -0.41, -0.52, -0.16, respectively, and Threshold 4 consists of item 12 with a threshold value of 0.61. In contrast, the four degrees of competency of the CKL dimension for statistics and probability content constituent are: Threshold 1 consists of item 2 and item 5 with threshold values of -0.49 and -0.52, respectively; Threshold 2 consists of items 1, 6, 15, 16, and 17 with threshold values of 0.11, 0.30, 0.12, -0.13, and 0.39, respectively; Threshold 3 consists of only item 14 with a threshold value 0.94, and Threshold 4 has item 4 with a threshold value of 3.16.

Degrees of Student Mathematical Competency Using a Multidimensional Test Response Model

The mean threshold of each dimension of the degree of mathematical competency for the three studied content constituents was applied to create a standardized mathematical competency measurement model. The transition point was calculated from the mean of the item thresholds in each degree of the dimension, as elucidated in Tables 6, 7, and 8 for the respective number and algebra content constituent, the measurement and geometry content constituent, and the statistics and probability content constituent. The researchers then created measurement standards by taking into account the transition in conjunction with the criteria area on the Wright map for every degree of mathematical competency, as shown in Figures 4, 5, and 6. The mean threshold is determined with regard to the equivalent degree of mathematical competency for every content constituent of the two dimensions of mathematics. The researchers made a comparison of students and items by using the Wright map to verify the transition point through

the criteria area setting (Lunz, 2010) to understand the measurement of mathematical competency specifically. Based on this line of reasoning, the benchmarks for diagnosing students' mathematical competency levels in every dimension would follow the intersection point to cluster degrees of student mathematical competency in accordance with the categorization recommended by Junpeng et al. (2019). Table 9 shows the construct map of the MME and CKL dimensions.

Figure 4. Wright map indicates the transition points of MME and CKL dimensions for number and algebra content constituent
Source: Authors' work

Table 9
Construct maps of the MME and CKL dimensions of mathematical competence standards

Dimension Level	Learning Progress Level	Diagnostic Descriptions
MME Degree 5	Strategic/extended thinking	Students can solve a complex problem using multiple steps. The procedure is clearly described and illustrated in various methods that can be utilized to a higher degree.
MME Degree 4	Simple skills and concepts	Students can solve a complex problem using more than one step. The procedure involves applying principles and theories, but the answer is not matched due to miscalculation.
MME Degree 3	Basic memory and reproduction	Students can use their basic understanding of mathematical procedures to find the answer. However, the answer is incomplete due to a lack of application.

Table 9 *(continue)*

Dimension Level	Learning Progress Level	Diagnostic Descriptions
MME Degree 2	Unrecalled memory	Students can use their understanding of simple mathematical procedures to solve the problem. However, the answer reflects that the student is unable to remember important content as a basis for solving the problem.
MME Degree 1	Non-response/ irrelevance	Students cannot use proper mathematical procedures to obtain the answer.
CKL Degree 5	Extended abstract structure	Students can relate concepts, summarize references as abstract knowledge, and use them to conclusions.
CKL Degree 4	Relation structure	Students can connect complex concepts, describe each part, and relate to problem situations but cannot summarize as an abstract body of knowledge to analyze or integrate the concepts.
CKL Degree 3	Multi-structure	Students can connect simple concepts but are unable to identify what they have learned in the form of memory to solve problems.
CKL Degree 2	Uni-structure	Students can solve an example-like problem without adding conditions or concepts. However, they often misunderstand concepts and lack skills in interpreting information.
CKL Degree 1	Pre-structure	Students lack the skills to interpret problem situations and are unable to form concepts.

Source: Authors' work

Degrees of Student Mathematical Competency Using a Multidimensional Test Response Model for the Number and Algebra Content Constituent

Based on the determination of the cut-off point in measuring the mathematical competencies of the students' test in terms of the number and algebra content constituent, the findings indicated that the transition point could be allocated into four cut-off points of five degrees each in ascending order. For instance, the intersections of the MME dimension were discovered from Degrees 1 to 2, Degrees 2 to 3, Degrees 3 to 4, and Degrees 4 to 5 as -0.54, 0.13, 0.71, and 2.43, respectively. Conversely, the intersections of the CKL dimension were recognized from Degrees 1 to 2, Degrees 2 to 3, Degrees 3 to 4, and Degrees 4 to 5 as -0.19, 0.21, 0.98, and 2.28, respectively. Figure 4 shows the transition point in every dimension from the Wright map with regard to the number and algebra content constituent.

The researchers then used the measurement criterion findings from the Wright map to design the specification adjustments for a mathematical competency measurement model for the number and algebra content constituent. Thus, the researchers determined five score ranges, which were transformed from evaluation mathematical competency parameters into scale scores and raw scores in that order. Table 10 demonstrates that those students who achieve their logits smaller than -0.54 and -0.19 in the MME and CKL dimensions, respectively, are

Table 10
Findings of determination of mathematical competence standards for the number and algebra content constituent

Dimension	Intersection θ	θ range	Scale scores	Raw scores
MME Degree 5	1.16	>1.16	>61.60	8–11
MME Degree 4	0.71	0.71<θ<1.16	57.10–61.59	7
MME Degree 3	0.13	0.13<θ<0.70	51.30–57.09	5–6
MME Degree 2	-0.54	0.54<θ<0.12	44.60–51.49	4
MME Degree 1	-	<0.54	<44.60	1–3
CKL Degree 5	2.28	>2.28	>72.80	8–10
CKL Degree 4	0.98	0.98<θ<2.28	59.80–72.80	6–7
CKL Degree 3	0.21	0.21<θ<0.97	52.10–59.79	4–5
CKL Degree 2	-0.19	-0.19<θ<0.20	39.10–52.09	3
CKL Degree 1	-	<-0.19	<39.10	0–2

Source: Authors' work

identified as possessing the lowest degree of mathematical competency. Likewise, if their logits are greater than 1.16 and 2.28 with regard to the MME and CKL dimensions, respectively, those students are identified as achieving the highest degree of mathematical competency.

Degrees of Student Mathematical Competency Using a Multidimensional Test Response Model for the Measurement and Geometry Content Constituent

According to the determination of the cut-off point in measuring the mathematical competencies of the students' test in the measurement and geometry content constituent, the findings indicated that the transition point could be allocated into four cut-off points of five degrees in ascending order. For instance, the intersections of MME dimensions were discovered from Degrees 1 to 2, Degrees 2 to 3, Degrees 3 to

4, and Degrees 4 to 5 as -1.00, -0.03, 0.42, and 0.83, respectively. On the other hand, the intersections of the CKL dimension were identified from Degrees 1 to 2, Degrees 2 to 3, Degrees 3 to 4, and Degrees 4 to 5 as -1.09, -0.51, 0.17, and 0.76, respectively. Figure 5 shows the transition point in
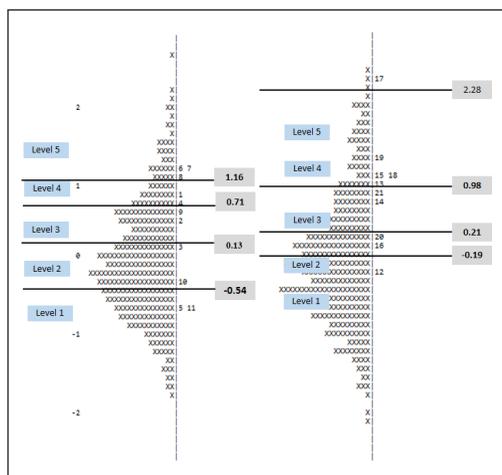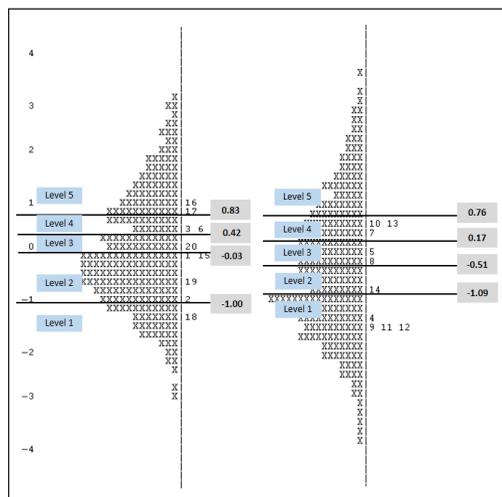

Figure 5. Wright map indicates the transition points of MME and CKL dimensions for measurement and geometry content constituent
Source: Authors' work

every dimension of the Wright map for the measurement and geometry content constituent.

The researchers then used the measurement criterion findings from the Wright map to design the specification adjustments for a mathematical competency measurement model for the measurement and geometry content constituent. Thus, the researchers identified five score ranges, which are transformed from evaluation mathematical competency parameters into scale scores and raw scores correspondingly. Table 11 reveals that those students who attain logits lower than -1.00 and -1.09 in the MME and CKL dimensions, respectively, are identified as possessing the lowest degree of mathematical competency. Similarly, if their logits are greater than 0.83 and 0.76 with regard to the MME and CKL dimensions, those students are identified as possessing the highest degree of mathematical competency.

Degrees of Student Mathematical Competency Using a Multidimensional Test Response Model for the Statistics and Probability Content Constituent

As stated by the determination of the cut-off point in gauging the mathematical competencies of the students' test in terms of the statistics and probability content constituent, the findings indicated that the transition point could be allotted into four cut-off points of five degrees in ascending order. For occurrence, the intersections of the MME dimension were discovered from Degrees 1 to 2, Degrees 2 to 3, Degrees 3 to 4, and Degrees 4 to 5 as -2.85, -1.46, -0.33, and 0.61, respectively. Conversely, the intersections of the CKL dimension were identified from Degree 1 to 2, Degree 2 to 3, Degree 3 to 4, and Degree 4 to 5 as -0.51, 0.16, 0.94, and 3.16, respectively. Figure 6 shows the transition point in every dimension from the Wright map with regard to the statistics and probability content constituent.

Table 11

Findings with regard to the determination of mathematical competence standards for the measurement and geometry content constituent

Dimension	Intersection θ	θ range	Scale scores	Raw scores
MME Degree 5	0.83	>0.83	>58.30	8–10
MME Degree 4	0.42	0.42<θ<0.83	54.20–58.29	7
MME Degree 3	-0.03	-0.03<θ<0.41	49.70–54.19	4–6
MME Degree 2	-1.00	-1.00<θ<-0.04	40.00–49.69	2–3
MME Degree 1	-	<-1.00	<40.00	0–1
CKL Degree 5	0.76	>0.76	>57.60	8–10
CKL Degree 4	0.17	0.17<θ<0.76	51.70–57.59	7
CKL Degree 3	-0.51	-0.51<θ<0.16	44.90–51.69	4–6
CKL Degree 2	-1.09	-1.09<θ<-0.52	39.10–44.89	2–3
CKL Degree 1	-	<-1.09	<39.10	0–1

Source: Authors' work
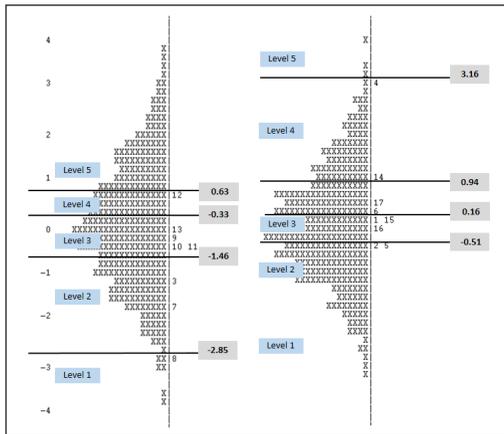
Figure 6. Wright map indicates the transition points of MME and CKL dimensions for statistics and probability content constituent
Source: Authors' work

parameters into scale scores and raw scores in that order. Table 12 indicates that those students who achieved logits lower than -2.85 and -0.51 in MME and CKL dimensions are separately deemed to possess the lowest degree of mathematical competency. Correspondingly, those students possess the highest degree of mathematical competency if their logits are greater than 0.63 and 3.16 in the MME and CKL dimensions.

Findings of Heuristic Evaluation

The heuristic evaluation findings are based on the quality of the real-time digital learning platform in terms of its usefulness, feasibility, and accuracy, as reported by six experts. The six experts reported that the real-time digital learning platform for diagnosing the degree of student mathematical competency in number and algebra, measurement and geometry, and statistics and probability content constituents was highly applicable in terms of its usefulness, feasibility, and

The researchers then used the measurement criterion findings from the Wright map to identify the specification adjustments for a mathematical competency measurement model for the statistics and probability content constituent. Consequently, the researchers determined five score ranges, which were transformed from evaluation mathematical competency

Table 12

Findings of determination of mathematical competence standards for the statistics and probability content constituent

Dimension	Intersection θ	θ range	Scale scores	Raw scores
MME Degree 5	0.63	>0.63	>56.30	7–8
MME Degree 4	-0.33	-0.33<θ<0.63	46.70–56.29	6
MME Degree 3	-1.46	-1.46<θ<-0.34	35.40–46.69	4–5
MME Degree 2	-2.85	-2.85<θ<-1.47	21.50–35.39	2–3
MME Degree 1	-	<-2.85	<21.50	0–1
CKL Degree 5	3.16	>3.16	>81.60	7–9
CKL Degree 4	0.94	0.94<θ<3.16	59.40–81.59	5–6
CKL Degree 3	0.16	0.16<θ<0.93	51.60–59.39	3–4
CKL Degree 2	-0.51	-0.51<θ<0.15	44.90–51.59	1–2
CKL Degree 1	-	<-0.51	<44.90	0

Source: Authors' work

accuracy. The highest mean score related to the usefulness of the real-time digital learning platform (mean score = 4.50, *SD* = 0.50) and applicability. The finding implies that the real-time digital learning platform can react to the requirements of students, and students' feedback can be used to diagnose their degree of mathematical competency concretely. In addition, the report the platform provides can then be utilized to prepare students to improve their learning outcomes.

It is observed with regard to the accuracy of the real-time digital learning platform, which displayed a high degree of applicability with a mean score of 4.44 and a standard deviation of 0.52. Therefore, the finding implies that the real-time digital learning platform can accurately specify the measurement objectives, the measurement process is accurate, and the feedback reports accurately match the degree of student competency. Although the feasibility aspect was found to have the lowest mean score (4.17) with a standard deviation of 0.26, it was still deemed to be most applicable. The findings imply that the real-time digital learning platform is a clearly defined method for reporting the degree of student mathematical competency. In addition,

the scope of reporting content applies to a comprehensive measurement report. Finally, the overall quality of the real-time digital learning platform (mean score = 4.37; *SD* = 0.42) is consistent with current learning measurement guidelines. Table 13 illustrates the findings of the heuristic evaluation in detail.

CONCLUSION

Our findings indicate a need for mathematics teachers to respond swiftly to technology adoption by using the real-time digital learning platform for diagnosing degrees of mathematical competency with regard to seventh-grade students so that they can overcome the challenges posed by the COVID-19 pandemic. The key findings of this research relate to the quality of the real-time digital learning platform in terms of its usefulness, accuracy, and feasibility. Our findings correspond to Clark-Wilson's (2020) findings on teacher use of digital technology in mathematics learning and how this has been accomplished. Therefore, the researchers would like to suggest to the Thailand Ministry of Education that there is a need to adopt it nationwide to support student learning and foster their academic growth so that students can be assessed

Table 13
Findings of the heuristic evaluation

Quality Aspect	Mean Score	Standard Deviation	Applicability Degree
Usefulness	4.50	0.50	Most applicable
Feasibility	4.17	0.26	Most applicable
Accuracy	4.44	0.52	Most applicable
Overall	4.37	0.42	

Source: Authors' work

using an online environment that allows for automated and timely scoring. As a result, the real-time digital learning platform will allow a much faster turnaround of results, which is crucial for mathematics teachers when making instructional decisions at the individual student and instructional unit levels (Seo et al., 2021).

Implications for Theory and Practice

Existing educational theories often focus on generalized teaching methods, overlooking the diverse learning needs of individual students, especially in mathematics (Inprasitha, 2022). A real-time digital learning platform can address this gap by tailoring instructional content based on the specific competencies and learning styles of each student (Inprasitha, 2023). Besides, traditional assessment models may not effectively capture the dynamic nature of mathematical competency development. A real-time platform can incorporate continuous and adaptive assessment techniques, allowing for a more accurate and real-time evaluation of a student's mathematical proficiency.

In addition, the researchers employed by IRT in developing a real-time digital learning platform have allowed for creating adaptive assessments within a real-time digital learning platform. As students answer questions, the system can dynamically adjust the difficulty of subsequent items based on their previous responses. It ensures that the assessment is finely tuned to each student's level of competency, providing a more accurate and efficient diagnosis.

Moreover, IRT can be integrated into the platform to analyze the strengths and weaknesses of individual students. Based on their responses to specific items, the system can recommend personalized learning paths, focusing on areas where the student needs improvement. This adaptive and tailored approach enhances the efficiency of the learning process. In conclusion, the digital platform can continuously monitor a student's progress over time rather than relying solely on periodic assessments, and the system can gather data from each interaction to update the estimate of the student's mathematical competency. This real-time monitoring enables timely intervention and support.

Limitations and Recommendations for Future Research

Developing a real-time digital learning platform for diagnosing degrees of mathematical competency is a complex task that involves various challenges and considerations. In other words, the effectiveness of the platform heavily relies on the quality and diversity of the data used for training and testing, which will be the limitation. Future research should focus on collecting high-quality and diverse datasets representing various mathematical competencies across different demographics.

Another limitation is that some students may have unique learning styles, and a one-size-fits-all approach might not be optimal. Therefore, future researchers are recommended to investigate methods for personalization and adaptability in the

platform to cater to individual learning preferences and needs. Moreover, providing accurate real-time feedback is challenging, and there may be instances where the platform misinterprets a student's competency level. Future researchers can invest in advanced algorithms and artificial intelligence techniques to enhance the accuracy of real-time feedback, minimizing false positives and negatives.

The researchers would like to suggest that future research investigate the incorporation of various pedagogical approaches, such as inquiry-based learning or collaborative learning, to enhance the overall effectiveness of the platform. On the other hand, future research should stay updated with technological advancements, such as machine learning and artificial intelligence advancements and leverage these technologies to improve the platform's capabilities.

As never before, students with limited mathematical competencies are indicated to have underdeveloped mathematical skills. It is a formidable barrier to participation in their future learning (Michael & Ritzelda, 2021) with the use of real-world requirements with regard to the mathematics knowledge used in daily life and their future academic and economic pursuits. The real-time digital learning platform is proven to be effective in helping students acquire the fundamental mathematical competencies they need to achieve success in their current and future learning, as emphasized by previous researchers (Kantahan et al., 2020; Kesorn et al., 2020; Maisa et al., 2021). The research indicates that a real-time digital learning platform not only diagnoses degrees of mathematical competency but also provides the knowledge needed to stimulate their individual needs and the competencies required in this new normal era.

The researchers concluded by addressing these limitations and pursuing these recommendations. Future researchers can contribute to developing a more effective, inclusive, and ethically sound real-time digital learning platform for diagnosing degrees of mathematical competency.

ACKNOWLEDGMENTS

REFERENCES

Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation, 31*(2-3), 162-172. https://doi.org/10.1016/j.stueduc.2005.05.008

Adams, R. J., & Khoo, S. T. (1996). *ACER Quest: the interactive test analysis system* (Version 2.10) [Computer software]. Australian Council for Educational Research. https://research.acer.edu.au/measurement/3/

Adams, R. J., Wilson, M., & Wang, W. (1997). The multidimensional random coefficient multinomial logit model. *Applied Psychological Measurement, 21,* 1-23. https://doi.org/10.1177/0146621697211001

Adom, D., Mensah, J. A., & Dakes, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education (IJERE), 9*(1), 109-119. https://doi.org/10.11591/ijere.v9i1.20457

Briggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy.* Academic Press. https://doi.org/10.1016/C2013-0-10375-3

Buakhamphu, S., Thawinkarn, D., Vatanasapt, P., Tang, K. N. (2024). Development of digital culture model for small size schools in the northeast region of Thailand. *Kurdish Studies, 12*(2), 724-732. https://doi.org/10.58262/ks.v12i2.056

Clark-Wilson, A., Robutti, O., & Thomas, M. (2020). Teaching with digital technology. *ZDM – International Journal of Mathematics Education, 52,* 1223-1242. https://doi.org/10.1007/s11858-020-01196-0

Michael, D. G., & Ritzelda, D. A. (2021). Improving the least mastered competencies on number sense of Grade 7 learners. *International Journal of Advanced Research, 10*(10), 127-139. https://doi.org/10.21474/IJAR01/15471

Fowler, S., O'Keeffe, L., Cutting, C., & Leonard, S. (2019). The mathematics proficiencies: A doorway into spatial thinking. *Annals of Punjab Medical College (APMC), 24*(1), 36-40.

Harman, M. (2021, May 8). The role of digital learning platforms in the academic growth of students. *Kitaboo.* https://kitaboo.com/the-role-of-digital-learning-platforms-in-the-academic-growth-of-students/

Hayes, A. (2024, February 23). Stratified random sampling. *Investopedia.* https://www.investopedia.com/terms/stratified_random_sampling.asp

International Commission on the Futures of Education. (2020). *Education in a post-COVID world: Nine ideas for public action.* United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000373717

Inprasitha, M. (2022). Lesson study and open approach development in Thailand: A longitudinal study. *International Journal for Lesson and Learning Studies, 11*(5) 1-15. https://doi.org/10.1108/IJLLS-04-2021-0029

Inprasitha, M. (2023). Blended learning classroom model: A new extended teaching approach for new normal. *International Journal for Lesson and Learning Studies, 12*(4), 288-300. https://doi.org/10.1108/IJLLS-01-2023-0011

Junpeng, P., Inprasitha, M., & Wilson, M. (2018). Modeling of the open-ended items for assessing multiple proficiencies in mathematical problem solving. *The Turkish Online Journal of Educational Technology, 2,* 142-149.

Junpeng, P., Krotha, J., Chanayota, K., Tang, K. N., & Wilson, M. (2019). Constructing progress maps of digital technology for diagnosing mathematical proficiency. *Journal of Education and Learning, 8*(6), 90−102. https://doi.org/10.5539/jel.v8n6p90

Kanjanawasi, S. (2011). *New theory of testing* (3rd ed.). Chulalongkorn University Press. https://www.car.chula.ac.th/display7.php?bib=b1770973

Kantahan, S., Junpeng, P., Punturat, S., Tang, K. N., Gochyyev, P., & Wilson, M. (2020). Designing and verifying a tool for diagnosing scientific misconceptions in genetics topic. *International Journal of Evaluation and Research in Education (IJERE), 9*(3), 564-571. https://doi.org/10.11591/ijere.v9i3.20544

Kepner, H. S., & Huinker, D. A. (2012). Assessing students' mathematical proficiencies on the common core. *Journal of Mathematics Education at Teacher College, 3*(1), 26-32. https://doi.org/10.7916/jmetc.v3i1.735

Kesorn, K., Junpeng, P., Marwiang, M., Pongboriboon, K., Tang, K. N., Bathia, S., & Wilson, M. (2020). Development of an assessment tool for mathematical reading, analytical thinking and mathematical writing. *International Journal of Evaluation and Research in Education (IJERE), 9*(4), 955-962. https://doi.org/10.11591/ijere.v9i4.20505

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. National Academy Press. https://doi.org/10.17226/9822

Linacre, J. M. (2005). Rasch dichotomous model vs. one-parameter logistic model. *Rasch Measurement Transactions, 19*(3), Article 1032.

Lunz, M. E. (2010). *Measurement research associates test insights*. https://www.rasch.org/mra/mra-01-10.htm

Maisa, S. D., & Musthfa, B. (2021). Charging personal growth of preservice teachers with 21st century career and life skill during the COVID-19 pandemic and new era. *The Asian ESP Journal, 17*(7.2), 60-82.

Maoto, S., Masha, K., & Mokwana, L. (2018). Teachers' learning and assessing of mathematical processes with emphasis on representations, reasoning and proof. *Pythagoras – Journal of the Association for Mathematics Education of South Africa, 39*(1), 1−10. https://doi.org/10.4102/pythagoras.v39i1.373

Manmai, T., Inprasitha, M., & Changsri, N. (2021). Cognitive aspects of students' mathematical reasoning habits: A study on utilizing lesson study and open approach. *Pertanika Journal of Social Sciences & Humanities, 29*(4), 2591-2614. https://doi.org/10.47836/pjssh.29.4.27

Milgram, R. J. (2007). What is mathematical proficiency? In A. H. Schoenfeld (Ed.), *Assessing mathematical proficiency* (pp. 31-58). Cambridge University Press. https://doi.org/10.1017/CBO9780511755378.007

Phaniew, S., Junpeng, P., & Tang, K. N. (2021). Designing standards-setting for levels of mathematical proficiency in measurement and geometry: Multidimensional item response model. *Journal of Education and Learning, 10*(6), 103-111. https://doi.org/10.5539/jel.v10n6p103

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press. https://archive.org/details/probabilisticmod0000rasc

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner-instructor interaction in online learning. *International Journal of Educational Technology in Higher Education, 18,* Article 54. https://doi.org/10.1186/s41239-021-00292-9

Thailand Ministry of Education. (2008). *The basic education core curriculum*. http://academic.obec.go.th/images/document/1525235513_d_1.pdf

Thailand Ministry of Education. (2017). *Learning standards and indicators learning of mathematics (revised edition) according to the Core Curriculum of Basic Education, 2008*. Agricultural Cooperative of Thailand. https://drive.google.com/file/d/1MDQEDkqGs01PnyzqEnyTVVNTS776ObCz/view

Thisopha, P., Thawinkarn, D., Wachrakul, C., & Tang K. N. (2023). An investigation on coding educational management for small-sized elementary schools in northeastern of Thailand. *Remittances Review, 8*(4), 3057-3071. https://doi.org/10.33182/rr.v8i4.211

Vongvanich, S. (2020). *Design research in education*. Chulalongkorn University Printing House. https://www.chulabook.com/education/102930

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Routledge. https://doi.org/10.4324/9781410611697

Wilson, M., Allen, D. D., & Li, J. C. (2006). Improving measurement in health education and health behavioral research using item response modeling: Comparison with the Classical Test Theory Approach. *Health Education Research, 21*, 19-32. https://doi.org/10.1093/her/cyl108

Wilson, M., & De Boeck, P. (Eds.). (2004). Descriptive and explanatory item response models. *Explanatory item models: A generalized linear and nonlinear* (pp. 43-41). Springer. https://link.springer.com/book/10.1007/978-1-4757-3990-9

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208. https://doi.org/10.1207/S15324818AME1302_4

Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement.* Mesa Press. https://research.acer.edu.au/measurement/1/

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACERConQuest Version 2: Generalized item response modeling software.* Australian Council for Educational Research. https://www.researchgate.net/publication/262187496_ConQuest_Version_2_Generalised_Item_Response_Modelling_Software

Yao, L., & Schwarz, R. D. (2006). A multidimensional partial credit model with associated item and test statistics: An application to mixed-format tests. *Applied Psychological Measurement, 30*(6), 469-492. https://doi.org/10.1177/0146621605284537

Yulian, V. R., & Wahyudin. (2018). Analyzing categories of mathematical proficiency based on Kilpatrick opinion in junior high school. *Journal of Physics: Conference Series, 1132,* Article 012052. https://doi.org/10.1088/1742-6596/1132/1/012052

Case Study

Short-term Virtual Mobility Program: Student's Experiences and Perceptions

Hooi Sin Soo[1]*, Saliza Abdul Aziz[2], Badariah Din[3] and Nur Khairiel Anuar[4]

[1]*School of Business Management, College of Business, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*
[2]*Tunku Puteri Intan Safinaz School of Accountancy, College of Business, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*
[3]*School of Government, College of Law, Government & International Studies, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*
[4]*School of Technology Management and Logistics, College of Business, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*

ABSTRACT

Virtual mobility programs have gained popularity in higher education institutions. It can be a viable complement to physical mobility. There is a lack of research on international mobility students' perceptions and experiences of virtual mobility programs, particularly during the transition from physical to online/virtual mobility amid the COVID-19 pandemic. This study explores students' perceptions and experiences within short-term virtual mobility programs. The participants of the short-term virtual mobility programs are from Southeast Asia. The findings of this study demonstrated that students have a positive perception of short-term virtual mobility programs' effectiveness via a descriptive analysis. Students' experiences were analyzed by thematic analysis, and positive and negative experiences toward short-term virtual mobility programs were found. This study suggests that virtual mobility programs should not be abandoned in the post-pandemic era. Higher education institutions may consider incorporating intercultural virtual mobility into academic programs to achieve the internationalization agenda.

INTRODUCTION

In contemporary higher education, internationalization is pivotal in shaping global rankings and fostering interactions

among diverse cultural groups within and across nations (Fernandes et al., 2022; Spencer-Oatey & Dauber, 2019). Internationalization promotes cultural exchange to enhance the quality of education and research for students and academics, and this phenomenon encompasses not only cross-border engagements but also intercultural exchanges at regional and national levels. The pursuit of internationalization in higher education is driven by a myriad of objectives, including enhancing the quality of education and research, fostering cultural exchange, and contributing to societal advancement (De Wit, 2020; Knight, 2008).

With its profound impact on various facets of academic institutions, including teaching, research, and student experiences, internationalization is a continuous process that necessitates sustained efforts for effective implementation. Recognizing its significance, many universities, including those in Malaysia, have embraced internationalization initiatives in their research and teaching endeavors (Munusamy & Hashim, 2019; Mustapha et al., 2021). These initiatives encompass a range of activities, such as attracting international students and scholars, fostering academic collaborations, and promoting international publications, all aimed at advancing the global engagement agenda set forth by national authorities.

The internationalization agenda aims to support the nation's development and economic growth, as outlined by Malaysia's Ministry of Higher Education (2015). Many universities in Malaysia have incorporated international dimensions in their research and teaching components (Munusamy & Hashim, 2019; Mustapha et al., 2021), focusing on the number of international students, international scholars, mobility exchange abroad, co-authored international publications in high-impact academic journals, mutual agreements and memorandums of understanding to support the internationalization agenda. These efforts promote the exchange of knowledge and support international collaborations, reputation building, and branding for universities in Malaysia.

In the context of higher education rankings, the presence of international students has emerged as a significant criterion in ranking systems, such as Times Higher Education World University Rankings and QS World University Rankings, prompting universities to devise strategies to attract students from diverse cultural backgrounds (Yee et al., 2018). One such strategy involves academic mobility programs, which facilitate the exchange of students and scholars across borders. These programs contribute to students' professional development and enhance the quality of human capital, thereby bolstering national economies (Slipchuk et al., 2021).

The outbreak of the COVID-19 pandemic has posed unprecedented challenges to academic mobility programs worldwide. According to a survey by the International Association of Universities, international student mobility has been impacted by 89% (Marinoni et al., 2020). With restrictions on

international travel and physical gatherings, universities have been compelled to explore alternative modes of engagement to sustain their internationalization efforts (Marinoni et al., 2020). In this context, virtual mobility programs have emerged as an effective solution, allowing students to engage in academic exchanges without physical travel. The transition towards a "new normal" is increasingly characterized by digitalization, with traditional teaching methods being transformed into online platforms (Recio & Colella, 2020).

The disruption caused by the COVID-19 pandemic has necessitated the adoption of virtual alternatives to traditional academic mobility programs. Yildirim et al. (2021) found that while students' interest in physical mobility has decreased amidst the pandemic, their interest in online mobility or participation in online international courses has risen. Short-term mobility programs have become more attractive as they allow students to continue their studies at home while participating in academic mobility programs in host countries (Slipchuk et al., 2021). Yang et al. (2022) proposed that diversifying academic mobility programs by either joint degree programs or short-term visiting projects abroad shall continue after the pandemic as a long-term development strategy to promote international mobility programs.

However, there is a dearth of understanding regarding students' experiences and perceptions of short-term virtual mobility programs, particularly in terms of the potential impact on students from Indonesia, Malaysia, and Thailand. Studies have explored students' perspectives on virtual mobility, primarily through case studies, particularly in Japan (Enkhtur et al., 2023), Europe (Baranowski & Jabkowski, 2023), Brazil (Queiroz et al., 2023), and Singapore (Cheng et al., 2023). Research on how students perceive the value of virtual mobility programs compared to traditional study abroad programs, especially in Southeast Asia, is very limited. A further understanding of the students' experiences and perceptions toward short-term virtual mobility programs is warranted. Thus, this research aims to explore students' experiences of participating in short-term virtual mobility programs and the perception of how this program could help them in future learning.

This paper mainly includes an introduction to introduce the background of studies related to international academic mobility, followed by the literature review section to discuss types of academic mobility and analyze the studies' findings. The third section is related to the method used to conduct the research. The final section is a summary of the main points and suggestions for future research on international academic mobility.

LITERATURE REVIEW

Academic Mobility Program

Academic mobility refers to a set of attitudes and willingness to move, as well as the physical movement itself, such as academic migration, for education, advanced professional training, and

academic activities for both students and teaching personnel in the education and science system, as defined by Ryazantsev et al. (2019). Academic mobility, also known as mobility exchange, allows international students to study and experience life outside their home country. It immerses students in different countries' social and academic cultures. Apart from gaining academic knowledge, participation in a mobility exchange program develops students' soft skills. Mobility exchange provides the opportunity to enhance language skills, be more independent, increase confidence, and learn and improve social and communication skills. The value of mobility exchange can invariably be seen in personal and professional development, frequently by incorporating an intercultural dimension into an academic path (Blankvoort et al., 2019; Chan et al., 2021; Woldegiorgis & Doevenspeck, 2015).

Apart from gaining different perspectives of experience, mobility exchange benefits students with cultural awareness. Students who participated in mobility exchange gained a better understanding of different cultural norms and appreciated the diversity in that country. Students' cultural development increases their understanding of foreign languages, eventually adding value to their future careers (Dixon & Tahmaz, 2020; Heirweg et al., 2020; Rahim, 2021). It also helps to develop international skills, increases chances of employability, and increases the likelihood of a person working abroad later in life. The richness of a mobility exchange program stems from the combination of academic experience and social-cultural experiences gained while living in a foreign country. Indeed, the significance of out-of-classroom experiences cannot be overstated, as they have been shown to significantly contribute to the transformative learning of international students. Academic mobility benefits students and home and host institutions, communities, and society (Knight, 2012).

An academic mobility program can be interpreted as a student exchange program and is determined by the time international students spend in the host country. Mobility usually refers to students who cross borders within/outside nations to receive short-term academic training in higher education and is associated with professional development, competence development, and broadening one's horizons (Blankvoort et al., 2019; Woldegiorgis & Doevenspeck, 2015). The initiative of an academic exchange program benefits students' professional growth and improves the quality of human capital, contributing to the national economy (Slipchuk et al., 2021). Participating in academic mobility exchange programs can increase the likelihood of individuals working abroad later in life, adding value to their future careers (Heirweg et al., 2020).

Inbound Mobility vs. Outbound Mobility

The distinction between mobility exchange, such as inbound and outbound mobility, is based on the mobility flows. Inbound mobility is classified as individual and/ or groups of students (undergraduate or

postgraduate) and academic staff from a foreign institution joining a local institution for educational or internship purposes in a set period, normally in a semester or up to a year (Rostovskaya et al., 2021). Outbound mobility entails individual and/or group exchanges of students (undergraduate and postgraduate) and academic staff to foreign institutions for educational and internship purposes for a set period, normally in a semester or up to a year (Rostovskaya et al., 2021).

In short, inbound mobility refers to foreign students traveling from their home country to attend a local institution for educational purposes. In contrast, outbound mobility refers to local students traveling to a foreign country to attend the academic program hosted by higher education institutions.

Credit Mobility vs. Non-credit Mobility

Students can consider participating in the academic mobility program with or without credit transfer at the end of the program. Credit mobility refers to those students who complete a semester of study at a host university and are eligible for credit transfer to the qualification they are pursuing at their home university. The credit transfer process occurs after students have completed a semester mobility exchange and received course results from the host university. The courses taken at the host university are recognized by their home university and transferable to their qualifications.

Non-credit mobility occurs when a student attends an academic mobility program but does not transfer credit to the qualification they are pursuing. Non-credit mobility is typically brief, lasting from a few hours to a few days or weeks. Students participate in non-credit mobility mainly to learn about and appreciate different cultures, volunteer in community service in another region, and build relationships and networks with international students. International students can attend a summer camp at a local institution or vice versa for two weeks. Non-credit mobility programs are designed to provide international experience and personal development rather than credit transfer.

Virtual Mobility Program amid the Pandemic

Prior to the COVID-19 pandemic, the mobility program was conducted on-site, where mobility exchange students met in person, allowing them to gain hands-on experience. Physical mobility exchange includes international transportability and the exchange of tertiary students and university staff attending a higher education program in a location other than their home country to broaden opportunities for intercultural enrichment and knowledge creation and transfer of knowledge (Rostovskaya et al., 2021). When the COVID-19 pandemic hit the world, universities adopted technology to ensure the continuance of teaching and learning and academic mobility was not excepted from the shift from physical to online mode, namely virtual mobility. The term "virtual mobility" is a coinage of the words virtual (information and

communication technology) and mobility (movement), and it refers to the ability to use ICT to replace the difficulties of movement. With the advancement of ICT, institutions can collaborate and compete more effectively, and students and academic lecturers who are unable to participate in physical exchanges can benefit from virtual mobility exchange programs.

Virtual mobility is considered a valuable alternative to physical mobility because it allows students to participate in courses at other institutions without leaving their home country and thus without incurring the risk of virus infection, financial costs, or visa application, although the time and space limitations (Cioltan-Drăghiciu & Stanciu, 2020; Yang et al., 2022). A university can provide an international experience for students and staff through virtual mobility by hosting an international webinar, an online international discussion, and a virtual international community support group. Moreover, virtual mobility exchange provides vast academic experience to the exchange students and develops their digital, linguistic, and intercultural competence (Helm, 2019; Yang et al., 2022); it also offers a new perspective for knowledge sharing.

Various activities have been initiated using online platforms to sustain the mobility collaboration with partner universities, i.e., online/hybrid classes, webinars, online/hybrid summer camps, and cultural events. For example, students are encouraged to discover and appreciate cultural diversity through a virtual cultural program. Students can share the various cultures of their home countries virtually by using multimedia or video presentations on the online platform. There will be an online discussion to exchange knowledge, such as sharing pre and post-cultural awareness or providing feedback on cultural reflection (Yang et al., 2022).

Challenges of Virtual Mobility

Virtual mobility has been popular among university faculty and students and has become a viable complement to physical mobility (Helm, 2019; Yang et al., 2022). Nonetheless, transitioning from physical to virtual mobility poses challenges for students and academic staff, administrative officials, institutions, and higher education. Although White and Lee (2020) stated that there is less space bound for the post-mobility world, there are some critical concerns about the quality of online learning (Teräs et al., 2020). From an institutional perspective, the academic staff's competence in online teaching pedagogies, IT infrastructure in place, and blended learning regulations will all contribute to the quality of online learning and determine the success of the mobility program. In contrast, connectivity, motivation, study skills, and time zone differences will also impact individual students' virtual mobility participation.

The transition from physical to virtual mobility has compelled university officials, faculty, and staff to quickly adapt to online solutions to ensure the mobility program's continuation. Changes imposed unexpected burdens on faculty and students, requiring them to quickly adapt to new forms of

teaching and learning. As this is a new practice for mobility students to become acquainted with the university's teaching practices and examination requirements, teaching and administrative staff play a critical role in assisting virtual mobility students.

Academic staff must act quickly to transition their lectures from classrooms and lecture halls to digital platforms, learn how to use digital tools to continue teaching and learning remotely, adjust both content and methodology online, find ways to incorporate critical thinking and reflection in learning and ensure student attention and participation in the online class. It will eventually cause stress for academic staff because most lecturers are digital immigrants (Prensky, 2001), which means they are not technologically savvy and are struggling to adapt to a new technology environment while refusing to abandon traditional teaching methods.

On the other hand, students thought their virtual mobility learning experience during the COVID-19 pandemic vastly differed from a typical mobility exchange. Because not all courses were designed to run synchronously (Koris et al., 2021), they are unable to perceive the physical interaction that should occur during a physical exchange. The students' adaptation to the host country's sociocultural environment is critical for an effective virtual mobility exchange program (Koris et al., 2021). Sociocultural adaptation is more difficult in the online mode of exchange because contact and interaction with the host country's culture

may be limited due to synchronous and asynchronous remote class sessions. The lack of social interaction between teacher and student made the virtual sociocultural immersion program even more difficult.

Additionally, students' discontent with technical difficulties experienced during online classes on digital platforms like Microsoft Teams, Cisco Webex, or the Zoom application has led to a pessimistic outlook on virtual mobility (Sandarenu, 2021). Students were also concerned about the costs of internet connectivity and data packages, online connectivity devices, and access to online resources (Sandarenu, 2021). Furthermore, long online lectures and insufficient break time caused students to become fatigued, and as a result, distractions developed among students (Koris et al., 2021).

Incorporating game-based activities becomes crucial in maintaining their interest and enhancing their participation to cater to the short attention spans of digital-native students. These activities foster engagement and contribute to developing important cognitive skills such as critical thinking, visual-spatial abilities, observation, and inquiry skills. Prensky (2001) suggested that activities should include components for reflection and critical thinking to facilitate the effective flow of information. Additionally, providing parallel discussion sessions allows students to engage in meaningful conversations and exchange ideas. However, it is important to note that remote learning requires students to be more self-directed in their studies, as it

often entails individual work. It necessitates students to dedicate additional time and effort to their learning endeavors.

Virtual mobility benefits students by eliminating the hassle of preparing the international travel requirements to the host country. It is flexible but faces some challenges, like a lack of IT skills, infrastructure support, and online interaction. Despite the challenges of implementing virtual mobility for a semester, short-term virtual mobility, where the duration is within a few days or weeks but less than three months, is an alternative to encourage cultural exchange and international collaboration. The process of planning virtual mobility is akin to planning physical mobility in terms of required efforts, with the only distinction being the utilization of technological tools in the former, whereas the latter involves face-to-face interactions. When designing short-term virtual mobility, the central focus is fostering digital interaction. A key question to address during the planning process is, "How can participants experience social interaction that emulates the sense of engagement found in physical activities?" Social constructivism is a learning theory that emphasizes knowledge development's social interaction and collaborative nature. It posits that students build knowledge and construct their understanding through their social and cultural experiences.

Short-term virtual mobility provides students with international exposure to strengthen their professional identities and develop cultural competencies, which leads to the following questions: How do students perceive the value of virtual mobility programs? What are the experiences reported by students after participating in virtual mobility programs? This study endeavors to shed light on these questions by exploring and investigating students' perceptions and experiences in the short-term virtual mobility program where they can engage with peers from different universities and countries to exchange ideas, share perspectives, and collectively explore cultural differences. This study postulated that social interaction could lead to students' positive experiences and perceptions of short-term virtual mobility program effectiveness. Guided by the research questions, two hypotheses were proposed: (1) students exhibit overall positive perceptions toward short-term virtual mobility programs, and (2) students will experience significant positive or negative effects on their engagement and comprehension during the short-term virtual mobility program.

Context of the Study

The collaboration between Malaysian universities and the universities in Indonesia and Thailand is rooted in an established partnership, fostering academic cooperation across borders. This collaboration is a short-term virtual mobility program designed to facilitate international exchange and collaboration. Throughout the year, participating universities engage in six planned activities meticulously crafted in consultation with partner institutions. These activities are focal points for collaborative

efforts, fostering meaningful interactions and knowledge exchange among students and faculty.

Central to the program's ethos is its alignment with Sustainable Development Goals (SDGs), ensuring that each series of activities tackles pressing global challenges from various angles. This thematic coherence provides a unifying framework and underscores the program's commitment to addressing real-world issues. Each student is paired with a mentor with expertise in the assigned SDG-related topic to enrich the learning experience. This mentorship ensures that participants receive guidance and support as they delve into their respective research areas, laying the groundwork for informed discussions and presentations.

During the virtual mobility program, students from diverse backgrounds and geographical locations converge in discussion groups, fostering cross-cultural understanding and collaboration. Participants collectively explore solutions to complex challenges through shared dialogue and presentations, leveraging their unique perspectives and expertise. Overall, this dynamic and collaborative approach to international education enriches participants' academic experience and cultivates a global mindset, empowering them to become agents of positive change in their communities and beyond.

METHODS

This study explores students' experiences with short-term virtual mobility programs and their perceptions of how these programs could enhance their future learning. A three-week short-term virtual mobility program was conducted using an online Zoom application platform in September 2021. The participants comprised 375 university students and staff from Indonesia, Malaysia, and Thailand. All participants were recruited as the sample of this study. An online survey questionnaire was employed to collect the data. The questionnaire comprises demographic information, nine questions adapted from training evaluation forms (Greenfield et al., 2015) to gauge students' perceptions, and an open-ended question to capture their experiences. The participants were required to answer the survey questionnaire at the end of the virtual mobility, and those who did not submit the survey were given one week to complete it. A reminder was sent three days after the events, and all responses were recorded at the end of one week.

The collected data were analyzed using descriptive statistics and thematic analysis. Thematic analysis is a useful and flexible method for qualitative research in and beyond psychology (Braun & Clarke, 2006), and it allows for a comprehensive understanding of students' experiences and perceptions within the context of virtual mobility programs.

RESULTS

Given the study's focus on students' experiences and perceptions of short-term virtual mobility, university staff were excluded from the sample. The final number of respondents for the study amounted to

361 students. The respondent's profile is presented in Table 1.

Table 1
Profile of respondents

Details	Total respondents	Total percentage (%)
Gender		
Male	70	19.4
Female	291	80.6
Country		
Indonesia	34	9.4
Malaysia	92	25.5
Thailand	235	65.1
First time participating in virtual mobility		
Yes	186	51.5
No	175	48.5
Channel to learn about this virtual mobility		
Social Media	106	29.4
University official channel	199	55.1
Friend recommendation	51	14.1
Other	5	1.4
Preference to the mode of mobility program		
Online/virtual	252	69.8
Physical/face-to-face	99	27.4
Hybrid/both online and physical	10	2.8

Source: Authors' work

The study revealed that female participation outweighed that of males, with 80.6% (291) of participants being female students and 19.4% (70) male students. Geographically, a significant proportion of respondents hailed from Thailand (65.1%, 235 students), followed by Malaysia (25.5%, 92 students), with a smaller cohort originating from Indonesia (9.4%, 34 students). Notably, most participants (51.5%) were newcomers to short-term virtual mobility programs, lacking prior experience in such initiatives, as illustrated in Table 1.

Further, about 55.1% of students recognized the international office's concerted efforts in promoting the short-term virtual mobility program through official university channels. Alternatively, 29.4% of participants discovered this virtual mobility program via social media channels, while 15.5% of students were introduced to it through friend recommendations and other sources.

Interestingly, most students (69.8%) preferred short-term mobility programs delivered online or virtually, whereas only 27.4% (99 students) favored face-to-face mobility, and a mere 2.8% leaned towards a hybrid approach. This inclination towards virtual modalities aligns with the evolving landscape of education, particularly amidst the pandemic, which has accelerated the adoption of online learning and collaboration platforms.

Students' perceptions of short-term virtual mobility programs were analyzed using descriptive analysis across nine statements, ranging from program organization to satisfaction with program content, and presented in Table 2. The data provided represents responses from 361 participants regarding their perceptions. Each item was rated on a 5-point Likert scale, ranging from "1 = Strongly Disagree" to "5 = Strongly Agree."

Most respondents (over 90%) either agreed or strongly agreed that the program was well organized, indicating a high level

Table 2
Student's perceptions of virtual mobility program

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1) This program was well organized.	1	3	24	128	205
2) The moderators facilitated the program effectively.	1	0	23	136	201
3) The keynote speaker provided information I can use to be a success.	2	1	31	144	183
4) This program has created an efficient learning environment.	4	1	27	156	173
5) The program is valuable.	1	0	30	128	202
6) This program improved my understanding of the topics.	2	4	28	147	180
7) Participating in this program increased my interest in the topics.	2	2	26	146	185
8) I can apply what I learned in this program.	3	2	30	152	174
9) I am satisfied with the program content.	1	2	25	157	176
					Total respondents: 361

Source: Authors' work

of satisfaction with how the program was structured and managed. Similarly, most respondents (over 90%) agreed or strongly agreed that the moderators effectively facilitated the program, suggesting that the facilitators played a crucial role in guiding the program activities. While most respondents still agreed that the keynote speaker provided useful information, there were slightly fewer strong agreements compared to the previous statements.

When asked about the efficient learning environment, a substantial number of respondents (over 80%) agreed or strongly agreed that the program created an efficient learning environment, indicating that the format and delivery of the program online were conducive. Few students did not agree that the virtual mobility program created an efficient learning environment,

probably because some students might have experienced technical difficulties or limitations during the virtual sessions, such as poor internet connectivity, audio/video issues, or platform usability issues. These technical challenges can disrupt the learning process and diminish the overall efficiency of the virtual environment.

A significant majority of respondents (over 90%) agreed or strongly agreed that the program was valuable, and over 80% of respondents agreed or strongly agreed that the program improved their understanding of the topics covered, suggesting that they perceived the importance of the program, and the content was impactful. Notably, a significant majority of respondents (over 80%) agreed or strongly agreed that participating in the program increased their interest in the topics discussed, indicating

that the program successfully engaged participants and stimulated their curiosity.

While still a majority, slightly fewer respondents (around 75%) agreed or strongly agreed that they could apply what they learned in the program, suggesting that there may be some room for improvement in terms of practical application. Even if students recognize the relevance of the content, they may feel that there are limited opportunities to apply what they learned in their current academic or professional contexts. Without clear guidance on transferring the knowledge into action, students may doubt their ability to apply what they learned.

Most respondents (over 80%) agreed or strongly agreed that they were satisfied with the program content, indicating overall positive perceptions among participants regarding various aspects of the short-term virtual mobility program.

Students' experiences with short-term virtual mobility were analyzed using thematic analysis. Thematic analysis is a powerful method for understanding experiences and thoughts across a data set (Braun & Clarke, 2006). This study allows participants to share their experiences in their own words in the open-ended question. About 334 students were provided written responses to record their positive or negative experiences with this short-term virtual mobility program.

The recorded responses were then transformed into a thematic network following the steps. Firstly, the responses are coded by breaking the text into digestible and coherent segments. The second step is to identify themes. In this step, organizing themes were formulated from the coded segment and categorized into two global themes. i.e., positive and negative experiences. The organizing themes were then subdivided into two global themes based on related conceptual content. Lastly, the thematic network was interpreted and presented in Table 3.

The findings from Table 3 highlight both the positive and negative experiences reported by students participating in the virtual mobility program. Many students expressed enjoyment and excitement in the learning process. They found the program to be engaging and stimulating, with activities that they found enjoyable. It showed that this virtual mobility program effectively captured students' interest and motivated them to participate actively. Further, students reported that they like to apply the knowledge gained from the program in real-life situations. It suggests that the content of the virtual mobility program was relevant and practical, which enabled students to transfer their learning to their daily lives. Nevertheless, students appreciated the opportunity to interact with peers from different countries and gain insights from international perspectives, meaning the students value the cultural exchange aspect of the program.

On the other hand, some students faced difficulties understanding English, which hindered their participation and comprehension of the program. In addition, technical issues, such as internet connectivity, were frustrating for some students. Poor internet connection disrupted

Table 3
Global themes for student's experiences toward virtual mobility

Organizing Theme	Description	Representative quote
Global Theme: Positive Experiences		
Fun Learning	Students described enjoyment and excitement in the learning process.	"I want to have this program again because it's so much fun. I like it." "I like everything because it is so exciting, and fun, and gives me more knowledge." "I like it when there is a dance to the song. It's fun."
Knowledge application	Students apply the knowledge they have learned in real life.	"I like everything about this program. It is an educational activity about improving mental health. I am able to apply the knowledge gained in my daily life and make my daily life happy." "I extremely like the content of this presentation because I can apply what I learned from it to use in real life."
International networking and knowledge	Students appreciate new friendships and knowledge from international perspectives.	"Interaction and sharing between students from various countries." "The things I like is that we can have new experiences and friends from another university." "This program makes me know what happened in my neighboring country during the COVID-19 pandemic."
Global Theme: Negative Experiences		
Language Barrier	Students described difficulty in understanding English.	"I wish there were more languages spoken because some people might not understand the language." "What I don't like is that I don't understand English." "Some words or sentences I do not understand because I am not very good at English."
Internet Connection problem	Students expressed frustration with the internet connection issue.	"As a digital tool, sometimes the connection is not good enough to hear what is being said." "I dislike the problem of the internet."

Source: Authors' work

their ability to fully engage in the program and affected their learning experience. Overall, while the virtual mobility program provided many positive experiences, there were also challenges related to language barriers and technical issues that need to be addressed to enhance such programs' overall effectiveness and inclusivity.

DISCUSSION

The findings of this study shed light on the student's perceptions and experiences of short-term virtual mobility programs, revealing valuable insights for both academic institutions and policymakers. Firstly, students' awareness of short-term virtual mobility programs, primarily through the university's official channels, indicates the importance of proactive promotion by university officers to promote such programs, whether long-term or short-term, physical or virtual. Given that the pandemic has accelerated the adoption of online academic mobility, students have reported a positive experience with short-term virtual mobility and intend to participate in the

future. Hence, a virtual mobility or hybrid program initiative needs to be sustained beyond COVID-19 (Yang et al., 2022).

Secondly, the aggregated results demonstrate overwhelmingly positive feelings regarding their short-term virtual mobility program participation. Most students, above 90%, perceived that this short-term virtual mobility program is well organized and that they benefit from it in terms of understanding the topics, increasing interest in learning, and learning in an efficient environment. They are pleased with this short-term virtual mobility program because it was valuable, and the keynote speaker provided useful information. Most students exhibit positive perceptions of the virtual mobility program, indicating high satisfaction with various aspects of the program beyond just the content.

Besides that, students also shared their experiences with this short-term virtual mobility program. The experiences were coded and aggregated into a global theme (Braun & Clarke, 2006). There are three organizing themes for good experiences and two for bad experiences. Students reported they had much fun learning new information, and some even requested that a similar program be implemented. Furthermore, students valued the knowledge gained during the program because it applies to their daily lives. Most importantly, students have the opportunity to broaden their networking to other countries, make new friends, and stay up to date on the latest information in the neighboring countries. The experiences of mobility students in short-term virtual

mobility programs promote personal and professional development by strengthening their academic courses, cultural awareness, and digital literacy, which prepare them for the labor market in the future.

Notwithstanding, several students raised the possibility of a language barrier and admitted to a lack of English proficiency. The English language used throughout the program was a challenge for their understanding and communication failures. One possible solution could be to provide translators fluent in Thai and Indonesian in the future with the same mobility program when English is the primary medium of communication to avoid this negative effect. Despite technological advancements such as virtual meeting technology, the internet's stability is one factor that disrupts the program's effectiveness. Students are dissatisfied with the internet connection problem, which causes them to be 'kicked out' of the online program regularly. In other words, students with negative online learning experiences are more inclined to perceive virtual mobility programs negatively and vice versa.

The findings of this study provide empirical evidence that short-term virtual mobility programs can play a significant role, just like physical mobility. Participating in a short-term virtual mobility program can influence awareness of cultural diversity, enhance intercultural competence, encourage personal and professional development, promote diversification, build character and attitudes, as well as develop social and communication skills

(Blankvoort et al., 2019; Slipchuk et al., 2021; Yang et al., 2022). It has also been demonstrated that the short-term virtual mobility program fosters an effective learning environment and has aided in acquiring new learning environments and knowledge (Helm, 2019; Koris et al., 2021). Indeed, as international students share their culture with other students, knowledge sharing has the potential to effectively support the lifelong learning process. It is also reflected as more enjoyable while providing significant benefits outside the typical classroom setting.

Short-term virtual mobility programs are a useful tool for universities to promote internationalization in higher education, as highlighted by Knight (2012), Munusamy and Hashim (2019), Mustapha et al. (2021), and Yee et al. (2018). By offering such programs, universities can attract students worldwide, regardless of their location or travel restrictions. It provides educational opportunities to a wider audience and enhances the institution's link to the global community.

One of the major advantages of short-term virtual mobility programs is their flexibility, making them more appealing to students. With such programs, students can continue their studies while participating in academic mobility programs in host countries (Slipchuk et al., 2021). This exposure to different cultures, languages, and ways of thinking can help students become more globally aware, preparing them better for the global workforce. Moreover, students can gain valuable international experience

and cultural exchange without committing to long-term physical mobility, making it an attractive option for many.

Another benefit of short-term virtual mobility programs is their use of technology to create immersive learning experiences, such as virtual reality or gamification. It can help students engage with their learning and connect with others, building networks supporting their future career opportunities. Furthermore, virtual mobility programs are more cost-effective than traditional physical mobility programs, making international experiences more accessible to a wider range of students, including those from low-income backgrounds. Hence, it is important to note that virtual mobility programs should not be abandoned after the pandemic. In fact, Yang et al. (2022) also suggest incorporating virtual mobility with physical mobility to provide online and offline academic and cultural activities.

Implications of Findings

This study contributes to the understanding of student's experiences and perceptions of virtual mobility programs by extending the current literature and providing insight for higher education institutions when planning for the internationalization agenda. The findings serve as a blueprint for institutions globally to design impactful international mobility programs focusing on students' experiences with short-term virtual mobility. Highlighting the lack of integration between universities under a Memorandum of Understanding (MOU) underscores the necessity for well-defined international

student recruitment policies. It is particularly important due to the imbalance in the number of participants from Indonesia and Malaysia compared to Thailand. Higher education institutions should invest more time in developing virtual mobility programs incorporating interactive activities and meaningful engagement to maximize student satisfaction and learning outcomes (Blankvoort et al., 2019; Koris et al., 2021). It includes investing in technological infrastructure and creating support systems to address the specific needs of students engaging in virtual mobility. It is important to ensure the opportunity for collaboration among mobility students and provide them with enriching educational experiences. Higher education institutions can also leverage these findings to promote virtual mobility programs as an alternative for students who face financial constraints or travel limitations for physical mobility.

CONCLUSION

This study provides a snapshot of Southeast Asian students' perceptions and experiences of a short-term virtual mobility program in the new normal. The findings highlight the potentially beneficial role that a short-term virtual mobility program could play in developing students' educational experiences in terms of personal and professional development and cultural awareness. Despite some difficulties, such as adjusting to the language and connection problems with the online program, some aspects could be better systematized, particularly in terms of the attractiveness

of mobility programs. Policymakers are encouraged to refine international student recruitment policies and integrate intercultural learning into virtual mobility curricula, fostering engagement with diverse perspectives and strengthening educational experiences. More student-friendly online and hybrid education can be suggested in the long run. While doing so, ensuring mutual collaboration and strengthening interinstitutional relationships is critical.

Recommendation

The COVID-19 pandemic has accelerated the adoption of digital education in higher institutions, marking a shift toward a more digital post-pandemic era. The development of virtual mobility programs, catalyzed by this transition, presents a significant challenge and an opportunity for innovative educational approaches. Embracing technological advancements, universities can ensure students' connection and engagement with their institutions while promoting internationalization. Universities must sustain virtual mobility programs beyond the pandemic, incorporating them into curriculum development to attract international students and bolster their internationalization efforts. This trajectory also offers a promising avenue for future research on the interplay between physical and virtual mobility in higher education.

Limitations and Future Research

The sample size of this study may not be sufficient to draw definitive conclusions about the virtual mobility program. Future

research may include a large sample size and international mobility students from the Middle East and European countries. Further, students' discontent with technical difficulties experienced during online classes on digital platforms like Microsoft Teams, Cisco Webex, or the Zoom application has led to a pessimistic outlook on virtual mobility. Future research may delve deeper into the different platforms, pedagogical approaches, training and development within virtual mobility programs and also consider the specific aspects of virtual mobility, such as the impact of virtual mobility on students' academic performance and long-term career outcomes and the impact of virtual mobility on higher education institution's rankings.

ACKNOWLEDGEMENT

REFERENCES

Baranowski, M., & Jabkowski, P. (2023). On virtual mobility in three Central European universities: Similar but different? *Knowledge Cultures, 11*(3), 82-100. https://doi.org/10.22381/kc11320235

Blankvoort, N., Kaelin, V. C., Poerbodipoero, S., & Guidetti, S. (2019). Higher education students' experiences of a short-term international program: Exploring cultural competency and professional development. *Educational Research, 61*(3), 356-370. https://doi.org/10.1080/00131881.2019.1632725

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77-101. https://doi.org/10.1191/1478088706qp063oa

Chan, H. K., Yaakob, A. R., & Pinjaman, S. (2021). Effect of support from lecturers and host country nationals on cross-cultural adjustment among international students: Evidence from a Malaysia university. *Pertanika Journal of Social Science & Humanities, 29*(2), 1213-1229. https://doi.org/10.47836/pjssh.29.2.25

Cheng, Y., Yeoh, B., & Yang, P. (2023). Virtual student mobility on Zoom: Digital platforms and differentiated experiences of international education and (im)mobilities in a time of pandemic. *Mobilities, 18*(5), 839-854. https://doi.org/10.1080/17450101.2023.2209824

Cioltan-Drăghiciu, A., & Stanciu, D. (2020). Virtual exchange: Romania and Hungary 100 years later. In F. Helm & A. Beaven (Eds.), *Designing and implementing virtual exchange: A collection of case studies* (pp. 37-46). Research-publishing.net. https://doi.org/10.14705/rpnet.2020.45.1113

De Wit, H. (2020). Internationalization of higher education. *Journal of International Students, 10*(1), i-iv. https://doi.org/10.32674/jis.v10i1.1893

Dixon, D., & Tahmaz, O. (2020). Intercultural competence for youth workers. In F. Helm & A. Beaven (Eds.), *Designing and implementing virtual exchange: A collection of case studies* (pp. 231-240). Research-publishing.net. https://doi.org/10.14705/rpnet.2020.45.1129

Enkhtur, A., Li, M., & Zhang, X. (2023). Virtual student mobility from student's perspectives: Case study from Japan. *Higher Education Forum, 20,* 59-78. https://doi.org/10.15027/53851

Fernandes, A. J., Shukla, B., & Fardoun, H. (2022). Fostering reputation of higher education institutions in international ranking by means of diversity of international collaboration. *Information Sciences Letters, 11*(4), 1137-1143. https://doi.org/10.18576/isl/110416

Greenfield, V. A., Shelton, S. R., Balkovich, E., Davis II, J. S., & Adamson, D. M. (2015). The federal voting assistance program and the road ahead: Achieving institutional change through analysis and collaborating (Report No. RR-882-FVAP). RAND National Defense Research Institute. https://www.rand.org/pubs/research_reports/RR882.html

Heirweg, S., Carette, L., Ascari, A., & Hove, G. V. (2020). Study abroad programmes for all? Barriers to participation in international mobility programmes. *International Journal of Disability, Development and Education*, 1-19. https://doi.org/10.1080/1034912X.2019.1640865

Helm, F. (2019). Languages and international virtual exchange: Introduction to the special issue. *European Journal of Language Policy, 11*(2), 139-142. https://doi.org/10.3828/ejlp.2019.9

Knight, J. (2008). An internationalization model: Meaning, rationales, approaches and strategies. In J. Knight (Ed.), *Higher education in turmoil: The changing world of internationalisation* (pp. 19-37). Sense Publishers. https://doi.org/10.1163/9789087905224_003

Knight, J. (2012). Student mobility and internationalization: Trends and tribulations. *Research in Comparative and International Education, 7*(1), 20-33. https://doi.org/10.2304/rcie.2012.7.1.20

Koris, R., Mato-Díaz, F. J., & Hernández-Nanclares, N. (2021). From real to virtual mobility: Erasmus students' transition to online learning amid the COVID-19 crisis. *European Educational Research Journal, 20*(4), 463-478. https://doi.org/10.1177/14749041211021247

Marinoni, G., Land, H. V., & Jensen, T. (2020). *The impact of COVID-19 on higher education around the world.* The International Association of Universities Publications.

Ministry of Higher Education. (2015). *Malaysia Education Blueprint 2015-2025 (Higher Education).* https://jpt.mohe.gov.my/portal/index.php/en/corporate/policy-documents/16-malaysia-education-development-plan-2015-2025

Munusamy, M. M., & Hashim, A. (2019). Internationalisation of higher education in Malaysia: Insights from higher education administrators. *AEI Insights, 5*(1), 21-29.

Mustapha, S. M., Abdullah, N., Devarajoo, K., Ibrahim, F., Suid, S. H., & Amirrudin, S. (2021). Embracing internationalization in gearing Malaysian higher education towards global education. *Asian Journal of University Education, 17*(4), 132-147. https://doi.org/10.24191/ajue.v17i4.16224

Prensky, M. (2001). Digital natives, digital immigrants Part 1. *On the Horizon, 9*(5), 1-6. https://doi.org/10.1108/10748120110424816

Queiroz, F. C. B. P., da Silva, C. L., Lima, N. C., Thomaz, S. M., Queiroz, J. V. (2023). Virtual mobility in higher education in Brazil: Factors of influence on satisfaction and continuity. *Higher Education, Skills and Work-based Learning, 13*(6), 1354-1367. https://doi.org/10.1108/HESWBL-02-2023-0039

Rahim, N. A. (2021). Adjustment experiences of international graduate students to life in the United States. *Pertanika Journal of Social Sciences & Humanities, 29*(4), 2243-2262. https://doi.org/10.47836/pjssh.29.4.08

Recio, S. G., & Colella, C. (2020). *The world of higher education after COVID-19: How COVID-19 has affected young universities.* Young European Research Universities.

Rostovskaya, T. K., Skorobogatova, V. I., Pismennaya, E. E., & Bezverbny, V. A. (2021). Incoming and outgoing academic mobility in Russia and abroad: Main trends, administrative challenges. *International Journal of Sociology and Social Policy, 41*(1/2), 154-166. https://doi.org/10.1108/IJSSP-03-2020-0091

Ryazantsev, S. V., Rostovskaya, T. P., Skorobogatova, V., & Bezverbny, V. A. (2019). International academic mobility in Russia: Trends, types, state stimulation. *Ekonomika Regiona, 2*, 420-435. https://doi.org/10.17059/2019-2-9

Sandarenu, K. (2021). University student views on the impact of COVID-19 pandemic on their university life and education in Sri Lanka. *International Journal of Research Publications, 71*(1), 18-32. https://doi.org/100711220211753

Slipchuk, A., Yuzkiv, H., Batechko, N., Pisotska, M., & Klymenko, L. (2021). Academic mobility development among university students during COVID-19 pandemic. *International Journal of Health Sciences, 5*(3), 573-583. https://doi.org/10.53730/ijhs.v5n3.2490

Spencer-Oatey, H., & Dauber, D. (2019). Internationalisation and student diversity: How far are the opportunity benefits being perceived and exploited? *Higher Education, 78*, 1035-1058. https://doi.org/10.1007/s10734-019-00386-4

Teräs, M., Suoranta, J., Teräs, H., & Curcher, M. (2020). Post-Covid-19 education and education technology 'Solutionism': A seller's market. *Postdigital Science and Education, 2*, 863-878. https://doi.org/10.1108/10748120110424816

White, B., & Lee, J. J. (2020, April 18). The future of international HE in a post-mobility world. *University World News*. https://www.universityworldnews.com/post.php?story=20200417105255362

Woldegiorgis, E. T., & Doevenspeck, M. (2015). Current trends, challenges and prospects of student mobility in the African higher education landscape. *International Journal of Higher Education, 4*(2),105-115. https://doi.org/10.5430/ijhe.v4n2p105

Yang, Q., Shen, J., & Xu, Y. (2022). Changes in international student mobility amid the COVID-19 pandemic and response in the China context. *Fudan Journal of the Humanities and Social Sciences, 15*, 23-40. https://doi.org/10.1007/s40647-021-00333-7

Yee, C. P., Yean, T. S., & Yi, A. K, J. (2018). Verifying international students' satisfaction framework for the development of MISS-model in Malaysia. *Pertanika Journal of Social Sciences & Humanities, 26*(1), 1-18.

Yildirim, S., Bostanci, S. H., Yildirim, D. G., & Erdogan, F. (2021). Rethinking mobility of international university students during COVID-19 pandemic. *Higher Education Evaluation and Development, 15*(2), 98-113. https://doi.org/10.1108/HEED-01-2021-0014

SOCIAL SCIENCES & HUMANITIES

The Effects of Medicine Ball Training on Batting Velocity in Female Collegiate Softball Players

Maisarah Shari[1], Wan Ahmad Munsif Wan Pa[2], Mohamad Firdaus Ahmad[3], Ellail Ain Mohd Aznan[4], Ali Md Nadzalan[5], Noor Azila Azreen Md Radzi[6], Nurul Ain Abu Kasim[6] and Raja Nurul Jannat Raja Hussain[6]*

[1]Department of Sports Science, Faculty of Sports Science and Recreation, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

[2]Department of Education and Community Welfare Studies, Faculty of Educational, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

[3]Department of Sports Management, Faculty of Sports Science and Recreation, Universiti Teknologi MARA, 70300 Seremban, Negeri Sembilan, Malaysia

[4]Department of Sports Management, Faculty of Sports Science and Recreation, Universiti Teknologi MARA, 02600 Arau, Perlis, Malaysia

[5]Department of Coaching Science, Faculty of Sports Science and Coaching, Universiti Pendidikan Sultan Idris, 35900 Tanjung Malim, Perak, Malaysia

[6] Department of Sports Science, Faculty of Sports Science and Recreation, Universiti Teknologi MARA, 70300 Seremban, Negeri Sembilan, Malaysia

ABSTRACT

Softball batting performance, characterised by bat velocity and rotational strength, is crucial for competitive success. This study explores the effect of medicine ball training, a widely used yet understudied method, on these key performance metrics. This study examines the impact of an 8-week medicine ball training program on torso rotation strength and batting velocity in female collegiate softball players. Forty participants were divided into a control group and a medicine ball group, performing 100 bat swings thrice weekly. Additionally, the medicine ball group undertook full-body medicine ball exercises thrice weekly. Results showed a significant main effect, $F(1, 38) = 5.00$, $p = 0.03$, partial eta squared = 0.12 for torso rotation strength test and $F(1, 38) = 4.01$, $p = 0.05$, partial eta squared = 0.10 for batting velocity, signifying there is a difference in the effectiveness of the two training approaches. Pre- and post-intervention tests showed significant improvements in

E-mail addresses:
maisarahshari@uitm.edu.my (Maisarah Shari)
munsif@ukm.edu.my (Wan Ahmad Munsif Wan Pa)
firdaus466@uitm.edu.my (Mohamad Firdaus Ahmad)
ellailain@uitm.edu.my (Ellail Ain Mohd Aznan)
ali.nadzalan@fsskj.upsi.edu.my (Ali Md Nadzalan)
azila_azreen7172@uitm.edu.my (Noor Azila Azreen Md Radzi)
nurulain755@uitm.edu.my (Nurul Ain Abu Kasim)
nuruljannat@uitm.edu.my (Raja Nurul Jannat Raja Hussain)
* Corresponding author

Maisarah Shari, Wan Ahmad Munsif Wan Pa, Mohamad Firdaus Ahmad, Ellail Ain Mohd Aznan, Ali Md Nadzalan, Noor Azila Azreen Md Radzi, Nurul Ain Abu Kasim and Raja Nurul Jannat Raja Hussain

both groups, with the medicine ball group exhibiting notably greater enhancements. Torso rotation strength and batting velocity increased significantly in the medicine ball group, demonstrating large effect sizes (Cohen's d=1.15 and Cohen's d=1.17, respectively). In contrast, the control group showed a substantial improvement in torso rotation strength (Cohen's d=0.86) but a minor increase in batting velocity (Cohen's d=0.22). These results suggest that medicine ball training effectively boosts key performance metrics in female softball players, offering valuable insights for designing more effective athletic training programmes.

Keywords: Batting velocity, collegiate female, medicine ball, softball, torso rotation strength

INTRODUCTION

Softball is a sport that demands a high level of competence in a range of skill sets, including throwing, batting, and fielding. These skills necessitate focused attention to optimise performance. Within the context of player skills, batting stands out as an area that demands considerable attention, focusing on enhancing bat velocity (Hussain, Mea, et al., 2019; Hussain & Shari, 2021). Various elements must be considered. These include swift awareness, reactive reflexes, proper swing technique, and bat velocity to excel in softball batting (Rivera et al., 2018). Batting in softball is characterised by a complex kinetic link, as discussed in previous research (Hussain, Mea, et al., 2019; Larson, 2021). To excel in batting, one must possess

physical strength and a deep understanding of the intricate mechanics involved in the swinging technique. Enhancing rotational strength via targeted training can improve bat speed, directly influencing overall batting performance.

Extensive research has been conducted on using medicine balls in baseball and other swing- and throw-related sports. Research has shown that using weighted baseball training programmes can improve pitch velocity, arm biomechanics, and passive range of motion while decreasing the likelihood of injuries (Reinold et al., 2018). Furthermore, research has shown that implementing acute weighted baseball throwing programmes can substantially affect shoulder range of motion and elbow varus torque, suggesting their profound impact on the biomechanics of the throwing motion (Reinold et al., 2020). Moreover, research has shown that incorporating medicine ball training into tennis players' workout regimens can significantly improve their upper-body explosive strength. This finding underscores the efficacy of this training method in enhancing athletic performance, particularly in sports that involve swinging or throwing motions (Kong & Xu, 2023).

In baseball, research has shown that training with weighted baseballs can increase arm speed without significantly increasing elbow and shoulder joint kinetics, implying a potential for performance improvement without putting undue strain on the throwing arm (O'Connell et al., 2022). In addition, research has demonstrated

the beneficial impact of medicine ball throws on ball velocity and accuracy among young competitive tennis players. It further reinforces the effectiveness of medicine ball training in improving sports performance, as highlighted by Terraza-Rebollo and Baiget (2021). The necessity of conducting a new study on medicine ball training in softball arises from the sport's distinct biomechanics and skill demands, despite the extensive documentation of its benefits in baseball and other sports. Softball players exhibit unique hitting and throwing techniques, potentially resulting in different effects of medicine ball training on their overall performance. Conducting studies with softball presents an opportunity to develop training programmes specifically designed to meet the biomechanics and skill demands of the sport. It ensures that the training methods are tailored to the unique requirements of softball, optimising their effectiveness. Furthermore, exploring the impact of medicine ball training on softball players can greatly contribute to advancing sport-specific training protocols, thereby enriching the collective knowledge of training methodologies in softball.

Studies showing improved athletic performance have supported using medicine balls in baseball and other swing- and throw-related sports. The effectiveness of medicine ball training has been well-documented in different sports. However, conducting a specific study on softball is crucial due to its distinct biomechanics and skill requirements. It emphasises the significance of designing training programmes catering to softball's demands. Despite their frequent incorporation into training regimens, more research needs to be conducted on the impact of medicine ball exercises, particularly in the context of female softball players. Our study seeks to fill a void in existing research by examining the effects of medicine ball training on bat swing velocity in female collegiate softball players. By doing so, we aim to enhance our knowledge of effective training techniques in this sport.

MATERIALS AND METHODS

Research Area and Duration

This research was conducted to determine the effect of eight-week medicine ball training on improving female collegiate softball players' torso strength and batting velocity. The research area occurred at the Universiti of Teknologi MARA Selangor, Shah Alam campus, Malaysia.

Research Design

The study employed an experimental design, incorporating a pre-test and post-test randomised design.

Participants

The sample size for this study was determined through a thorough calculation of the effect size, utilising the formula proposed by Cohen (1992). According to the findings of a previous study by Higuchi et al. (2013), this method has a large effect size (0.75) on batting velocity, as seen in 12 participants. In a study conducted by Schoenfeld et al. (2014), a

group of 10 participants was examined, revealing a moderate effect size (0.56) in batting velocity. Twenty of the sixty female softball players who initially volunteered for the study were excluded because of predetermined inclusion requirements. Forty healthy female athletes were left in the final cohort. The participants in the study had an average age of 23.32 years (±1.87), a height of 156.42 cm (±1.69), and a weight of 53.49 kg (±3.06). The following were the requirements for inclusion in the study: (1) having a valid collegiate softball team roster, (2) having prior experience with resistance training, and (3) not having any self-reported medical conditions, neurological disorders, mental illnesses, or significant past or present injuries that could put participants at risk during training or exercises. A bat speed measurement was performed on each participant to guarantee baseline comparability. Participants were randomly assigned to the Medicine Ball (MB) or Control (CT) group by fish-bowl drawing, with no initial batting velocity differences. The methodological approach utilised in this study received approval from the Research Ethics Committee of Universiti Teknologi MARA (600-RMI 5/1/6), underscoring the study's commitment to upholding ethical research standards.

Instrument

Torso Rotation Strength Test

The Cybex Torso Rotation Machine™ has been widely utilised for conducting 3RM torso rotation strength tests, as demonstrated in the study by Szymanski et al. in 2010. Before the 3RM rotational strength test, participants were instructed to engage in standing torso exercises with a lightweight medicine ball as a warm-up. After the warm-up, participants were provided with a brief one-minute rest period. The participants assumed a seated position on the machine and gently positioned their feet on the footplates. They applied pressure to the machine's adductor pads to ensure their knees were at a 90° angle. The post-test involved recording the number of seat adjustments made by each participant. Participants assumed an upright position, firmly grasping the handles with both hands and applying pressure to their chests against the machine's chest pad. The study aimed to evaluate the participants' strength on the dominant side of their torso, specifically focusing on the direction of their bat swing. Before the test, each participant performed a single bat swing. Participants were instructed to execute rotations within a range of motion that closely mimicked the movements involved in a softball swing to assess the strength of their torso rotation. This range spanned from -30º to +75º, resulting in a total range of 105º. This range was established utilising the Cybex Torso Rotation Machine™ and has been used in softball and baseball batting experiments (Szymanski et al., 2010). Each participant's movement during the test was carefully monitored and guided by a qualified strength and conditioning instructor. Participants were given a 2-minute rest after failing to complete the full range of motion (105°) or perform the

test properly before trying again with a 2.3-kg load. The test-retest reliability for these variables was statistically significant, with a correlation coefficient greater than 0.90 ($p \leq 0.05$).

Batting Velocity

After conducting the 3RM Torso rotation strength test, a 30-minute break was provided before proceeding with the batting velocity test. Before assessing the bat swing, a comprehensive warm-up routine lasting 5 minutes was implemented, consisting of intense full-body exercises. Five dry swings were conducted using the designated softball bat for the upcoming testing as part of the warm-up routine. The bat utilised during the testing phase was constructed from aluminium, with a length of 84 cm and a weight of 0.68 kg. The bat swing velocity was measured using a portable swing analyser (ZEP-BT-000002; Zepp Company, Cupertino, California, USA). According to a study conducted by Hussain, Kee, et al. (2019), this swing analyser has been found to have reliability ranging from $r = 0.822$ to 0.988. The participants performed a series of five swings, with a brief 30-second rest period between each swing. The participants were instructed to execute rapid swings while ensuring a consistent stance and technique. Participants were asked to focus on the external environment rather than their bodily movements to optimise performance (Gray, 2018). The velocities of each of the five swings were meticulously recorded for analysis.

Procedure

Participants were required to attend a one-day session for assessment. The initial evaluation involved measuring body composition, including height, weight, and body mass index, as well as conducting a 3-Repetition Maximum (3RM) torso rotation strength test. The pre-test for batting velocity was conducted at the same location, with a 30-minute rest period following the 3RM torso rotation strength test. Participants were instructed to abstain from high-intensity physical activity before the test session. All the tests were recorded for each participant—the MB group was involved in regular practice, performing 100 dry swings three times a week. Nevertheless, the group engaged in medicine ball training in the gymnasium following the completion of swing training. The training was structured using the stepwise periodised method, as seen in previous studies on resistance training for enhancing strength and batting velocity (Szymanski et al., 2007). The training regimen began with a high volume of training at a low intensity (300 kg), gradually transitioning to a lower volume of training at a higher intensity (320 kg), as illustrated in Table 1.

During each training session, participants engaged in a sequence of exercises, completing a predetermined number of repetitions. The exercises mentioned encompassed various throwing techniques, such as the hitter's throw, figure-of-eight throw, speed rotation, standing side throw, and squat and hurl. The throw executed by the hitter follows a conventional

Table 1
Medicine ball training program

Variables	Week 1 to Week 4			Week 5 to Week 8		
	Sets	Reps		Sets	Reps	
Warm-up	2	10		2	10	
Swing practice	5	20		5	20	
Medicine Ball Exercise	Sets	Reps	Load (kg)	Sets	Reps	Load (kg)
Hitters throw	2	6	5	2	8	4
Standing figure of eight	2	6	5	2	8	4
Speed rotation	2	6	5	2	8	4
Standing side throw	2	6	5	2	8	4
Squat and hurl	2	6	5	2	8	4

Source: Szymanski et al. (2007)

batting stance, resembling the motion of swinging a softball bat while incorporating a medicine ball. The figure-of-eight throw involves a dynamic exchange of medicine balls between two participants positioned back-to-back, requiring them to rotate swiftly. The partner pass entails swiftly rotating the torso and throwing a medicine ball to a designated partner. A notable exercise is the standing side throw, in which individuals vigorously propel a medicine ball following a swift rotation, emulating a batter's stance. The rotational exercises involving the medicine ball were executed in both directions to promote equilibrium and enhance the development of the muscles in the torso. The squat and hurl exercise involves participants squatting down and rapidly extending their arms to throw the ball forcefully. In week five, modifications were implemented to replicate the conditions of a game by adjusting the weight and repetitions of the medicine ball by transitioning from heavier to lighter balls and increasing the number of repetitions. The CT group utilised the training programme Szymanski et al.

(2009) developed for the baseball swing. The batting velocity of baseball players experienced a notable increase due to the implementation of this programme.

Furthermore, the instructors utilise this programme during their regular softball practice sessions. Participants performed several sets of swing training, with each exercise session commencing with a warm-up consisting of two sets of ten swings using a standard bat. Next, the participants are required to complete five sets of 20 swings. During the eight-week training period, participants consistently stuck with the bat they had chosen.

Statistical Analysis

Before proceeding with the analysis, a Kolmogorov-Smirnov test was conducted to evaluate the normality of the data distribution. The findings revealed that all parameters exhibited a normal distribution. This study used a mixed within-group analysis to examine the changes in the 3RM torso rotational strength test and

batting velocity before and after an eight-week intervention. A pairwise comparison analysis was used to compare the pre-and post-test results for each variable. A test was conducted to validate the assumption of homogenous variances. The data was analysed using SPSS 23 software (IBM®, Armonk, NY, USA), with an alpha level of $p \leq 0.05$ to determine statistical significance.

RESULTS AND DISCUSSION

An analysis of variance was conducted to evaluate the effects of two interventions (medicine ball and bat swing) on participants' scores on the torso rotation strength test and batting velocity test over two periods (pre-intervention and post-intervention). The analysis revealed no significant interaction between training and time for the torso rotation strength test and batting velocity (Wilks' Lambda =0.59, $F(1, 38) = 26.43$, p =0.15, partial eta squared =0.42). Similarly, no significant effect was found for the same variables when considering Wilks' Lambda = 1.05, $F(1, 38) = 322.47$, p =0.16, and partial eta squared =0.89. The results

revealed a significant main effect for time, indicating a substantial increase in torso rotation strength test and batting velocity scores across the two periods (Table 2). The statistical analysis showed a strong effect size, with Wilks' Lambda =0.16, $F(1, 38) = 199.91$, p <0.001, partial eta squared =0.84 for the torso rotation strength test, and Wilks' Lambda =0.06, $F(1, 38) = 642.79$, p <0.001, partial eta squared =0.94 for batting velocity scores. The study's results revealed a significant main effect when comparing the two types of interventions. For the torso rotation strength test, the F-value was 5.00 (p =0.03, partial eta squared =0.12), while for the other test, the F-value was 4.01 (p =0.05, partial eta squared =0.10). These findings suggest a notable difference in the effectiveness of the two training approaches. Statistical analyses through pairwise comparisons indicated a significant enhancement in torso rotation strength for the Medicine Ball (MB) group when contrasted with pre-intervention values (p <0.001), exhibiting a large effect size (Cohen's d=1.15). A parallel increase was

Table 2

3RM Torso rotation strength test and batting velocity test scores for the medicine ball and control groups across two time periods

Parameters	Group		Mean	Std. deviation	N
Torso Rotation Strength Test	MB	Pre	30.36	7.16	20
		Post	39.21	8.23	20
	CT	Pre	28.29	5.16	20
		Post	32.42	4.39	20
Batting Velocity	MB	Pre	87.95	6.47	20
		Post	95.48	6.40	20
	CT	Pre	87.22	5.87	20
		Post	88.51	5.58	20

Source: Authors' work

observed in batting velocity, displaying a large effect size (Cohen's d=1.17). In the Control (CT) group, pairwise comparisons denoted a notable improvement in torso rotation strength compared to baseline measurements ($p < 0.001$), with a large effect size (Cohen's d=0.86). Conversely, the alterations in batting velocity from pre- to post-intervention for the CT group were associated with a small effect size (Cohen's d=0.22).

The present study utilised a between-subjects design to examine and compare the bat swing velocity of the medicine ball (MB) and the control (CT) groups by investigating the impact of medicine ball training, a specialised form of resistance training, on the bat speed of collegiate softball players. The experimental design employed in this research allowed for a thorough examination of this relationship. Previous studies have predominantly centred around training programmes to enhance the bat speed of high school and collegiate baseball players. The training programmes discussed in the literature incorporate various sport-specific resistance training methods, including weighted implement warm-ups, weighted bat training, and total-body resistance training. These programmes have yielded favourable outcomes in enhancing batting velocity (Kobak et al., 2018; Rivera et al., 2018). The current study found that the MB group demonstrated significantly greater improvements in the 3RM torso rotation strength test (%) and batting velocity (%) compared to the CT group. The results of this study were anticipated, as the MB group

incorporated a structured medicine ball programme consisting of periodic rotations thrice a week for eight weeks. It is worth noting that the MB group also maintained the same number of total bat swings as the CT group.

In contrast to not incorporating medicine ball exercises, the CT group demonstrated a modest improvement in the 3RM torso rotation strength test and batting velocity. Felion and DeBeliso (2020) discovered that engaging in a daily routine of swinging a bat 100 times three days a week for six weeks significantly improved bat swing velocity. According to Szymanski et al. (2009), the relationship between strength and power and its impact on the batting velocity of adolescent, high school, and collegiate baseball players has been explored in 11 different studies. Based on the analysis, it is evident that enhancing strength and power directly correlates with increased batting velocity. Power is the result of combining strength and velocity. Power can be defined as multiplying the applied force by the applied velocity. The correlation between force and velocity is crucial in determining muscular power output. An inverse relationship exists between the force-generated quantity and application velocity.

The relationship between force and velocity presents a trade-off. Muscle strength and speed have a positive correlation, although this relationship has limits. As a result, the highest level of muscle strength is achieved by effectively combining force and velocity. The findings indicate that the

robust and fluid movements of the kinetic link components, namely the hips and thorax, significantly impact increasing bat speed. The findings of this study align with the research discussed in a previous review article (Szymanski et al., 2009). The results of this study demonstrate a significant improvement in both the torso rotation strength test and batting velocity following strength-power (medicine ball) training. The post-test results of torso rotation strength considerably impact the batting velocity, with a substantial effect size exceeding 1.0. The study highlights two key factors contributing to the increased torso rotation strength test. Firstly, the use of a standard softball bat as a resistance tool proves to be effective.

Additionally, engaging in medicine ball training also plays a significant role in enhancing strength. The study participants performed 100 swings with a standard softball bat daily, thrice a week, over eight weeks. In addition, a stepwise medicine ball training programme was implemented three days a week for eight weeks. The study's findings suggest a significant enhancement in torso rotation strength among female collegiate softball players due to engaging in this combination of activities. Medicine ball training has been widely utilised in the athletic community for its effectiveness as a targeted power training technique (Kobak et al., 2018). The current study suggests that the movements performed with a medicine ball resemble those utilised in rotational power sports. The primary goal of these sports is to efficiently transfer the highest possible rotational speed to an external object through a combination of large and precise body movements. Medicine ball throws have been recognised in previous research as an effective method of resistance training due to their capacity to enhance speed and power compared to conventional athletic activities (Bragazzi et al., 2020).

This study explored the effects of increased force on power generation at a particular point on the force-velocity curve. The findings of this research could be valuable for athletes, as they shed light on the significance of powerful high-velocity movements that are often neglected in sports and resistance training programmes. The medicine ball training programme was conducted in a manner that closely replicated explosive, rotating softball-specific batting actions. The programme followed a ballistic and sequential approach to achieve this simulation. While striking a ball, the trunk flexors exert significant effort to rotate and stabilise the spine. Consequently, throughout swing practice and medicine ball training, the trunk flexors of softball players may undergo numerous contractions. The softball swing encompasses a circular kinetic link motion, engaging various body components such as the legs, trunk, shoulders, and arms. Therefore, it is imperative to implement sport-specific routines to replicate the swing accurately. To optimise bat swing velocity, it may be beneficial to incorporate methods such as adding resistance to the bat or engaging in explosive medicine ball throws.

The current study investigated the impact of an 8-week progressive medicine

ball training regimen on the strength of torso rotation and batting velocity in female collegiate softball athletes. The study's results indicated a notable enhancement in torso rotation strength and batting velocity due to this training method. These findings underscore the advantages of integrating medicine ball exercises into softball training programmes. The findings of this study provide valuable insights into the effects of movement-specific resistance training on softball performance, enhancing our understanding of this training approach. The swift rotation of the hips and torso greatly influences the effectiveness of softball actions. The body's role as a kinetic link is paramount in enhancing both batting and throwing skills. Incorporating medicine ball exercises into training regimens can greatly improve softball players' explosive rotational power while fortifying their core muscles through various movement patterns. The training method presented in this study provides a practical and adaptable strategy for improving performance in sports that require rotational power. This approach offers a cost-effective solution that benefits athletes in these specific sports.

Nevertheless, it is important to acknowledge that the study did not fully consider all potential confounding factors that could have influenced the results, potentially impacting the outcome. Further investigation is necessary to overcome these limitations and enhance the validity and dependability of the findings. It is recommended that forthcoming studies incorporate control groups with varying training interventions to strengthen the comprehension of medicine ball training and its utilisation in softball and other rotational power sports. Direct comparisons could be facilitated, and the specific effects of medicine ball training could be better isolated. In addition, thoroughly examining the lasting impacts of these training programmes and their applicability to competitive performance could provide valuable insights for athletes and coaches.

CONCLUSION

Based on the findings, incorporating medicine ball training into the preparation of softball athletes yields noteworthy enhancements in torso rotation strength and batting velocity. Coaches and trainers can effectively utilise this training method to improve female collegiate softball players' performance and potentially positively affect athletes in other rotational power sports. However, it is essential to consider the study's limitations. Further studies should expand on these findings to enhance our comprehension of movement-specific resistance training and its practical applications for athletes.

ACKNOWLEDGEMENTS

REFERENCES

Bragazzi, N. L., Rouissi, M., Hermassi, S., & Chamari, K. (2020). Resistance training and handball players' isokinetic, isometric and

maximal strength, muscle power and throwing ball velocity: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health, 17*(8), Article 2663. https://doi.org/10.3390/ijerph17082663

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101. https://doi.org/10.1111/1467-8721.ep10768783

Felion, C. W., & DeBeliso, M. (2020). The effects of core training on high school baseball performance. *Athens Journal of Sports, 7*(3), 173-188. https://doi.org/10.30958/ajspo.7-3-3

Gray, R. (2018). Comparing cueing and constraints interventions for increasing launch angle in baseball batting. *Sport, Exercise, and Performance Psychology, 7*(3), 318-332. https://doi.org/10.1037/spy0000131

Higuchi, T., Nagami, T., Mizuguchi, N., & Anderson, T. (2013). The acute and chronic effects of isometric contraction conditioning on baseball bat velocity. *The Journal of Strength & Conditioning Research, 27*(1), 216-222. https://doi.org/10.1519/JSC.0b013e318252ddba

Hussain, R. N. J. R., & Shari, M. (2021). Effects of resistance training and whole-body electromyostimulation on muscular strength in female collegiate softball players. *Pertanika Journal of Social Sciences & Humanities, 29*(3), 1939-1955. https://doi.org/10.47836/pjssh.29.3.24

Hussain, R. N. J. R., Kee, K. M., Razman, R. M., Ismail, S. I., Shari, M., Idris, N. M., Razak, R. A., & Kassim, N. A. M. (2019). Reliability of Zepp baseball on batting velocity. Malaysian. *Journal of Sport Science and Recreation, 15*(1), 21-29.

Hussain, R. N. J. R., Mea, K. K., Razman, R., Ismail, S. I., Shari, M., & Idris, N. M. (2019). The comparison between major muscle activations during different phases in softball batting. *Pertanika Journal of Social Sciences & Humanities, 27*(S3), 147-157.

Kobak, M. S., Rebold, M. J., Buser, S. L., Kappler, R. M., & Otterstetter, R. (2018). The effects of medicine ball training on bat swing velocity in prepubescent softball players. *International Journal of Exercise Science, 11*(4), 75-83.

Kong, H., & Xu, Z. (2023). Impacts of medicine ball training on explosive strength in tennis players' upper bodies. *Revista Brasileira De Medicina Do Esporte, 29.* https://doi.org/10.1590/1517-8692202329012023_0013

Larson, C. (2021). *The biomechanics of the softball swing in seven stages: Optimizing exit velocity* [Master's thesis, Lawrence University]. Lux the Lawrence University Institutional Repository. https://lux.lawrence.edu/luhp/155

O'Connell, M. E., Lindley, K., Scheffey, J., Caravan, A., Marsh, J. A., & Brady, A. (2022). Weighted baseball training affects arm speed without increasing elbow and shoulder joint kinetics. *Journal of Applied Biomechanics, 38*(5), 281-285. https://doi.org/10.1123/jab.2021-0339

Reinold, M. M., Macrina, L. C., Fleisig, G. S., Aune, K. T., & Andrews, J. R. (2018). Effect of a 6-week weighted baseball throwing program on pitch velocity, pitching arm biomechanics, passive range of motion, and injury rates. *Sports Health: A Multidisciplinary Approach, 10*(4), 327-333. https://doi.org/10.1177/1941738118779909

Reinold, M. M., Macrina, L. C., Fleisig, G. S., Drogosz, M., & Andrews, J. R. (2020). Acute effects of weighted baseball throwing programs on shoulder range of motion. *Sports Health: A Multidisciplinary Approach, 12*(5), 488-494. https://doi.org/10.1177/1941738120925728

Rivera, M., Leyva, W. D., Archer, D. C., Munger, C. N., Watkins, C. M., Wong, M. A., Dobbs, I. J., Galpin, A. J., Coburn, J. W., & Brown, L. E. (2018). No effect of assisted hip rotation on bat velocity. *International Journal of Exercise Science, 11*(4), 68-74.

Schoenfeld, B. J., Ratamess, N. A., Peterson, M. D., Contreras, B., Sonmez, G. T., & Alvar, B. A. (2014). Effects of different volume-equated resistance training loading strategies on muscular adaptations in well-trained men. *The Journal of Strength & Conditioning Research, 28*(10), 2909-2918. https://doi.org/10.1519/JSC.0000000000000480

Szymanski, D. J., DeRenne, C., & Spaniol, F. J. (2009). Contributing factors for increased bat swing velocity. *The Journal of Strength & Conditioning Research, 23*(4), 1338-1352. https://doi.org/10.1519/jsc.0b013e318194e09c

Szymanski, D. J., Szymanski, J. M., Bradford, T. J., Schade, R. L., & Pascoe, D. D. (2007). Effect of twelve weeks of medicine ball training on high school baseball players. *The Journal of Strength & Conditioning Research, 21*(3), 894-901. https://doi.org/10.1519/r-18415.1

Szymanski, D. J., Szymanski, J. M., Schade, R. L., Bradford, T. J., McIntyre, J. S., DeRenne, C., & Madsen, N. H. (2010). The relation between anthropometric and physiological variables and bat velocity of high-school baseball players before and after 12 weeks of training. *The Journal of Strength & Conditioning Research, 24*(11), 2933-2943. https://doi.org/10.1519/jsc.0b013e3181f0a76a

Terraza-Rebollo, M., & Baiget, E. (2021). Acute and delayed effects of strength training in ball velocity and accuracy in young competition tennis players. *Plos One, 16*(12), Article e0260825. https://doi.org/10.1371/journal.pone.0260825

Review Article

Unveiling the Landscape of Cruise Tourism Research: A Bibliometric Analysis and Visualisation

Md. Tariqul Islam[1], Biswabhusan Pradhan[2], Rozi Nor Haizan Nor[3]*, Rakib Ahmed Saleh[3] and Jeetesh Kumar[4]

[1]*School of Hospitality, Tourism & Events, Taylor's University, 47500 Subang Jaya, Selangor, Malaysia*
[2]*Department of Tourism and Travel Management, Central University of Jammu, Jammu 181143 (J&K), India*
[3]*Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[4]*Faculty of Social Sciences and Leisure Management, Sustainable Tourism Impact Lab, Centre for Research and Innovation in Tourism (CRiT), Taylor's University, 47500 Subang Jaya, Selangor, Malaysia*

ABSTRACT

Cruise tourism comprises having a vacation on a large passenger ship or boat, typically visiting several destinations along a predetermined itinerary, and it has taken the attention of academicians for the past two decades. This research study conducts a comprehensive longitudinal analysis of cruise tourism research covering 2003 to 2022. A meticulous selection process identifies 590 seminal articles from the Scopus database, which are then subjected to thorough bibliometric analysis. The R package (Bibliometrix R-package) is utilised to conduct the bibliometric analysis that incorporates various indicators to comprehend the overall trends in cruise tourism research, including annual output, top journal, prolific author, influential articles, author keywords, co-citations, and social structures. The analysis findings indicated that the United States and China emerged as the leading countries in cruise tourism research. Moreover, keywords like cruise tourism, cruise industry, cruise ships, cruise travel, and cruise experience appeared frequently, indicating their prominence within the domain. The insights gained from this analysis provide valuable strategic information for researchers, aiding them in developing and planning their future studies in cruise tourism.

Keywords: Bibliometric analysis, bibliometrix, co-citation analysis, cruise tourist, cruise tourism

INTRODUCTION

Cruise tourism has emerged as a popular form of travel in recent years, with millions of passengers taking to the seas every year. Cruise Lines International Association reported that this industry has experienced significant growth over the past few years, with the increase of cruise passengers from 17.8 million to 30 million in a decade (2009–2019; Micallef, 2020). This growth is expected to continue in the coming years, driven by increasing consumer demand, expanding into new markets, and introducing new technologies and innovations in the industry. Cruise tourism involves passengers embarking on a ship for a period, usually several days to several weeks, during which they visit various destinations (Manning, 2006). The ships offer various amenities and activities, including dining, entertainment, and recreational facilities. The destinations cruise ships visit vary widely, from tropical beach destinations to historical cities and cultural landmarks.

Moreover, the cruise tourism industry has a substantial effect on the global economy, with the industry generating astronomical sums in revenue each year. The direct economic impact of cruise tourism includes spending by cruise lines, passengers, and crew members, as well as employment opportunities in the industry (Li et al., 2022). Meanwhile, indirect economic impacts include spending on goods and services by suppliers to the industry and the ripple effect of cruise tourism spending on local economies. However, the cruise tourism industry has

also been subject to criticism and scrutiny in recent years. Environmental concerns, such as the impact of cruise ships on marine ecosystems and air quality, have been raised by environmental activists and governments (Brida & Zapata, 2010; Copeland, 2008; GhulamRabbany et al., 2013; Klein, 2010). In addition, there have been concerns about crew members' labour practices, working conditions, and safety and security issues on board cruise ships (Dahl, 2020; Oldenburg et al., 2010; Vukonić et al., 2016). Despite these challenges, the cruise tourism industry remains a significant and growing sector of the global tourism industry. As such, there is a need for continued research into various aspects of the industry, including its economic impact, environmental sustainability, and social and cultural implications.

The cruise industry's growth has led to increased interest from researchers, resulting in a large body of literature on various aspects of cruise tourism. However, despite the wealth of research, there is still a need for a systematic review of the literature to identify critical research gaps and emerging trends. Systematic literature review and bibliometric analysis are two different approaches to synthesising knowledge. Systematic reviews assist the researcher in analysing the conceptual structure of the articles analysed (Linnenluecke et al., 2020). On the other hand, bibliometric analysis is an approach that uses a set of quantitative methods to measure, track, and analyse scholarly literature. Moreover, it allows the researchers to take a more scientific approach

and synthesise comprehensive knowledge. Few systematic reviews were conducted on a specific theme of cruise tourism (Rungroueng & Monpanthong, 2023); however, to obtain a more comprehensive understanding of cruise tourism from a broader perspective, it is imperative to undertake a bibliometric analysis. This paper aims to conduct a bibliometric analysis of the literature on cruise tourism, using a range of bibliometric tools and techniques to identify patterns and trends in the research and also analyse the publication outputs, citation patterns, and collaboration networks of authors and institutions to gain insights into the most active areas of research and the most influential authors and institutions. Moreover, this bibliometric study also employed the sensemaking concept to interpret results by following the procedure of Lim and Kumar (2024). This analysis also focuses on identifying gaps in the literature and areas where further research is needed. Overall, this bibliometric analysis will provide a comprehensive overview of the research landscape on cruise tourism, highlighting key trends, knowledge gaps, and emerging research areas. Researchers, practitioners, and policymakers can use this information to identify further research opportunities and inform decision-making in the cruise tourism industry.

METHODOLOGY

Bibliometric analysis is a powerful tool that utilises relational and evaluative methodologies to assess the scientific output of countries, authors, and institutions (Benckendorff & Zehrer, 2013; R. Singh et al., 2021). This statistical approach provides valuable insights into a specific subject's epistemology and knowledge structure, with far-reaching implications for research and scholarship. By leveraging publication and citation data, bibliometric analysis creates an evolutionary structure of a research field, incorporating statistics and econometrics into the process. The quantitative analysis enables researchers to identify key trends and patterns in research output, including the most influential authors, institutions, and countries in a given field. This study used bibliometric analysis to analyse data using the R language, utilising the Bibliometrix R-package (Aria & Cuccurullo, 2017; R. Singh et al., 2021). This method enabled researchers to generate a comprehensive summary of the most impactful authors, institutions, and countries in a specific field of research.

Scopus, the largest abstract and citation database, was selected for this study to retrieve the relevant bibliographic information. Despite Web of Science being a prominent database for article retrieval, the author opted for Scopus due to its broader scope of collections. V. K. Singh et al. (2021) mentioned that approximately 99.11% of journals indexed in Web of Science are also indexed in Scopus. Moreover, Scopus provides approximately 20% more comprehensive coverage than Web of Science. Prior researchers have preferred to extract peer-reviewed literature for bibliometric and scientometric analyses in specific research fields (Bashir & Singh, 2023; Kokol et al., 2021). The search strategy

Md. Tariqul Islam, Biswabhusan Pradhan, Rozi Nor Haizan Nor, Rakib Ahmed Saleh and Jeetesh Kumar

included the keywords ("cruise" OR "cruises" OR "passenger ship" OR "cruise industry" OR "cruising" OR "cruise passenger" OR "cruise port" AND "tourism" OR "travel" OR "tourist"), following the standard search protocol. The initial search yielded 2873 articles related to cruise tourism research. The search results were further refined by setting the publication year between 2003 and 2022, considering only articles published in English; the refined process resulted in 946 articles. Finally, after screening the titles, abstracts, and keywords, a total of 590 articles were included in the analysis. Furthermore, the current study has used annual scientific production, most prolific journals, highly productive researchers, nations, institutions, most impactful articles, author's keywords, intellectual structure and social structure to comprehend the cruise tourism research publication trend. These indicators are the most commonly used to measure the performance of authors in specific research fields (Kumar et al., 2020; Mulet-Forteza et al., 2019).

RESULTS

Annual Scientific Production

Figure 1 illustrates the evolution and structure of publications related to cruise tourism research, with 590 articles published in various journals from 2003 to 2022. The graph shows that cruise tourism research publication frequency could have been more consistent between 2003 and 2013, with only 134 articles published during this period. However, the growth rate of publications started to increase in 2016. Notably, 2021 received the highest number of publications, with 79 documents published on this subject. The increasing number of publications suggests that cruise tourism is becoming an increasingly popular research topic. It may be due to the growing importance of the cruise industry in the global tourism market, which has led to a greater interest in understanding its impact on various aspects, such as the environment, local communities, and the economy.
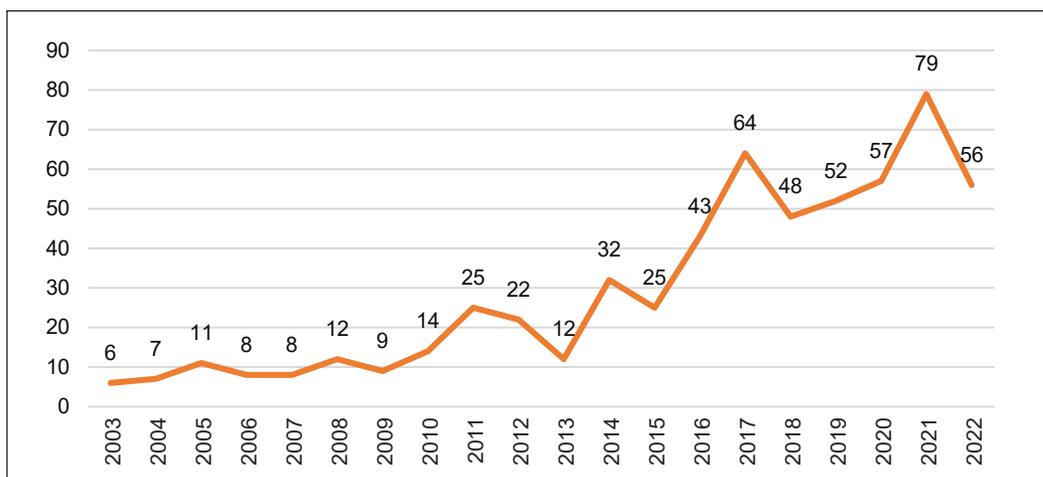
Figure 1. Annual scientific production of articles (*Source:* Authors' work)

Most Productive Journals

Figure 2 presents a comprehensive list of the top ten journals that have published cutting-edge research articles on cruise tourism across various contexts. Notably, the journal Tourism in Marine Environments has secured the top spot, with a remarkable 35 scientific publications to its credit. It demonstrates the pivotal role played by this journal in advancing research and scholarship in the cruise tourism domain. Other leading journals in this list include Sustainability (Switzerland) and Tourism Management; both journals have published 33 and 29 research articles, respectively, highlighting their significant contributions to the cruise tourism literature. In addition, the International Journal of Tourism Research, Tourism Economics, Current Issues in Tourism, and Annals of Tourism Research have also made noteworthy contributions to the field and published 18, 16, 15, and 14 research articles, respectively. It is worth noting that the Journal of Travel Research and Maritime Policy and Management has published an equal number of research articles, demonstrating their growing significance in the cruise tourism domain. The wealth of research published in these top-tier journals underscores the vital role played by scholarly research in advancing the understanding of cruise tourism and charting a sustainable future for the industry.

Most Impactful Authors, Institutions and Countries

Table 1 comprehensively summarises the most influential authors, institutions, and countries in the cruise tourism research domain. These metrics are regarded as major indicators in bibliometric analysis, which assist practitioners and researchers in identifying and fostering a collaborative and supportive environment for academic exchange in a particular domain. Upon analysing the final dataset, 1115 authors were identified, with Petrick JF securing the top position in the list with the highest number of citations, h-index, g-index, and publications. Han H and Hosany S hold the second and

Figure 2. Most productive journals (*Source:* Authors' work)

Table 1
Most impactful authors, institutions and countries

Author	h_index	g_index	m_index	TC	NP	PY_start
Petrick JF	13	20	0.619	1543	20	2003
Han H	11	13	1.1	709	13	2014
Hosany S	2	2	0.143	677	2	2010
Hung K	8	9	0.571	531	9	2010
Brida JG	14	17	1	507	17	2010
Mattila AS	3	3	0.143	446	3	2003
Witham M	1	1	0.071	437	1	2010
Duman T	2	2	0.095	423	2	2003
Hyun SS	6	6	0.667	349	6	2015
Hsu CHC	4	5	0.267	333	5	2009
Affiliation				**Articles**		
Hong Kong Polytechnic University				19		
Texas A&M University				16		
University of Genoa				14		
Sejong University				13		
Free University of Bolzano				12		
Griffith University				11		
Dalian Maritime University				9		
Lincoln University				9		
University of Brighton				9		
University of Palermo				9		
Country				**Frequency**		
USA				185		
China				110		
Spain				94		
Italy				85		
Australia				62		
South Korea				60		
Canada				54		
UK				53		
Croatia				31		
Germany				30		

Source: Authors' work

third positions on the list, with 709,677 citations, respectively. In terms of the most productive countries in cruise tourism research, the USA emerged as the leader with 3032 citations and 185 publications, followed by China with 596 citations and 110 publications. These findings highlight the significant contributions made by researchers from these countries in cruise tourism. Furthermore, the most impactful institutions in cruise tourism research were identified as The Hong Kong Polytechnic

University and Texas A&M University, which held the first and second positions, respectively, with 19 and 16 publications. These institutions have played a vital role in advancing research and scholarship in the cruise tourism domain, underscoring the importance of academic institutions in driving innovation and progress in the field.

Most Impactful Articles

Table 2 provides a detailed summary of the top 10 highly influential articles with the total number of citations each article receives. According to the Scopus data, out of 590 articles, "Dimensions of Cruisers' Experiences, Satisfaction, and Intention to Recommend" is the most influential, with 437 citations. Hosany and Witham (2010) study the links between cruisers' experiences, satisfaction, and intention to recommend, identifying the underlying aspects of cruisers' experiences and determining what influences those qualities. Similarly, the article "The Role of Affective Factors on Perceived Cruise Vacation Value" occupied second place with 411 citations. In this phenomenal work, Duman and Mattila (2005) investigated the influence of chosen

Table 2
Most impactful articles

Title	Author(s), Publication Year, Journal	Total Citations	TC per Year
1. Dimensions of cruisers' experiences, satisfaction, and intention to recommend	Hosany & Witham (2010), Journal of Travel Research	437	31.21
2. The role of affective factors on perceived cruise vacation value	Duman & Mattila (2005), Tourism Management	411	21.63
3. Are loyal visitors desired visitors?	Petrick (2004a), Tourism Management	249	12.45
4. The impact of customer-to-customer interaction on cruise experience and vacation satisfaction	Huang & Hsu (2010), Journal of Travel Research	245	17.50
5. Self-image congruence in consumer behaviour	Hosany & Martin (2012), Journal of Business Research	240	20.00
6. First timers' and repeaters' perceived value	Petrick (2004b), Journal of Travel Research	234	11.70
7. Why do you cruise? Exploring the motivations for taking cruise holidays and the construction of a cruising motivation scale	Hung & Petrick (2011), Tourism Management	206	15.85
8. Antarctic cruise tourism: the paradoxes of ambassadorship, "last chance tourism" and greenhouse gas emissions	Eijgelaar et al. (2010), Journal of Sustainable Tourism	195	13.93
9. The geography of cruises: Itineraries, not destinations	Rodrigue & Notteboom (2013), Applied Geography	172	15.64
10. Beyond the tourist bubble: Cruise ship Passengers in Port	Jaakson (2004), Annals of Tourism Research	162	8.10

Source: Authors' work

emotive elements (such as hedonics, control, and novelty) on the value within the context of cruise holiday experiences. In addition to this, it also highlighted how the level of client satisfaction impacts the link between effect and value. The findings of this study reveal that the emotional aspects have a significant role in determining the perceived value of cruise services. In particular, hedonics, often known as pleasure, has a significant connection to cruise vacationers' beliefs, values, and plans for how they would behave. The sensemaking emphasises the significance and impact of two highly influential articles (Duman & Mattila, 2005; Hosany & Witham, 2010) on cruise vacation experiences and satisfaction. Both studies offer valuable insights into the multifaceted nature of cruisers' experiences and the factors contributing to their satisfaction and perceived value.

Analysis of Author's Keywords

The keywords in the word cloud shown in Figure 3 between 2003 and 2022 represent the authors' keywords. The authors of this research classified their work using 1815 different keywords. Which occurred 2944 times; hence, only frequently occurring keywords are selected for further analysis. According to the word cloud that was created from the author's keywords, the keywords "cruise tourism," "cruise industry," "cruise ships," "cruise travel," and "cruise experience" occur the most often. The only thing that can be observed happening in the word cloud is a gradual accumulation of keywords. Because of this, it is insufficient on its own. As a result, trend themes analysis was used to investigate the level of popularity possessed by various ideas over periods to investigate the development of ideas presented in published works. The

Figure 3. Word cloud of author's keywords (*Source:* Bibliometrix R-package)

trend analysis revealed the logarithmic frequencies that were employed for a variety of terms across the articles. Figure 4 provides a detailed summary of the themes that are currently trending. It is clear from looking at Figure 4 that the terms "cruise tourism" (169), "cruise" (51), "cruise industry" (37), "cruise ships" (20), "cruise travel" (8), and "cruise experience" (8) have been used quite frequently. Moreover, sensemaking suggests that the research community has been consistently engaged with cruise tourism over the years, with certain themes and keywords maintaining their prominence throughout the period studied.

Intellectual Structure

Co-citation analysis is a methodology used in science mapping which operates under the assumption that publications which are frequently cited together have similar thematic content; in other words,

this technique examines the co-occurrence of citations among publications to identify relationships and patterns within a particular field of research, the basic idea is that if two publications are frequently cited together, they are likely to be thematically similar and share common research interests (Donthu et al., 2021; Hjørland, 2013; Kumar et al., 2020). Figure 5 reflects the authors' co-citation network consisting of two cluster solutions.

Cluster 1 (red) consists of 23 documents, including Hung and Petrick's phenomenal work. In this study, a measurement scale for motivation to cruise was developed. The study results showed that the measurement scale was reliable and valid. Additionally, the findings indicated that cruising motivation positively affects the intention to go cruising. This study also underscores the importance of understanding motivation's role in shaping potential cruise customers'
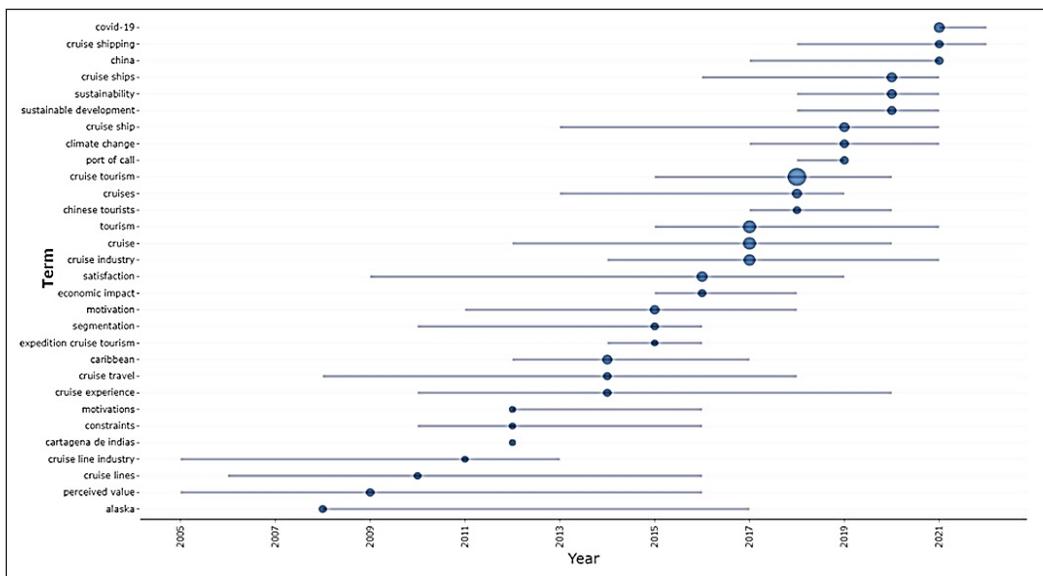
Figure 4. Trend topics (*Source:* Bibliometrix R-package)

intentions. Similarly, Jones (2011) explored three potential factors that may motivate tourists to take a cruise vacation, including their sources of information, the attributes of the vacation, and their motives for leisure activities as measured by the leisure motivation scale. The research analysis was organised according to the respondents' specific cruise itinerary and their experience of the cruise itself. The results showed that the influence of tested motivations varied with cruise experience but not cruise itinerary. Among cruise tourists, personal and internet-based information sources were found to be the most influential, and the need for stimulus avoidance primarily drove their desire to cruise. Furthermore, this cluster also contains various work on cruise tourism linked with cruise business performance and policy (Sun et al., 2014), cruise tourist experience (Andriotis & Agiomirgianakis, 2010; Hosany & Witham, 2010; Kwortnik, 2008), tourist purchase behaviour (Petrick, 2004b), development and growth of cruise industry (Cartwright & Baird, 1999), motivation factors and satisfaction level of cruise tourist (Qu & Ping, 1999) and price sensitivity (Petrick, 2005). Sense-making of this cluster suggests a comprehensive understanding of the motivations driving cruise tourism, including the development of measurement scales, the influence of various factors on cruising decisions, and implications for cruise industry stakeholders such as businesses, policymakers, and marketers. It highlights the multidimensional nature of cruise tourism research and its relevance
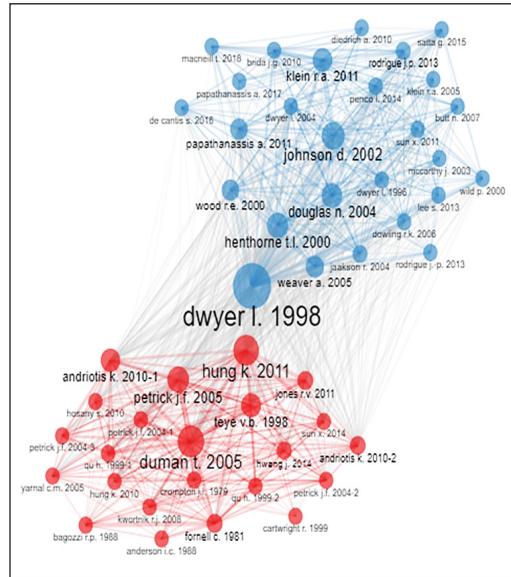
Figure 5. Co-citation analysis (*Source:* Bibliometrix R-package*)*

for addressing practical concerns within the industry.

Cluster 2 (blue) consists of 28 documents, including the major work of De Cantis et al. (2016), which conducted a study to segment cruise passengers based on their behaviour at the destination, using traditional survey methods and GPS (Global Positioning System) technologies to collect data on passengers' mobility. The study identified seven distinct activity patterns among passengers and found that socio-demographic characteristics and other factors were associated with movement patterns at the destination. The study highlights the importance of understanding cruise passengers' behaviour to better serve their needs and expectations. However, other major work includes a natural experiment conducted by MacNeill and Wozniak (2018), which measured the consequences of cruise

tourism on a local community. Despite industry estimates predicting economic gains, the study found little evidence of improvement in employment or income. The study also identified negative impacts on the community's ability to provide for necessities, increased corruption, and negative environmental impacts. The study suggests that large cruise tourism projects may fail to benefit local populations without community development and involvement initiatives. Moreover, the cluster is also linked with the work related to the economic significance of cruise tourism (Dwyer & Forsyth, 1996, 1998; Dwyer et al., 2004; Henthorne, 2000; Papathanassis & Beckmann, 2011; Sun et al., 2011; Weaver, 2005; Wood, 2000), cruise ship itinerary (Lee & Ramdeen, 2013; Rodrigue & Notteboom, 2013) and cruise tourist behavioural intention (Satta et al., 2015). Sensemaking of Cluster 2 indicates a comprehensive examination of cruise tourism, encompassing aspects of passenger behaviour, community impacts, economic implications, and itinerary planning. It underscores the importance of understanding these dynamics for cruise tourism destinations' sustainable development and management.

Social Structure

In order to determine the social structure of a topic, an analysis of the cooperation network is needed. The research of collaboration networks helps to discover structures like regular study groups, hidden groups of academics, and key authors (Aria &

Cuccurullo, 2017). The research groups are representative of the many social groupings that may be found in the region which has been examined. Figure 6 represents the author's collaboration network consisting of ten distinct clusters. The study groups have not been connected comprehensively concerning their social structures, which indicates that the cooperation between the various research groups is not particularly robust. It has been determined that the authors who contributed to the clusters worked together to publish more than one publication. For instance, Petrick JF and Papathanassis collaborate strongly due to the frequency of publishing within their network of colleagues. It can be observed that writers such as Petrick JF, Papathanassis A, Hung K and Wang S are actively participating in their social structure. Based on the sensemaking
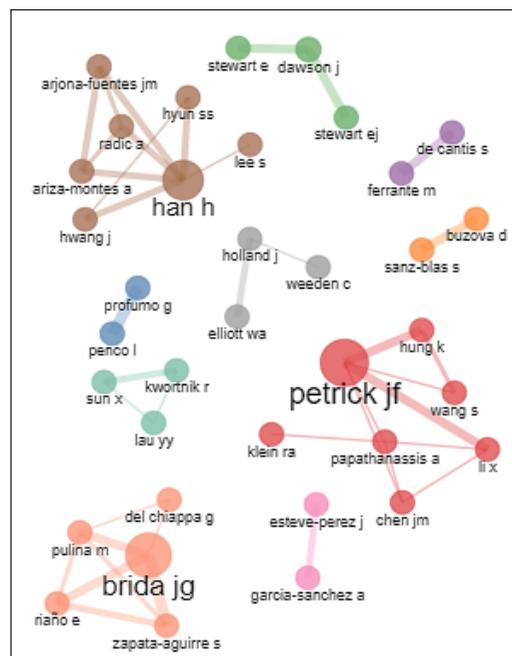
Figure 6. Collaboration network of authors (*Source:* Bibliometrix R-package)

suggestions, certain authors, such as Petrick JF, Papathanassis A, Hung K, and Wang S, are identified as actively participating in their respective social structures. It indicates their prominent roles within their research communities and potentially their influence on the collaborative dynamics within those groups.

The international cooperation networks between countries are shown in Figure 7, which consists of six clusters. The countries that makeup Cluster 1 (red) are as follows: Canada, Poland, Sweden, Oman, Egypt, Denmark, Switzerland, Hungary, France, Netherlands, the UK, Norway, Germany, Greece, Australia, New Zealand, Indonesia, Barbados and Iceland. Cluster 2 (green), cluster 3 (orange) and Cluster 4 (brown) are made up of two countries each. Cluster 5(blue) consists of six countries. In contrast, cluster 6 (purple) consists of thirteen countries. The USA is the most influential nation in this group and the second biggest

cluster in the collaboration network. The results of this investigation suggest that the USA and China are likely to be regarded as the most influential countries in cruise tourism research. The most productive nations, shown in Table 1, are ranked from most productive to least productive. The USA is ranked highest, followed by China. It can be assumed that most prestigious tourism and hospitality institutions are found in the United States, and China contributes to this predicament. In addition, Sensemaking of the international cooperation networks in cruise tourism research highlights the central role of influential countries like the USA and China, the dynamics of collaboration among nations, and the implications of productivity rankings for understanding the global landscape of cruise tourism research.

Anticipated Future Research Areas

The study offers a comprehensive overview of the scientific progress in cruise tourism
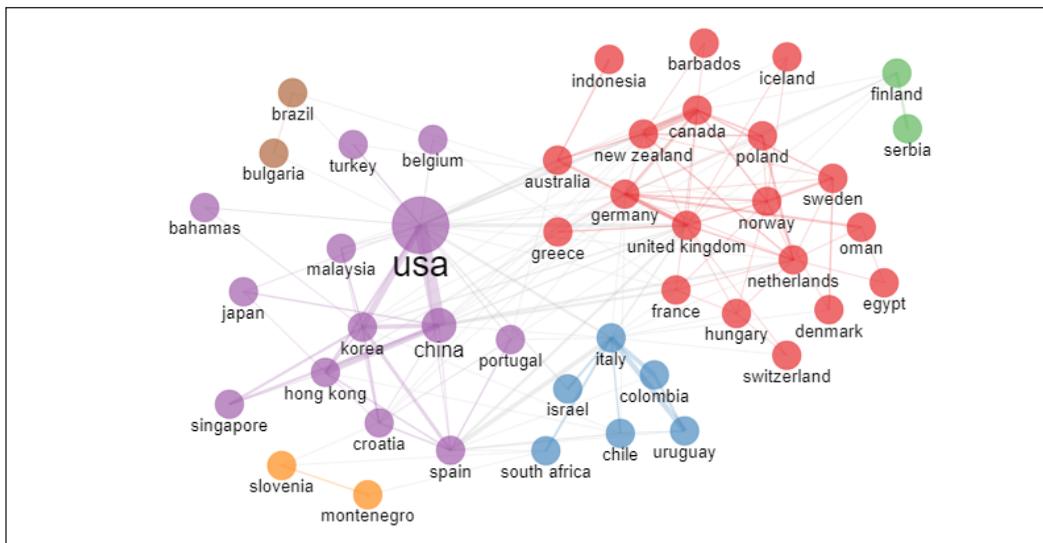
Figure 7. Collaboration network of countries (*Source:* Bibliometrix R-package)

research. Significant research gaps require attention in future endeavours. Addressing these gaps will enable future researchers and the scientific community to delve deeper into these areas, enriching the existing body of knowledge within this domain. The authors' extensive literature analysis shows opportunities for further exploration and investigation in these areas, fostering a better understanding of cruise tourism and its related factors. However, in a recent study, Risitano et al. (2022) investigated the segmentation of cruise passengers based on their onboard servicescape experience, yielding insightful managerial and scholarly implications. While limited by convenience sampling and a short observation period, future research should employ different sampling techniques, explore cross-cultural differences, and examine cruise companies' perspectives on experiential segmentation and its monetary and non-monetary returns. Longitudinal studies, technology integration, and a focus on sustainability and emotional aspects would further enrich the understanding of cruise passengers' experiences.

Chang et al. (2021) analysed how cruise tour attributes, such as leisure space and cabin comfort, influence tourists' preferences and satisfaction. To further improve customer retention, future research should also investigate the attractiveness of cruise routes and cruisers' willingness to revisit destinations. Understanding these additional factors will enhance the cruise tourism experience and increase customer return rates. Wu et al. (2021)

found that customer participation positively impacts service personal values, customer satisfaction, and loyalty in the context of cruise tourism. The study's limitations suggest the need for future research with longitudinal designs to strengthen the conclusions. Further exploration of Customer Perception and Service Product Values can offer valuable insights for cruise line operators to improve their strategies. Factors like environmental sustainability practices and personalised experiences may also enhance overall customer retention in the industry. However, research on this topic remains limited, providing only preliminary insights. The results revealed reputation and familiarity as the primary explanatory factors for passengers' intention with a port of call destination. Cognitive perception and affective evaluation were identified as antecedents of reputation and familiarity. Future research could expand on this by examining reputation and familiarity in different cultural contexts to understand variations in antecedent factors affecting destinations and passengers' future intentions. Additionally, comparative studies on passenger intention and behaviour could quantify the impact of reputation and familiarity on the sustainability of port-of-call destinations.

DISCUSSION

Examining existing literature is a crucial technique for shedding light on the current state of research and making recommendations for future studies that contribute to advancing the field. However,

this research suggests a comprehensive understanding of cruise tourism's conceptual and intellectual foundations is uncommon. The current study offers an evaluative and relational perspective of cruise tourism, focusing on its trends, foundations, and structural components to address this gap. The findings indicate a notable increase in research on cruise tourism since 2016. Scholars' interest in exploring various marketing strategies (Park et al., 2016) and consumer behavioural intentions (De Cantis et al., 2016) are the primary drivers of this growth in cruise tourism research. In terms of academic journals, Tourism in Marine Environments (35), Sustainability (Switzerland) (33), Tourism Management (29), International Journal of Tourism Research (18), Tourism Economics (16), Current Issues in Tourism (15), and Annals of Tourism Research (14) were the most prolific publishers of articles on cruise tourism. However, the number of scholars engaged in cruise tourism research still needs to be higher, with only Petrick JF, Brida JG, and Han H contributing a significant number of articles in this area. Therefore, more research is required to deepen the understanding of cruise tourism and the expectations of tourists regarding the various products and services the cruise industry offers.

The author's keywords analysis provides insight into the most commonly used terms in the published literature on cruise tourism. The word cloud generated from the author's keywords reveals that "cruise tourism," "cruise industry," "cruise ships," "cruise

travel," and "cruise experience" are the most frequently used terms. However, it is worth noting that the word cloud only shows the frequency of keywords without indicating their relationship or how they are used in the literature. Therefore, trend themes analysis was used to identify the development of ideas presented in published works over time. The trend analysis indicates that "cruise tourism" is the most popular theme in the literature, with 169 occurrences. It suggests that cruise tourism is a significant research area, and many scholars are interested in exploring it. Additionally, the term "cruise" has been used 51 times, indicating that cruise tourism is a popular subfield within the broader tourism industry. Furthermore, the terms "cruise industry," "cruise ships," and "cruise travel" were used 37, 20, and 8 times, respectively, indicating the growing interest in these topics in the cruise tourism literature. Finally, the term "cruise experience" was used eight times, highlighting the importance of studying cruise tourists' experiences and satisfaction levels.

The findings of the co-citation and collaboration network analyses yield significant insights into the intellectual and social structures of cruise tourism research. These findings shed light on the field's key authors, influential studies, thematic clusters, and international cooperation patterns. In terms of the intellectual structure, the co-citation analysis revealed two main clusters. Cluster 1 focused on the motivations and experiences of cruise tourists, including studies on motivation scales, information

sources, vacation attributes, and satisfaction levels. This cluster illustrated the importance of understanding motivation's role in shaping potential cruise customers' intentions. It also emphasised the importance of information sources and stimuli-avoidance in driving cruise tourists' desire to cruise. Furthermore, this cluster explored various aspects related to cruise business performance, policy, and the development of the cruise industry. On the other hand, Cluster 2 examined cruise passengers' behaviour at the destination, including their mobility patterns and the impacts of cruise tourism on local communities. This cluster highlighted the need to understand passengers' behaviour to better serve their needs and expectations, as well as the potential negative impacts of cruise tourism on employment, income, community well-being, and the environment. These findings indicate that cruise tourism research is diverse and covers various aspects, including motivations, experiences, business performance, policy, destination behaviour, and community impacts. The studies within each cluster share thematic similarities, as they are frequently cited together, indicating common research interests and relationships. It suggests that researchers within these clusters are building on and referencing each other's work, contributing to developing and advancing knowledge in their respective areas.

Moreover, moving on to the social structure, the collaboration network analysis revealed ten distinct clusters of researchers. These clusters represent different social groupings within the field

of cruise tourism research. While some clusters showed strong collaboration within their groups, their overall connectivity was not particularly robust. It suggests that there may be limited interaction and collaboration between different research groups in the field. However, it is worth noting that certain authors, such as Petrick JF, Papathanassis A, Hung K, and Wang S, stood out due to their active participation and frequent collaborations with other researchers. These authors play significant roles in their respective social structures and contribute to the overall knowledge production in cruise tourism research. However, the analysis of international cooperation networks highlighted the involvement of multiple countries in cruise tourism research. The USA emerged as the most influential nation, followed by China. The collaboration network revealed clusters of countries with shared research interests and collaborations. The presence of the USA and China as influential countries suggests they have prominent research institutions and contribute significantly to the cruise tourism research domain.

CONCLUSION

This extensive bibliometric analysis provides comprehensive insights into the evolving landscape of cruise tourism research. Notably, the field has witnessed substantial growth in recent years, reflecting the increasing significance of cruise tourism in the global travel industry. Key trends indicate a surge in research since 2016, primarily driven by inquiries into marketing

strategies and consumer behaviour. Prominent journals such as Tourism in Marine Environments, Sustainability, and Tourism Management have played pivotal roles in disseminating research findings. However, the community of scholars engaged in cruise tourism research still needs to be expanded. The co-citation and collaboration network analyses reveal two central thematic clusters: one focusing on the motivations and experiences of cruise tourists and the other on cruise passengers' behaviour at destinations. These findings underscore the multifaceted nature of cruise tourism research. Furthermore, international collaboration networks highlight the influential roles of the USA and China in this domain. This study not only unveils the current state of cruise tourism research but also identifies critical gaps and emerging areas, offering valuable insights for future scholars, practitioners, and policymakers to contribute to developing this dynamic field.

Implications of the Study

This research includes valuable theoretical and practical implications for tourism academia, cruise tourism stakeholders, and policymakers. It further enriches comprehension of cruise tourism's multifaceted aspects, laying a solid foundation for future research and facilitating strategic decision-making within the industry.

Theoretical Implications

Current research significantly contributes to the theoretical foundations of cruise tourism

research by meticulously highlighting critical research gaps for future research. Therefore, these highlighted areas enable researchers to formulate research directions adequately to tackle the current gaps in the field. By bridging these gaps, researchers can cultivate a more resilient and inclusive comprehension of cruise tourism research.

Further, a thorough analysis of annual scientific production and co-citation clusters offers a panoramic view of evolving trends and thematic contours for cruise tourism research. It facilitates a nuanced comprehension of cruise tourism research, enabling researchers to stay abreast of emerging research trajectories. With this knowledge, researchers can navigate the complex landscape of cruise tourism research with greater precision and agility.

Furthermore, co-citation and collaboration network analyses reveal the complex intellectual and social frameworks supporting cruise tourism research. Grasping these frameworks is crucial as it illustrates the interconnectivity among researchers and research communities, promoting knowledge sharing and creating conducive environments for potential collaborations. By strengthening linkages within the research community, researchers can magnify the impact of transformative advancements in cruise tourism research.

Additionally, identifying influential authors and journals in cruise tourism research explains the thought leaders and pivotal platforms for scholarly dissemination. With this knowledge, researchers can strategically navigate the vast expanse of literature

and forge meaningful collaborations with cruise tourism stakeholders. This strategic approach enhances their research's visibility and impact and fosters a vibrant knowledge dissemination and exchange ecosystem.

Moreover, the author's keywords and trend themes offer valuable insights into the dominant themes and conceptual underpinnings shaping cruise tourism research. By interpreting these trends, researchers can refine their research focus and align their work with the prevailing research paradigms in the field. This strategic alignment enhances the research's relevance and significance, fostering a cohesive and synergistic research ecosystem wherein researchers collectively contribute to advancing knowledge and driving innovation in cruise tourism research.

Practical Implications

This study presents the latest trends and current status of cruise tourism, offering valuable insights for practitioners to support effective decision-making and manage cruise tourism challenges. In addition, the findings of this research also offer valuable insights for public sector entities (policymakers) to make strategic decisions on environmental concerns, labour practices, safety regulations, and economic impacts.

Moreover, identifying influential authors and institutions can promote academic collaboration within cruise tourism. Researchers can establish partnerships with key figures and institutions to enhance the quality and impact of their work. Furthermore, educational institutions

offering programs related to tourism and hospitality can incorporate the insights from this study into their curriculum. It ensures that students are exposed to the latest trends and research findings in cruise tourism, preparing them for careers in the industry. Cruise companies can access cutting-edge research and expertise that informs their business practices by forming strategic partnerships with these critical entities. This collaboration can lead to adopting more sustainable approaches, such as minimising environmental impacts and enhancing community well-being, thereby meeting the evolving expectations of tourists who prioritise responsible travel experiences.

Furthermore, integrating the insights from this study into educational institutions' curricula ensures that future cruise tourism industry professionals are well-equipped with the knowledge and skills necessary to address sustainability challenges. It benefits students by providing them with practical and relevant training. It supports the industry by nurturing a workforce attuned to the importance of sustainability and equipped to drive positive change.

Furthermore, the study's emphasis on environmental impacts and community well-being provides actionable guidance for cruise companies aiming to bolster their sustainability efforts. By adopting the recommendations outlined in the study, such as minimising ecological footprints and cultivating positive ties with local communities, companies can bolster their reputation, appeal to conscientious travellers, and fortify their sustainability credentials.

Limitations and Future Research Directions

However, it is essential to acknowledge the constraints of this study. Firstly, the study only considered papers published between 2003 and 2022, which may limit the scope of the analysis and overlook earlier influential works. Future studies could consider expanding the time frame to gain a more comprehensive understanding of the historical development of cruise tourism research.

Several articles addressing diverse facets of cruise tourism, such as sustainability, resident perceptions, smart technology adoption, and tourist sentiment analysis, have emerged in the years 2023 and 2024 (Gonzáles-Santiago et al., 2024; Hoarau-Heemstra et al., 2023; Jordan et al., 2023; Löf et al., 2023; Sun et al., 2023). Notably, these recent publications were not incorporated into the present study. Consequently, it is recommended that future researchers undertake both bibliometric analyses and systematic literature reviews focusing on specific themes or dimensions within cruise tourism. These could include investigations into tourists' perceptions or attitudes towards cruise tourism, sustainable development within the cruise tourism sector, or adopting technological innovations in cruise tourism operations.

Secondly, the analysis focused solely on bibliographic data from articles, neglecting other types of publications such as books, conference papers, and reports. Including a broader range of publication sources could provide a more comprehensive overview of the research landscape in cruise tourism. Finally, the study relied on data collected from the Scopus database, which has limitations and may not cover all relevant publications in the field of cruise tourism. Future studies could consider using multiple databases (e.g., Web of Science) and bibliometric indicators to gather a more diverse and extensive dataset for analysis.

ACKNOWLEDGEMENT

REFERENCES

Andriotis, K., & Agiomirgianakis, G. (2010). Cruise visitors' experience in a Mediterranean port of call. *International Journal of Tourism Research, 12*(4), 390-404. https://doi.org/10.1002/jtr.770

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959-975. https://doi.org/10.1016/j.joi.2017.08.007

Bashir, A., & Singh, R. (2023). Mapping the research journey of tourism analysis: A bibliometric study. *Tourism Analysis, 28*(1), 155-159. https://doi.org/10.3727/108354222X16534530194787

Benckendorff, P., & Zehrer, A. (2013). A network analysis of tourism research. *Annals of Tourism Research, 43*, 121-149. https://doi.org/10.1016/j.annals.2013.04.005

Brida, J. G., & Zapata, S. (2010). Cruise tourism: Economic, socio-cultural and environmental impacts. *International Journal of Leisure and Tourism Marketing, 1*(3), 205-226. https://doi.org/10.1504/IJLTM.2010.029585

Cartwright, R., & Baird, C. (1999). *The development and growth of the cruise industry.* Elsevier.

Chang, M.-Y., Wang, C.-H., & Chen, H.-S. (2021). Exploring cruise tourists' preferences and satisfaction: The case of Taiwan. *Water, 13*(22), Article 3183. https://doi.org/10.3390/w13223183

Copeland, C. (2008). Cruise ship pollution: Background, laws and regulations, and key issues. In *CRS Report for Congress* (pp. 1-30). Congressional Research Service, Washington, DC.

Dahl, E. (2020). Coronavirus (Covid-19) outbreak on the cruise ship Diamond Princess. *International Maritime Health, 71*(1), 5-8. https://doi.org/10.5603/mh.2020.0003

De Cantis, S., Ferrante, M., Kahani, A., & Shoval, N. (2016). Cruise passengers' behavior at the destination: Investigation using GPS technology. *Tourism Management, 52*, 133-150. https://doi.org/10.1016/j.tourman.2015.06.018

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

Duman, T., & Mattila, A. S. (2005). The role of affective factors on perceived cruise vacation value. *Tourism Management, 26*(3), 311-323. https://doi.org/10.1016/j.tourman.2003.11.014

Dwyer, L., & Forsyth, P. (1998). Economic significance of cruise tourism. *Annals of Tourism Research, 25*(2), 393-415. https://doi.org/10.1016/S0160-7383(97)00098-4

Dwyer, L., & Forsyth, P. J. (1996). Economic impacts of cruise tourism in Australia. *Journal of Tourism Studies, 7*(2), 36-43.

Dwyer, L., Douglas, N., & Livaic, Z. (2004). Estimating the economic contribution of a cruise ship visit. *Tourism in Marine Environments, 1*(1), 5-16. https://doi.org/10.3727/154427304774865841

Eijgelaar, E., Thaper, C., & Peeters, P. (2010). Antarctic cruise tourism: The paradoxes of ambassadorship, "last chance tourism" and greenhouse gas emissions. *Journal of Sustainable Tourism, 18*(3), 337-354. https://doi.org/10.1080/09669581003653534

GhulamRabbany, M., Afrin, S., Rahman, A., Islam, F., & Hoque, F. (2013). Environmental effects of tourism. *American Journal of Environment, Energy and Power Research, 1*(7), 117-130.

Gonzáles-Santiago, M. S., Loureiro, S. M. C., Langaro, D., & Ali, F. (2024). Adoption of smart technologies in the cruise tourism services: A systematic review and future research agenda. *Journal of Hospitality and Tourism Technology, 25*(2), 285-308. https://doi.org/10.1108/JHTT-06-2022-0159

Henthorne, T. L. (2000). An analysis of expenditures by cruise ship passengers in Jamaica. *Journal of Travel Research, 38*(3), 246-250. https://doi.org/10.1177/004728750003800306

Hjørland, B. (2013). Facet analysis: The logical approach to knowledge organization. *Information Processing & Management, 49*(2), 545-557. https://doi.org/10.1016/j.ipm.2012.10.001

Hoarau-Heemstra, H., Wigger, K., Olsen, J., & James, L. (2023). Cruise tourism destinations: Practices, consequences and the road to sustainability. *Journal of Destination Marketing & Management, 30*, Article 100820. https://doi.org/10.1016/j.jdmm.2023.100820

Hosany, S., & Martin, D. (2012). Self-image congruence in consumer behavior. *Journal of Business Research, 65*(5), 685-691. https://doi.org/10.1016/j.jbusres.2011.03.015

Hosany, S., & Witham, M. (2010). Dimensions of cruisers' experiences, satisfaction, and intention to recommend. *Journal of Travel Research, 49*(3), 351-364. https://doi.org/10.1177/0047287509346859

Huang, J., & Hsu, C. H. (2010). The impact of customer-to-customer interaction on cruise experience and vacation satisfaction. *Journal of Travel Research*, *49*(1), 79-92. https://doi.org/10.1177/0047287509336466

Hung, K., & Petrick, J. F. (2011). Why do you cruise? Exploring the motivations for taking cruise holidays, and the construction of a cruising motivation scale. *Tourism Management, 32*(2), 386-393. https://doi.org/10.1016/j.tourman.2010.03.008

Jaakson, R. (2004). Beyond the tourist bubble?: Cruiseship passengers in port. *Annals of Tourism Research*, *31*(1), 44-60. https://doi.org/10.1016/j.annals.2003.08.003

Jones, R. V. (2011). Motivations to cruise: An itinerary and cruise experience study. *Journal of Hospitality and Tourism Management, 18*(1), 30-40. https://doi.org/10.1375/jhtm.18.1.30

Jordan, E. J., Vieira, J. C., Santos, C. M., & Huang, T. Y. (2023). Do residents differentiate between the impacts of tourism, cruise tourism, and Airbnb tourism? *Journal of Sustainable Tourism, 31*(2), 265-283. https://doi.org/10.1080/09669582.2020.1833894

Klein, R. A. (2010). The cruise sector and its environmental impact. In C. Schott (Ed.), *Tourism and the implications of climate change: Issues and actions* (Vol. 3, pp. 113-130). Emerald Group Publishing Limited. https://doi.org/10.1108/S2042-1443(2010)0000003009

Kokol, P., Blažun, H. V., & Završnik, J. (2021). Application of bibliometrics in medicine: A historical bibliometrics analysis. *Health Information & Libraries Journal, 38*(2), 125-138. https://doi.org/10.1111/hir.12295

Kumar, S., Sureka, R., & Vashishtha, A. (2020). The Journal of Heritage Tourism: A bibliometric overview since its inception. *Journal of Heritage Tourism, 15*(4), 365-380. https://doi.org/10.1080/1743873X.2020.1754423

Kwortnik, R. J. (2008). Shipscape influence on the leisure cruise experience. *International Journal of Culture, Tourism and Hospitality Research, 2*(4), 289-311. https://doi.org/10.1108/17506180810908961

Lee, S., & Ramdeen, C. (2013). Cruise ship itineraries and occupancy rates. *Tourism Management, 34*, 236-237. https://doi.org/10.1016/j.tourman.2012.03.009

Li, H., Zhu, L., & Zheng, M. (2022). Is the cruise enterprise giant also a giant of social responsibility for sustainable development? *Marine Policy, 138*, Article 105009. https://doi.org/10.1016/j.marpol.2022.105009

Lim, W. M., & Kumar, S. (2024). Guidelines for interpreting the results of bibliometric analysis: A sensemaking approach. *Global Business and Organizational Excellence, 43*(2), 17-26. https://doi.org/10.1002/joe.22229

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management, 45*(2), 175-194. https://doi.org/10.1177/0312896219877678

Löf, A., Sokolíčková, Z., & Steins, N. A. (2023). Experiencing Svalbard sustainably? Reflecting on what we can learn about polar cruise tourism from the SEES expedition. *The Polar Journal, 13*, 177-185. https://doi.org/10.1080/2154896X.2023.2205249

MacNeill, T., & Wozniak, D. (2018). The economic, social, and environmental impacts of cruise tourism. *Tourism Management, 66*, 387-404. https://doi.org/10.1016/j.tourman.2017.11.002

Manning, T. (2006). *Managing cruise ship impacts: Guidelines for current and potential destination communities*. Tourisk Inc. http://tourisk.com/wp-content/uploads/2017/10/Managing-Cruise-Ship-Impacts.pdf

Micallef, J. V. (2020, January 20). *State of the cruise industry: Smooth sailing into the 2020's*. Forbes. Retrieved April 24, 2023, from https://www.forbes.com/sites/joemicallef/2020/01/20/state-of-the-cruise-industry-smooth-sailing-into-the-2020s/

Mulet-Forteza, C., Genovart-Balaguer, J., Mauleon-Mendez, E., & Merigó, J. M. (2019). A bibliometric research in the tourism, leisure and hospitality fields. *Journal of Business Research, 101*, 819-827. https://doi.org/10.1016/j.jbusres.2018.12.002

Oldenburg, M., Baur, X., & Schlaich, C. (2010). Occupational risks and challenges of seafaring. *Journal of Occupational Health, 52*(5), 249-256. https://doi.org/10.1539/joh.K10004

Papathanassis, A., & Beckmann, I. (2011). Assessing the 'poverty of cruise theory' hypothesis. *Annals of Tourism Research, 38*(1), 153-174. https://doi.org/10.1016/j.annals.2010.07.015

Park, S. B., Ok, C. M., & Chae, B. K. (2016). Using Twitter data for cruise tourism marketing and research. *Journal of Travel & Tourism Marketing*, *33*(6), 885-898. https://doi.org/10.1080/10548408.2015.1071688

Petrick, J. F. (2004a). Are loyal visitors desired visitors? *Tourism Management*, *25*(4), 463-470. https://doi.org/10.1016/S0261-5177(03)00116-X

Petrick, J. F. (2004b). First timers' and repeaters' perceived value. *Journal of Travel Research, 43*(1), 29-38. https://doi.org/10.1177/0047287504265509

Petrick, J. F. (2005). Segmenting cruise passengers with price sensitivity. *Tourism Management, 26*(5), 753-762. https://doi.org/10.1016/j.tourman.2004.03.015

Qu, H., & Ping, E. W. Y. (1999). A service performance model of Hong Kong cruise travelers' motivation factors and satisfaction. *Tourism Management, 20*(2), 237-244. https://doi.org/10.1016/S0261-5177(98)00073-9

Risitano, M., Romano, R., & Sorrentino, A. (2022). Segmenting cruise passengers from the experiential marketing perspective: An explorative study. *Research in Transportation Business & Management, 45*, Article 100590. https://doi.org/10.1016/j.rtbm.2020.100590

Rodrigue, J. P., & Notteboom, T. (2013). The geography of cruises: Itineraries, not destinations. *Applied Geography*, *38*, 31-42. https://doi.org/10.1016/j.apgeog.2012.11.011

Rungroueng, T., & Monpanthong, P. (2023). Updating cruise tourism theme: A methodology of systematic literature review. *ABAC Journal, 43*(3), 241-267. https://doi.org/10.59865/abacj.2023.41

Satta, G., Parola, F., Penco, L., & Persico, L. (2015). Word of mouth and satisfaction in cruise port destinations. *Tourism Geographies, 17*(1), 54-75. https://doi.org/10.1080/14616688.2014.938689

Singh, R., Sibi, P. S., & Sharma, P. (2021). Journal of ecotourism: A bibliometric analysis. *Journal of Ecotourism*, 1-17. https://doi.org/10.1080/14724049.2021.1916509

Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics, 126*, 5113-5142. https://doi.org/10.1007/s11192-021-03948-5

Sun, X., Cao, X., & Lau, Y. Y. (2023). Exploring cruise tourists' sentiment expression pattern from online reviews: An insight into market positioning. *Tourism Management Perspectives, 49*, Article 101195. https://doi.org/10.1016/j.tmp.2023.101195

Sun, X., Feng, X., & Gauri, D. K. (2014). The cruise industry in China: Efforts, progress and challenges. *International Journal of*

Hospitality Management, 42, 71-84. https://doi.org/10.1016/j.ijhm.2014.05.009

Sun, X., Jiao, Y., & Tian, P. (2011). Marketing research and revenue optimization for the cruise industry: A concise review. *International Journal of Hospitality Management, 30*(3), 746-755. https://doi.org/10.1016/j.ijhm.2010.11.007

Vukonić, D., Bielić, T., & Russo, A. (2016). Organizational factors in management of "mega cruise ships" from crowd management control aspect. *Pomorstvo, 30*(1), 58-66. https://doi.org/10.31217/p.30.1.8

Weaver, A. (2005). Interactive service work and performative metaphors: The case of the cruise industry. *Tourist Studies, 5*(1), 5-27. https://doi.org/10.1177/1468797605062713

Wood, R. E. (2000). Caribbean cruise tourism: Globalization at sea. *Annals of Tourism Research, 27*(2), 345-370. https://doi.org/10.1016/S0160-7383(99)00073-0

Wu, H., Lv, X., Cavusoglu, M., & Cobanoglu, C. (2021). Chinese cruise customers loyalty: The impact of customer participation and service personal values. *Tourism Review, 76*(1), 255-276. https://doi.org/10.1108/TR-03-2019-0085

The Effect of English Proficiency on the Production of English Intonation by Chinese EFL Learners

Qingbo Tian[1,3], Ngee Thai Yap[1]*, Chwee Fang Ng[2] and Yasir Yahya[1]

[1]*Department of English, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[2]*Department of Foreign Languages, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[3]*School of Foreign Languages, Yuxi Normal University, 653100 Hongta District, Yuxi, Yunnan, China*

ABSTRACT

English intonation is an integral component of English pronunciation teaching. However, as students' proficiency levels in English pronunciation improve, it remains unclear whether their intonation levels also develop. The present study, based on the second language intonation learning theory, aims to investigate the influence of English proficiency on Chinese EFL learners' production of English pitch accents, edge tones, and intonation patterns from the perspective of phonological representation. Two language groups of participants took part in a reading task: native English speakers (12) and Chinese EFL learners (36). The learners were classified into three groups based on their scores in the Chivox National Spoken English Test, ranked from high to low: the advanced, intermediate, and elementary groups. The reading task comprised 90 dialogue pairs. The participants were required to read part B of each dialogue pair aloud, but afterwards, only the Chinese EFL learners attended the semi-structured interview. The results showed that the native English speakers only demonstrated significant differences from each learner group in four of the ten intonation types involving the three aspects of English intonation, which may indicate regional variations in American English and difficulties distinguishing (H*) and (L+H*). In addition, there were no significant differences between the three learner groups in producing the ten intonation types, which maybe attributed to their similar learning experiences.

Keywords: Autosegmental Metrical theory, Chinese EFL learners, English proficiency, L2 intonation learning theory, production of English intonation

INTRODUCTION

English intonation is a key element in pronunciation and refers to "the use of suprasegmental phonetic features to convey post-lexical or sentence-level pragmatic meanings in a linguistically structured way" (Ladd, 2008, p. 4). It conveys a speaker's linguistic information and influences their attitude (Chen, 2008). Moreover, it is an essential indicator of "speaker identity, reflecting the speaker's physical state, age, gender, psychological state, and sociolinguistic membership" (Mennen, 2007, p. 53). The recently published Teaching Guideline for English Majors has made requirements for English intonation in four core courses: oral English, comprehensive English, audio-visual-oral English, and English public speaking and debating, and has explicitly stated that students are supposed to master standard English intonation (Guidance Committee for English Major Teaching, 2020).

In the past two decades, a substantial body of empirical research has been done on the production of second language (henceforth "L2") intonation (e.g. Barto, 2015; Bi & Chen, 2013; Bu, 2016; Chen, 2008; H. Chien & Fon, 2019; S. Chien & Fon, 2020; Li, Qu, & Zhi, 2020; Lu & Miran, 2016; Meng & Wang, 2009; Nguyen & Dao, 2018; Xia & Mu, 2008). These studies cover various factors influencing learners' production of English intonation, such as English proficiency and first language (henceforth "L1") intonation. Among them, English proficiency is of great value to learners' production of English intonation.

It can reveal their developmental processes of acquiring English intonation and provide references for future English intonation teaching.

Previous studies have identified that English proficiency influences learners' production of English intonation; however, the extent of this influence varies (Albin, 2015; Chen, 2006; Graham & Post, 2018; Liu & Chen, 2016; McGory, 1997; Nguyen & Dao, 2018). For example, Albin (2015) presented a significant analysis and discussion. The researcher investigated three kinds of intonation transfer among Japanese learners of English: L1-transferred phrasal H- (a high tone located at the left edge of a phrase, which is similar to English %H), use of low fundamental frequency (henceforth "F0") targets at prosodic boundaries, and simple boundary rise from the final syllable. The results indicated the effect of English proficiency only for the last one. In another major study, Graham and Post (2018) examined the production of English intonation with a narrow focus on late bilinguals of different L1 backgrounds. The effect of English proficiency was limited to the choice of plateau-type H* accent produced by Japanese learners of English. A third study by Nguyen and Dao (2018) investigated intonation patterns of statements and questions produced by Vietnamese speakers at two different proficiency levels. The results showed that most intonation patterns produced by the advanced learners were the same as those of the native English speakers. However, the beginners used a variety of patterns, some

of which were different from those of the native English speakers.

In contrast, there is one study in which the impact of English proficiency has not been found on the production of English intonation. Meng and Wang (2009) explored the characteristics of boundary tones in spoken English by Chinese EFL learners. They found no significant differences in the production of boundary tone patterns between the learners with high and low proficiency in English.

Through further analysis, we discover that these studies employ diverse methods to evaluate the English proficiency levels of learners, resulting in varying degrees of reliability in their classification. For example, in Albin's (2015) study, the researcher did not provide information regarding rates, nor did the researcher test the inter-rater reliability. The only measure the researcher took was that prior to evaluating the learners' recordings, the raters were implicitly encouraged to listen to a model speaker and establish a baseline. In Graham and Post's (2018) study, three native English speakers were instructed to score 1–3 (1, advanced; 2, basic; 3, other) based on their independent judgments. The participants classified into the advanced or basic groups by at least two of the three raters were included in their study. In Nguyen and Dao's (2018) study, the distinction between advanced learners and beginners was primarily based on whether they had any experience studying abroad. In Meng and Wang's (2009) research, the raters consisted of graduate advisors and their students and foreign teachers from Britain. The composition of the raters was complicated, and the reliability of the raters was not mentioned.

In addition, concerning the research on learners' production of English intonation, there are primarily three approaches: the British configurational approach, the American level approach, and the Auto-segmental Metrical (henceforth "AM") approach (Chen, 2009).[1] Researchers have consistently used the first and third approaches. With the in-depth investigation of L2 intonation, a theory based on the AM approach has emerged for cross-language analysis of intonation, offering valuable insights for a more profound explanation of the production of L2 intonation.

The L2 Intonation learning theory (henceforth "LILt") was proposed by Mennen (2015), which holds that deviations observed in the production of L2 intonation by learners are primarily influenced by their L1 intonation. In this theory, four dimensions for cross-language studies of intonation are proposed:

1. The inventory and distribution of categorical phonological elements ('systemic dimension')
2. The phonetic implementation of these categorical elements ('realizational dimension')

---

[1] The AM approach, developed by Pierrehumbert (1980), is characterized by the key perspective that an intonation contour consists of three structural elements: pitch accents, phrase accents, and boundary tones. Moreover, the three structural elements can be described by two pitch events, H and L (pitch accents, e.g., H*, L*; phrase accents, H- and L-; boundary tones, H% and L%).

3. The functionality of the categorical elements or tunes ('semantic' dimension)

4. The frequency of use of the categorical elements ('frequency' dimension). (Mennen, 2015, p. 173)

Moreover, this theory holds that when learners have limited experience with the use of L2 intonation, they will rely on their L1 intonation to produce L2 intonation. However, as learners' proficiency levels in L2 improve, their production of L2 intonation will be enhanced, at least in some of the four dimensions.

However, few previous studies have systematically investigated the effect of English proficiency on the production of English intonation under the AM framework. For example, in Albin's (2015) study, the researcher focused on three types of intonational transfer. Graham and Post's (2018) study centered on the pitch accent (L+)H* (H* and L+H*). In McGory's (1997) study, the researcher investigated two intonation patterns, H*L-L% and L*H-H%.

From the discussion above, we can draw the following conclusions: firstly, the methods utilized by the researchers to assess learners' proficiency levels in English pronunciation require further improvement; secondly, there is a lack of systemic research on L2 intonation production. Therefore, in this study, we will employ a relatively objective approach to assess learners' proficiency levels in English pronunciation. Furthermore, based on the LILt, we will investigate the effect of English proficiency on Chinese EFL learners' production of English intonation from the perspective of phonological representation, involving the systemic and semantic dimensions.[2] Specifically, the research questions are as follows: (1) Are there any significant differences in the production of English pitch accents, edge tones, and intonation patterns between the native speaker group and each of the three L2 learner groups? (2) Are there any significant differences in the production of English pitch accents, edge tones, and intonation patterns between the three L2 learner groups?

METHODS

This study obtained approval from the ethics committee of UPM [JKEUPM-2022-249]. The ethical review included assessments of ethical issues, an informed consent form, and scientific soundness. Prior to the survey, the participants were informed about detailed information related to the study and required to sign a consent letter.

Participants

The study participants were native English speakers and Chinese EFL learners. We recruited the participants using social media platforms such as WhatsApp and WeChat through a student studying in the United States and the first author's colleague. A total of 13 native English speakers and 50 Chinese EFL learners accepted our invitation.

---

[2] This study is limited to the phonological level of English intonation rather than the phonetic level, as the former may change the meaning of an utterance (H. Chien & Fon, 2019; Mennen et al., 2010).

We excluded one native English speaker and fourteen Chinese EFL learners based on the following criteria. Firstly, we excluded one native English speaker who was not born in Ohio, USA. According to Kretzschmar (2004), the state of Ohio is often regarded as a model for the "General American." In addition, we excluded 4 Chinese EFL learners who were non-Han Chinese from the Yunnan Chinese dialect regions to ensure homogeneity among the participants. Yunnan is a multi-ethnic province. In addition to the Han ethnic group, there are 24 ethnic minorities. Lastly, 46 Chinese EFL learners took an English proficiency test. Based on the test results, we selected 36 from the learners, 12 for each English proficiency level.

The 12 native English speakers were all undergraduates from a university in Ohio, USA. They consisted of seven female and five male students. The ages of these participants ranged from 20 to 22. They were all born in Ohio, USA. Two of the participants received training in English pronunciation.

The 36 Chinese EFL learners were English majors from a provincial university in Yunnan Province in southwest China. There were 34 female and 2 male students. They were Han Chinese from the Yunnan-Chinese dialect regions. Their ages ranged from 18 to 23. They had no experience of living abroad. They reported that they had normal hearing and no issues with language expression.

Chivox National Spoken English Test

The Chivox National Spoken English Test (henceforth "CNSET") is China's first oral English proficiency test. It uses artificial intelligence speech technology for scoring and relies on big data corpus technology to comprehensively assess candidates' pronunciation, pragmatic, and expressive abilities. The marks of this test can be used to compare a candidate's oral ability at different stages and to compare candidates' abilities in other areas at the same stage (China Information Association, 2021).

This test is hierarchical. Levels 1-2, 3, and 4 correspond to elementary school students, junior high school students, and senior high school students, and level 5 corresponds to college students. Level 5 has six aspects: paper reading, listening and repeating, listening and retelling, listening and answering, listening and summing up, and thinking and speaking. We used the level 5 test materials to assess the learners' English proficiency levels.

The test process and the selection of participants in this study were carefully arranged. Before the test, the participants were given 10 minutes to familiarize themselves with the test's question types. After getting these participants' scores, the researcher sorted their results and selected 12 participants with the highest, 12 with the middle, and 12 with the lowest scores.[3] The ten participants who fell outside the three groups were excluded from the study.

A one-way ANOVA was performed for their scores to examine the extent of the differences between the three groups. The test was significant at 0.5, $F(2,$

---

[3] We select 12 students for each group to compare with native English speakers. This method may effectively examine the differences between the three learner groups.

33)=67.316, p=0.000, η^2=0.803. Also, post-hoc analysis with Bonferroni correction revealed significant mean differences between the three groups (all p<0.001). To be specific, the mean score of the advanced group (M=68.16, SD=4.23) was significantly higher than that of the intermediate group (M=58.13, SD=2.06), which was significantly higher than that of the elementary group (M=45.98, SD=6.62).

Semi-structured Interview Questions

A semi-structured interview can provide an in-depth understanding of interviewees' views, experiences, and attitudes. Compared to a questionnaire, it can better explore interviewees' true feelings and thoughts because they have more opportunities to explain and clarify what they said (Safrudiannur, 2020).

We used this semi-structured interview to know the learners' learning of English pronunciation and intonation and interpret their actual performance. We designed the interview questions by referring to previous studies (Benrabah, 1990; McGory, 1997; Zárate-Sández, 2015) and taking into account the actual situation of learners. The questions consisted of three parts: 1) questions 1-3 examined the learners' experiences of English learning, 2) question 4 assessed the learners' knowledge of English pronunciation, and 3) questions 5-9 evaluated the learners' knowledge and the use of English intonation.

Materials

The materials used in this study consisted of 90 dialogue pairs. There were two primary sources of these dialogue pairs. One was from previous literature (Benrabah, 1990; Grice, 1995; Li, Wan, et al., 2020; Lian et al., 2013; Ward & Hirschberg, 1985; Wells, 2006; Works, 1985); the other was from dialogue pairs designed by the researcher. These materials were designed to examine the participants' production of English intonation. After completing these materials, we invited a native English speaker to review their grammatical issues, and he made minor revisions to five dialogue pairs.

The materials of this study were divided into five parts, which involved the three structural elements of English intonation and their combinations presented below (Bartels, 1999; Hirschberg & Beckman, 1994; Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990; Truckenbrodt, 2012).

> Pitch accents:
> H* (emphasis), L+H* (correction/contrast), L* (given information), L*+H (uncertainty)
>
> Edge tones:
> L-L% (not questioning a proposition), H-H% (putting up a proposition for question), L-H% (continuation dependence)
> Intonation patterns:
>
> H*L-L% (statements), L*H-H% (yes-no questions), H*L-H% (statements and non-finality)[4]

---

[4] In this study, the three structural elements and their combinations are summarized into three aspects: pitch accents, edge tones, and intonation patterns. Furthermore, each of the three aspects includes different types, and we selected the more representative ones.

We provide some examples below to better illustrate the use of these items in the materials.[5]

(1) A: Who arrived with the salad?
B: <u>Robbie</u> arrived with the salad.
H*, L-L%,   H*L-L%

(2) A: Bob made the salad.
B: No. <u>Bill</u> made the salad.
L+H*

(3) A: Mike loves his family.
B: Does <u>Molly</u> love her family?
L*, H-H%, L* H-H%

(4) A: Would anybody in their right mind marry Manny?
B: <u>Anna</u> may marry Manny. (Uncertainty about whether Anna is one of the people in their right mind.)
L*+H

(5) A: Who voted for Oliver?
B: <u>Gail voted for him</u>, but she did not like him.
L-H%, H*L-H%

Data Collection

To accurately reflect the learners' production of English intonation, we extracted a part of a native speaker's recording to elicit the targeted sentences. Before the survey, we invited two native English speakers (male and female) to read all the dialogue pairs aloud. We chose the male speaker's recording based on its audio quality and its suitability for our purposes. Part A of each dialogue pair was extracted from his recording, and then both the extracted sounds and the accompanying texts were integrated into presentation slides.

We provided this guidelines to the participants in the reading task to achieve the research objectives: (1) to familiarize the specific requirements of the reading task using the five dialogue pairs in the slides, (2) to listen to the recordings of Part A and then read B aloud at their normal speed, (3) to pay attention to the hint(s) in the bracket (if any), (4) to read Part B with the intonation patterns that they believe were appropriate, and (5) to read Part B at least twice, with the final reading being fluent.

Setting the location for the reading task and configuring the recording tool parameters were essential aspects of this investigation. This task was conducted in quiet classrooms at the universities where the participants were enrolled. The location was equipped with several laptops. The study's recording software was Audacity 1.3 Beta (Unicode), with a sampling frequency configured at 44100Hz. The sound file was saved as a mono soundtrack in WAV format.

After the recordings were completed, the learners were required to attend the semi-structured interview. First, we introduced the materials and the objectives of the interview. The interview comprised nine questions about their experiences learning English pronunciation and intonation. Second, we conducted individual interviews. Individual interviews contribute to a comprehensive understanding of participants' perspectives

---

[5] Since the use of pitch accents in English involves different contexts, we limit the scope of the study to English intonation with a narrow focus.

and protect their privacy. Finally, after obtaining their permission, we recorded the interview.

Data Analysis

We followed the Mainstream American English-Tones and Break Indices (henceforth "MAE-ToBI") convention in annotating English intonation (Beckman & Elam, 1997; Beckman et al., 2006; Hirschberg & Beckman, 1994; Silverman et al., 1992). This system has four tiers: an orthographic tier, a tone tier, a break-index tier, and a miscellaneous tier. Using Praat 6.2.14, we can easily identify and annotate the items mentioned above. Figure 1 provides some examples of using this system to annotate English intonation.

Figure 1(A-D) illustrates the four mentioned pitch accent types: H*, L+H*, L*, and L*+H. The pitch target H refers to the upper part of a speaker's pitch range, while the L refers to the lower part. The asterisk (*) following these pitch events indicates the position of the accented syllables. Unlike the other three pitch accents, the H in L*+H is situated on the syllable following the accented one rather than within the accented syllable. Figure 1(E) shows that the H* is preceded by the symbol "!", indicating that the H* in this syllable is realized at a lower pitch compared to the preceding accented syllable.

Figure 1 also illustrates three edge tone types: L-L% (A and B), H-H% (C), and L-H% (D and E), in which the H- and L- are phrase accents, typically found near the accented syllables and the H% and L% are boundary tones, typically located at the end of intonational phrases.

Intonation patterns are composed of pitch accents and edge tones. In Figure 1(A, C, and E), we can observe three distinct intonation patterns, namely, H*L-L%, L*H-H%, and H*L-H%, respectively.

Participants committed a variety of errors in producing English intonation. We classified these errors and illustrated their corresponding pitch contours, as depicted in Figure 2.

Figure 2 provides examples of error types in the production of pitch accents by speakers when conveying given information and emphasis. In Figure 2(A-C), the speakers are expected to use L* on the targeted word "volleyball," but they use H*, L+H*, and L*+H, respectively. In Figure 2(D), the speaker is supposed to use H* on "name" but use L* instead. In Figure 2(E), the expected accent on "name" is absent, resulting in the pitch accent error type Na.

In addition, in Figure 2(A-C and F), the speakers are supposed to use H-H%, but they use P, L-H%, L-L%, and H-L%, respectively. In Figure 2(A), the speaker accents the words "popular" and "at" after the targeted word "volleyball," leading to the edge tone error type P. In Figure 2(D), the speaker is anticipated to use L-H%, but he uses H-H%. In Figure 2(E), the speaker does not accent the targeted word "name," resulting in the edge tone error type NP.

The transcription of English intonation consisted of two steps. One was that the researcher was in charge of the words tier, the breaks tier, and the misc. tier. In
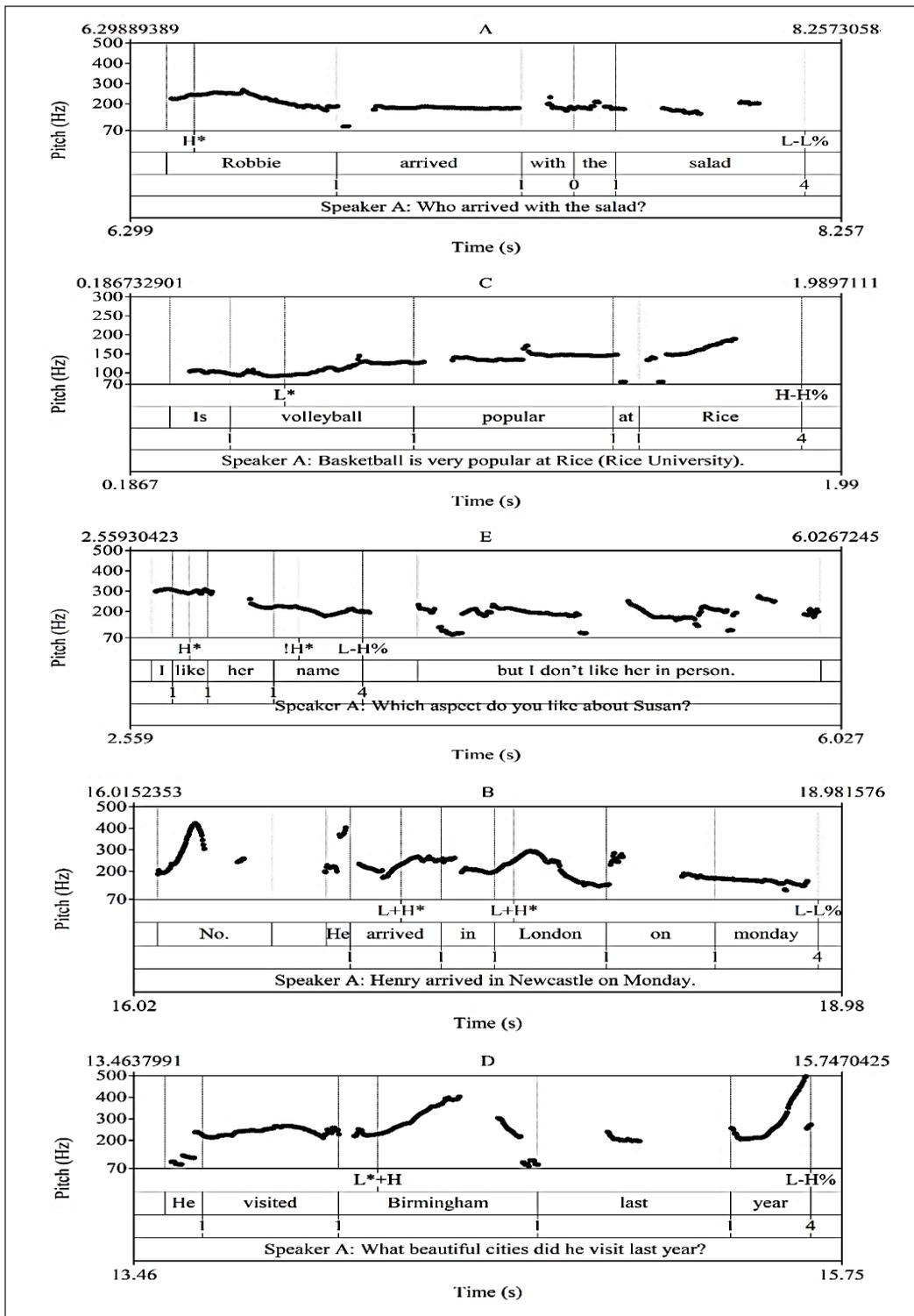
Figure 1. AM analysis of English intonation (*Source:* Authors' work)
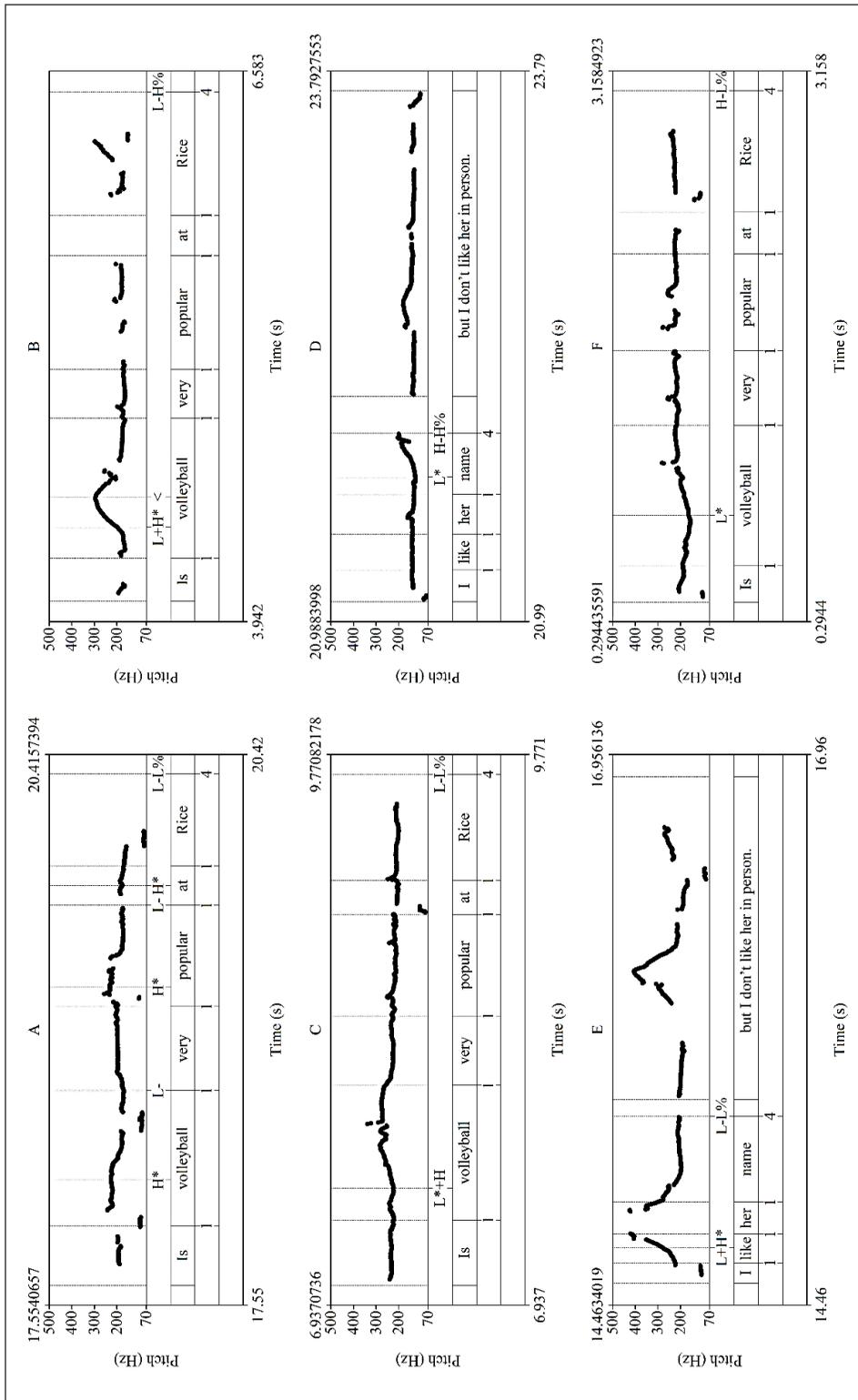
Figure 2. Error types in the production of English intonation by participants (*Source:* Authors' work)

Note. Na is where targeted words are not accented; P is where other accented words follow the targeted words; NP is the case of edge tone caused by Na, with or without accented words following it

addition to these tiers, he was responsible for determining the location of the accented words. In this process, a native English speaker was invited to identify the words the researcher was unsure whether they were accented. The other was that the researcher and his colleague annotated the tones independently. After the end of the annotation, the two raters discussed different tones between them by reviewing previous materials and seeking suggestions from four experts in the field. These experts have a wealth of experience in annotating English intonation. They reviewed the annotated materials we provided and offered detailed explanations.

To ensure the accuracy of the annotations, we tested the inter-reliability between the raters. We first converted the annotated tones into numbers 1 and 0, representing agreement and disagreement with the American model.[6] Subsequently, based on the ratings of the two raters, one-fourth of the recordings (12/48) were selected for a reliability test.[7] The test selected was Fleiss' Kappa. The results showed that the values of pitch accents, edge tones, and intonation patterns were 0.868, 0.8765, and 0.756, respectively. All the values were above 0.75, which

suggested a high inter-rater reliability.

For the analysis, we conducted both descriptive and inferential statistical analyses. English proficiency was used as the independent variable, whereas the accuracy scores of pitch accents, edge tones, and intonation patterns were used as the dependent variables. The results of the descriptive statistics for the variable, English proficiency, were offered. A Shapiro-Wilk test of normality was conducted to establish which statistical test should be used. A choice was made between a one-way ANOVA and a Kruskal-Wallis test, depending on whether the data presented a normal distribution and whether the homogeneity of variances test was met (if normally distributed; Cohen, 1988; Green & Salkind, 2014).

RESULTS AND DISCUSSION

Table 1 displays the means and standard deviations of ten types of intonation scores for the four participant groups. The results showed that the native English speaker group achieved higher scores than each of the three learner groups in the ten intonation types except for H*. Moreover, the differences between the learner groups were small for the ten intonation types.

Table 1 also provides the results from the normality tests. The tests revealed the presence of normal distribution for L+H*, L-L%, H-H%, H*L-L%, and L*H-H%. Nevertheless, the homogeneity of variances tests showed significant values of more than 0.05 for L-L% (p=0.158) and H*L-L% (p=0.301) for Levene statistic, but not for

---

[6] Figure 1 illustrates cases where the participants produced tones consistent with the American model, indicated by a value of 1. However, considering that H-L% in American English is considered a variant of H-H% (Bartels, 1999), we assign a value of 1 to H-L%. On the contrary, Figure 2 displays cases of different tones from the American model, marked with a value of 0.

[7] In Hua's (2019) study, after two raters had rated the recordings, one-fifth of the recordings were used to test the reliability between them.

Table 1

Descriptive statistics and normality test results of ten types of intonation scores for the four participant groups

Types	Group 1			Group 2			Group 3			Group 4		
	M	*SD*	*p*	*M*	*SD*	*p*	*M*	*SD*	*p*	*M*	*SD*	*p*
H*	10.1	3.6	**0.045**	12.2	2.2	0.113	11.7	2.8	0.406	10.9	2.2	0.583
L+H*	6.4	4.5	0.606	4.3	2.3	0.055	4.9	2.2	0.213	4.6	3.2	0.668
L*	14.4	3.7	0.075	6.2	2.4	0.834	4.7	4.5	**0.039**	4.3	2.1	0.486
L*+H	2.8	3.5	**0.014**	1.7	2.0	**0.016**	0.3	0.7	0.000	0.9	1	**0.028**
L-L%	13.9	1.8	0.635	9	1.7	0.350	7.9	2.8	0.135	8.7	1.7	0.149
H-H%	13.3	3.9	0.060	3.3	1.9	0.779	2.1	1.6	0.363	2.4	1.4	0.811
L-H%	2.1	2.5	**0.008**	1.3	1.5	**0.015**	0.2	0.4	0.000	0.8	1	**0.005**
H*L-L%	9	3.2	0.129	7	2.3	0.100	6.3	2.9	0.928	6.1	2.1	0.294
L*H-H%	12.7	4.6	0.193	3.3	1.9	0.779	1.7	1.4	0.082	2.3	1.2	0.495
H*L-H%	0.8	1.2	**0.001**	0.4	0.7	**0.000**	0.1	0.3	**0.000**	0.2	0.4	**0.000**

Note. Group 1, the native speaker group; Group 2, the advanced group; Group 3, the intermediate group; Group 4, the elementary group; M, Mean; SD, Standard deviation; p, Shapiro-Wilk sig.
Source: Authors' work

L* (*p*=0.013), H-H% (*p*=0.029), L*H-H% (*p*=0.000). Therefore, two different tests were used to evaluate the relationship between the four participant groups and the ten types of intonation scores: L-L% and H*L-L%, a one-way ANOVA test, and H*, L+H*, L*, L*+H, H-H%, L-H%, L*H-H% and H*L-H%, a Kruskal-Wallis test.

The tests showed that no significant differences were found among the four groups for H*, χ2 (3, N=48) = 2.855, *p*=0.415, L+H*, *F*(3, 44) = 1.090, *p*=0.363, and L*+H, χ2 (3, N=48) = 6.777, *p*=0.079, and H*L-H%, χ2 (3, N=48) = 6.668, *p*=0.083. However, there were statistically significant differences for L*, χ2 (3, N=48) = 25.399, *p*=0.000, η2=0.54, L-L%, *F*(3, 44)=21.364, *p*=0.000, η²=0.593, H-H%, χ2 (3, N=48) = 28.210, *p*=0.000, η2=0.6, L-H%, χ2 (3, N=48) = 9.801, *p*=0.020, η2=0.21, H*L-L%, *F*(3, 44)=2.886,

p=0.046, η²=0.164, L*H-H%, χ2 (3, N=48) = 28.212, *p*=0.000, η2=0.6.

Post-hoc analyses using pairwise comparisons were conducted with a Bonferroni correction. The results of the tests showed significant differences between the native speaker group and each of the three L2 learner groups for L*, L-L%, H-H%, and L*H-H%, and between the native speaker group and the intermediate group for L-H%. However, no significant differences were found between the groups for H*L-L%. For specific details, see Table 2.

Based on the results of these tests, there were significant differences in four of the ten intonation types between the native speaker group and each of the L2 learner groups. There were also no significant differences between the three learner groups in the ten intonation types.

Table 2

Pairwise comparisons of L*, L-L%, H-H%, L-H%, H*L-L% and L*H-H% scores for the four participant groups

Pairs	L*	L-L%	H-H%	L-H%	H*L-L%	L*H-H%
	p(adj.sig.)	*p*	*p*(adj.sig.)	*p*(adj.sig.)	*p*	*p*(adj.sig.)
Group 1 vs. Group 2	0.015	0.000	0.004	1.000	0.455	0.013
Group 1 vs. Group 3	0.000	0.000	0.000	0.018	0.117	0.000
Group 1 vs. Group 4	0.000	0.000	0.000	0.829	0.067	0.000
Group 2 vs. Group 3	1.000	1.000	1.000	0.128	1.000	0.416
Group 2 vs. Group 4	1.000	1.000	1.000	1.000	1.000	1.000
Group 3 vs. Group 4	1.000	1.000	1.000	0.829	1.000	1.000

Note. Group 1 is the native speaker group; Group 2 is the advanced group; Group 3 is the intermediate group; and Group 4 is the elementary group. p is the mean difference, which is significant at the 0.05 level; *p*(adj. sig.) is the significance level, which is 0.05, and significance values have been adjusted by the Bonferroni correction for multiple tests.
Source: Authors' work

Differences Between the Native English Speaker Group and Each of the Three Learner Groups

To explain the results comprehensively, we examined three main aspects: error types made by the native English speakers, results from the semi-structural interviews conducted with the learners, and previous studies.

Table 3 summarizes the error types of the four-pitch accents produced by native English speakers. Regarding H*, there were two main error types. One was that the native English speakers did not distinguish between H* and L+H* when expressing emphasis. The proportion of cases where H* was replaced with L+H* was 27%. The other was that the targeted words were not accented, which accounted for 13% of the total tokens. With respect to L+H*, the native English speakers did not distinguish between L+H* and H* when expressing contrast. A higher frequency of H* was used in comparison to L+H*, accounting for 55% and 36%, respectively. With regard to L*, though the native English speakers may use other pitch accents, they were not as high in proportion. Regarding L*+H, associating

Table 3

Types of errors in the production of pitch accents in different contexts by native English speakers

Types	Percentages of pitch accents produced and error types (%)				
	H*	L+H*	L*	L*+H	Na
H*	56	27	1	3	13
L+H*	55	36	1	2	7
L*	7	6	80	2	4
L*+H	38	29	10	16	8

Note. The numbers highlighted in grey refer to the percentages of correct responses from the participants; Na refers to the cases where targeted words are not accepted.
Source: Authors' work

this pitch accent with the meaning of uncertainty appeared to pose a challenge for the native English speakers. In this case, alternative pitch accents, H* (38%), L+H* (29%), and L* (10%) were used to substitute for this pitch accent.

The above results show that the native English speakers did not distinguish between H* and L+H* when expressing emphasis or contrast. It was consistent with Bartels's (1999) and Ladd and Schepman's (2003) claims. Bartels merged L+H* and H* into a single H*.

> One further simplification I will introduce here without extensive justification is transcribing all high-pitch accents without trailing tones as H*, whether they might be argued to show a leading L tone. In other words, there will be no distinction between H* and L+H*. The difference between these two patterns may not be a categorical one in phonology; the relatively greater perceptual prominence of L+H* may simply be a function of the relatively higher H* tone (Bartels and Kingston 1995). As to interpretation, any context in which L+H* might be used as a nuclear accent might also display the less emphatic H*, though the former is naturally more likely to promote inferences of 'contrast' arising from ostensibly greater speaker involvement (Bartels, 1999, pp. 19-20).

Ladd and Schepman (2003, p. 104) regarded H* and L+H* as "instances of a single accent category," which were described as (L+H)*.

> The F0 level of the F0min and the second H* accent is affected by the number of syllables intervening between the two accented syllables in a way that is not predicted by Pierrehumbert's "sagging transition" model, which is central to the distinction between H* and L+H*. We, therefore, argue that in both H* and L+H*, there are distinct L and H targets and that the two should be regarded as belonging to a single accent category (Ladd & Schepman, 2003, p. 81).

With regard to the issue above, we consulted with Nanette Veilleux (Professor, Computer Science and Informatics Program, Simmons University, personal communication, 2022/1/1) about the necessity of distinguishing between these two pitch accents. She replied that "Native English (either British or MAE) speakers can hear a distinction. Labelers usually describe hearing the Low tone throughout the beginning of the accented syllable, rising to a high tone later. Other labelers describe hearing the "scoop" or "rise" over the accented syllable."

To sum up, although the results showed that the native English speakers did not differentiate between these two pitch accents, it was still necessary to differentiate them due to their distinct meanings in American English.

The second issue related to pitch accents was L*+H. The results of this study did not receive support from other research (Ward & Hirschberg, 1985, 1986; Veilleux et al., 2012)**.** One possible reason for this difference was that the researcher did not intentionally prompt the native English speakers to produce this pitch accent. As mentioned earlier, the researcher simply instructed the native English speakers to read Part B of each dialogue pair based on Part A and the information in parentheses. Moreover, during the recording, the researcher did not interfere with the participants to gain insight into their real use of English intonation.

To gain an in-depth understanding of the use of L*+H by the native English speakers, we conducted a statistical analysis of the frequency with which different native English speakers employed this pitch accent. Table 4 shows the use of L*+H by the 12 native English speakers.

This table shows that some participants used L*+H multiple times, such as 11, 7, and 6. It indicates that these individuals were more inclined to use this pitch accent when expressing uncertainty. On the other hand, some participants used this pitch accent sparingly or not at all, like 2, 1, and 0, suggesting that these individuals were not familiar with the use of this pitch. The results may imply individual differences among native English speakers.

In addition, we found that pronunciation training may potentially impact the use of this pitch accent by native English speakers. In this study, the number of L*+H produced by the two native English speakers who received pronunciation training was 11 and 4, respectively.

Previous research has already established the existence of regional variants in American English (Arvaniti & Garding, 2007; Burdin, 2016; Burdin et al., 2018, 2022; Clopper & Smiljanic, 2011; Kelly, 2012; McLarty, 2018; Reed, 2020). Since the native English speakers in this study were all from Ohio, this result may be attributed to regional differences in American English.

Table 5 shows the error types of the four edge tones produced by the native English speakers. In terms of L-L%, NP and P were the primary error types, accounting for 22% of the total tokens. In relation to H-H%, in addition to NP and P, native English speakers may also use other edge tones, L-H% (8%) and L-L% (4%). Regarding L-H%, the percentage of L-L% was the highest, with a ratio of 59%. The percentages of P and NP accounted for 23%, making them the main error types.

According to the above description, the native English speakers exhibited some error types when producing L-L% and H-H%. Overall, their pronunciation

Table 4
L+H frequency analysis in the 12 native English speakers*

Frequency of occurrence	Participants
0	P02, P05, P09, P10
1	P11, P12,
2	P06, P07
4	P08
6	P04
7	P03
11	P01

Source: Authors' work

Table 5
Types of errors in the production of edge tones in different contexts by native English speakers

Types	Percentages of edge tones produced and error types (%)						
	L-L%	H-H%	H-L%	L-H%	!H-L%	P	NP
L-L%	77	0	0	0	0	9	13
H-H%	4	74	0	8	0	9	4
L-H%	59	3	3	12	0	9	14

Note. The numbers highlighted in grey refer to the percentages of correct responses from the participants. P refers to cases where other accented words follow the targeted words, and NP refers to edge tone caused by Na, with or without accented words following it.
Source: Authors' work

aligned with the American English model (Bartels, 1999; Hirschberg & Beckman, 1994). Nevertheless, there was a substantial deviation from the American English model due to the frequent use of L-L%.

Through further analysis, we found that L-L% was a potential edge tone of continuation for the native English speakers. Empirical findings from both perception and production studies revealed that among native speakers of British English, a notable prevalence of falling tone was observed (Hudson et al., 2019, 2022; Mok et al., 2016; Puga et al., 2017, 2018). In the perceptual experiment, falling tone reached 58%, while rising or level tone accounted for 42%. In the production experiments, the proportion of falling tones was 40%, while the proportion of rising or level tones was 60%. However, we also needed to recognize that this choice may be influenced by regional variations in American English for the same reason as L*+H.

From the discussion above, it can be observed that the native English speakers did not distinguish between H* and L+H*, and they rarely used L*+H and L-H%. These differences may be due to the influence of

regional variations in American English, which could be considered a significant factor affecting the differences between the native English speakers and the learners with varying levels of English proficiency.

Differences between the Three Learner Groups

The results of this study were not in line with previous studies (Albin, 2015; Bi & Chen, 2013; Graham & Post, 2018; Nguyen & Dao, 2018). This inconsistency was partly related to the methods used to analyze their results. Most studies had typically relied on frequency rather than inferential statistical analyses. In the study by Meng and Wang (2009), they used inferential statistics, and the results did not show the effect of English proficiency. We analyzed their interview recordings to identify the causes for the mismatch between the learners' performance in English intonation and pronunciation. As a result, we argued that the results could be attributed to their identical or similar learning experiences. In the following part, we will elaborate on this conclusion from four aspects.

Firstly, the learners were only exposed to native English speakers in college. The native English speakers mainly undertook the task of oral English teaching, with the purpose of cultivating the learners' ability in English communication. However, they rarely taught the learners knowledge of English intonation or corrected their problems in English intonation.

Secondly, they all took an English pronunciation course. The School of Foreign Languages offered this course when they were freshmen, and two Chinese English teachers undertook it. It covered all aspects of English pronunciation, such as English vowels, consonants, weak forms, stress, rhythm, and intonation. It may last for one or two semesters.

Thirdly, their knowledge of English intonation was limited. (1) Although English intonation is an integral part of this course, English teachers did not dedicate time and effort to teaching English intonation. There were only 2 or 4 lessons for this part. (2) The textbook introduces seven tones and their usages: high fall, low fall, low rise, high rise, level, fall-rise, and rise-fall (Zhu, 2003).[8] However, English teachers did not provide detailed explanations of these tones. (3) Though English teachers provided dialogues, passages, and songs to train the learners' English intonation, they focused on only two intonation patterns: rise and fall.

Fourthly, in addition to the textbook,

they had a handbook of English pronunciation exercises (Zhang, 2010). There are many materials about English intonation in it, including simple sentences, combined sentences, question tags, vocatives, parentheses, and reporting phrases. However, English teachers did not address these aspects in their pronunciation teaching.

Lastly, they had not developed the habit of using English intonation. (1) Most learners (55%) only remembered two tones, rise and fall, and associated using the two tones with sentence patterns, such as declarative and interrogative sentences. (2) A small number of learners (22%) could recall that a rising tone conveyed the meaning of uncertainty or continuation dependence. (3) Almost all the learners (97%) argued that English intonation was important and that English intonation had an impact on a speaker's meaning, attitude, or emotion. However, they reported that they had limited knowledge of English intonation. (4) Most learners (83%) seldom practiced oral English. They occasionally imitated what they heard while watching English movies or TV series or doing listening exercises.

Due to the similar learning experiences of the learners with varying levels of English proficiency, the extent to which their native language influenced them was also comparable. Chinese is a tone language. However, it has its own intonation system (Lin et al., 2020). There are significant differences between Chinese and English intonation systems, which may be a crucial

---

[8] English intonation presented in this textbook follows the standards of British English. In British English, no differentiation is made between H* and L+H* phonologically (Grabe, 1998; Grabe et al., 1998).

factor contributing to the poor performance of the learners with different English proficiency levels (Bartels, 1999; Lin et al., 2020; Lin & Li, 2011; Pierrehumbert & Hirschberg, 1990; Truckenbrodt, 2012).

CONCLUSION

In this study, we investigated the impact of English proficiency on the production of English pitch accents, edge tones, and intonation patterns by Chinese EFL learners under the AM framework. The results showed significant differences in four of the ten intonation types (L*, L-L%, H-H%, and L*H-H%) between the native English speaker group and each of the three learner groups, which may be related to the failure of the native English speakers to distinguish between H* and L+H* and the regional variations in American English (L*+H and L-H%). Furthermore, there were no significant differences between the three learner groups. However, we observed that the differences in mean scores varied between the three learner groups in the ten intonation types. The differences presented a complex situation where the influence of English proficiency was presented in H*, L*, and H*L-L% but not observed in other intonation types. One significant factor contributing to this result may be that the learners were not consistently exposed to English intonation.

Implication to Theory and Practice

The results of the study have significant theoretical and pedagogical implications. The theoretical implications can be demonstrated from two aspects. (1) This study reflects the independence between the learners' proficiency in spoken English and their proficiency in English intonation. In other words, the improvement of the learners' spoken English is not consistent with the enhancement of their intonation proficiency. It reflects the uneven development in learners' English pronunciation. (2) This study enriches the L2 intonation learning theory. As the learners at different proficiency levels have not gained enough experience in English intonation, they may be similarly influenced by their L1 intonation, resulting in no significant differences in the production of English intonation between the three learner groups.

The following suggestions on pedagogical implications may be considered based on insights from the findings and the interviews with the study participants. English teachers should consider making learners aware of the complexity of English intonation produced by native English speakers. Despite possible regional variations in the production of H*, L+H*, L*+H, and L-H% among native English speakers, we still recommend that learners strive to master these standard forms. It complies with the Teaching Guideline for English Majors (Guidance Committee for English Major Teaching, 2020) and contributes to understanding native English speakers' speech and expressing learners' viewpoints. It may also be helpful to increase learners' awareness of regional differences and differences in non-native varieties of

English. Also, English teachers should be encouraged to fully realize the importance of English intonation. Teaching is the main channel through which learners acquire English intonation knowledge. Therefore, English teachers should fundamentally understand the crucial role of English intonation in language communication to enhance learners' performance in English intonation. English teachers may need further training to ensure that they understand the similarities and differences between English and Chinese intonation systems so that they can transfer this awareness more effectively to learners. Moreover, English teachers should provide equally systematic English intonation instruction to learners of different proficiency levels, and they should adjust their instruction according to learners' varying degrees of difficulty in different intonation types. Last but not least, the textbooks may need to be updated to include the necessary components to increase learner awareness of the complexity of English Intonation. These textbooks should reflect the new changes in intonation research and provide more comprehensive practice materials.

Limitations and Recommendations for Future Research

This study has some limitations. We regarded the production of L*+H and L-H% by native English speakers as regional variations in American English. However, this conclusion still lacks direct evidence. Future studies should take into account the use of these intonation types among native English speakers in other regions to obtain a more comprehensive understanding of the production of English intonation by native English speakers.

In addition, there are great differences between Chinese and English intonation systems in terms of the structural elements and their meanings. However, we did not explore the specific L1 negative transfer. Future studies should systematically compare English and Chinese intonation systems to understand the potential impact of Chinese intonation on learners' production of English intonation.

ACKNOWLEDGEMENT

REFERENCES

Albin, A. L. (2015). *Typologizing native language influence on intonation in a second language: Three transfer phenomena in Japanese EFL Learners* [Doctoral dissertation, Indiana University]. IUScholarWorks Indiana University Libraries. https://hdl.handle.net/2022/20345

Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. In J. Cole & J. I. Hualde (Eds.), *Laboratory Phonology* (Vol. 9, pp. 547-576). Mouton de Gruyter.

Bartels, C. (1999). *The intonation of English statements and questions: A compositional interpretation*. Garland Publishing, Inc.

Barto, K. A. (2015). *Mandarin speakers' intonation in their L2 English* [Doctoral dissertation, University

of Arizona]. UA Theses and Dissertations. http://hdl.handle.net/10150/347161

Beckman, M. E., & Elam, A. G. (1997). *Guidelines for ToBI labelling*. The Ohio State University Research Foundation.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2006). The original ToBI system and the evolution of the ToBI framework. In S. A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 9-54). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199249633.003.0002

Benrabah, N. (1990). *Algerian intonational proficiency in English: An empirical study* [Doctoral dissertation, University College London]. UCL Discovery. https://discovery.ucl.ac.uk/id/eprint/10121651

Bi, R., & Chen, H. (2013). Longitudinal study of tone patterns by Chinese English learners. *Foreign Languages and Their Teaching*, *1*, 50-54.

Bu, Y. (2016). Chinese EFL learners' problems in their English intonation. *Foreign Language Teaching and Research*, *48*(4), 569-582.

Burdin, R. S. (2016). *Variation in form and function in Jewish English intonation* [Doctoral dissertation, The Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1470147757

Burdin, R. S., Holliday, N., & Reed, P. E. (2018). Rising above the standard: Variation in L+H* contour use across 5 varieties of American English. In *Proceeding of Speech Prosody 2018* (pp. 354-358). https://doi.org/10.21437/SpeechProsody.2018-72

Burdin, R. S., Holliday, N. R., & Reed, P. E. (2022). American English pitch accents in variation: Pushing the boundaries of mainstream American English-ToBI conventions. *Journal of Phonetics*, *94*, 1-19. https://doi.org/10.1016/j.wocn.2022.101163

Chen, H. (2006). Tone patterns in Chinese EFL learners' read speech. *Modern Foreign Languages, 4*, 418-425. https://doi.org/10.21437/SpeechProsody.2006-184

Chen, H. (2008). *On Chinese EFL learner's English intonation patterns*. Shanghai Education Press.

Chen, H. (2009). On English intonational studies in the 20th century. *Journal of PLA University of Foreign Languages*, *32*(3), 13-20.

Chien, H., & Fon, J. (2019). Are nuclear accents easier to acquire than prenuclear accents? — Using peak alignment in advanced Mandarin EFL learners as an example. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences 2019* (pp. 2856-2860). Australasian Speech Science and Technology Association Inc.

Chien, S., & Fon, J. (2020). On the learnability of nuclear and prenuclear accents? Using Taiwan Mandarin learners of English as an example. In *Proceeding of Speech Prosody 2018* (pp. 886-889). https://doi.org/10.21437/SpeechProsody.2020-181

China Information Association. (2021). *Standard of spoken English test on computer (T/CIIA 009—2021)*. National Group Standard Information Platform. https://www.ttbz.org.cn/StandardManage/Detail/48365/.

Clopper, C., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics*, *39*(2), 237-245. https://doi.org/10.1016/j.wocn.2011.02.006

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates, Publishers.

Grabe, E. (1998). *Comparative intonational phonology: English and German*. Ponsen and Looijen bv.

Grabe, E., Nolan, F., & Farrar, K. J. (1998). IViE - a comparative transcription system for intonational variation in English. In *Proceeding of 5th International Conference on Spoken Language Processing (ICSLP 1998;* paper 0099). https://doi.org/10.21437/ICSLP.1998-583

Graham, C., & Post, B. (2018). Second language acquisition of intonation: Peak alignment in American English. *Journal of Phonetics*, *66*, 1-14. https://doi.org/10.1016/j.wocn.2017.08.002

Green, S., & Salkind, N. (2014). *Using SPSS for Windows and Macintosh: Analysing and understanding data*. Pearson.

Grice, M. (1995). Leading tones and downstep in English. *Phonology*, *12*(2), 183-233. https://doi.org/10.1017/S0952675700002475

Guidance Committee for English Major Teaching. (2020). *Teaching guidelines for undergraduate programs in foreign language and literature majors at regular higher education institutions: English major teaching guidelines*. Foreign Language Teaching and Research Press.

Hirschberg, J., & Beckman, E. M. (1994). *The ToBI annotation conventions*. http://www.cs.columbia.edu/~julia/files/conv.pdf

Hua, C. (2019). *Acquisition of prosodic marking of information status by Mandarin-speaking EFL learners* [Doctoral dissertation, City University of Hong Kong]. CityU Scholar. https://scholars.cityu.edu.hk/en/theses/theses(e4a77858-733f-42fa-b076-454db444e4e5).html

Hudson, T., Setter, J., & Mok, P. (2019). Nuclear tones in Hong Kong and British English. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 320-323). Australasian Speech Science and Technology Association Inc.

Hudson, T., Setter, J., & Mok, P. (2022). English intonation in storytelling: A comparison of the recognition and production of nuclear tones by British and Hong Kong English speakers. *English World-Wide, 43*(3), 357-381. https://doi.org/10.1075/eww.21035.hud

Kelly, N. (2012). *Production and perception of intonation in Southern American English and West Coast American English*. https://www.academia.edu/2927808/Production_and_Perception_of_Intonation_in_Southern_American_English_and_West_Coast_American_English

Kretzschmar, W., Jr. (2004). Standard American English pronunciation. In E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A handbook of varieties of English 1: Phonology* (pp. 257-269). Mouton de Gruyter. https://doi.org/10.1515/9783110197181

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511808814

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, *31*(1), 81-112. https://doi.org/10.1016/S0095-4470(02)00073-6

Li, A., Qu, C., & Zhi, N. (2020). Intonation patterns of Wh-questions by EFL learners from Jinan dialectal region. In *Report of Phonetic Research* (pp. 82-87). Laboratory of Phonetics and Speech Science, Institute of Linguistics. http://paslab.phonetics.org.cn/wp-content/uploads/admin/2021-12.pdf

Li, A., Wan, X., Zhao, C., & Zhu, L. (2020). Phonological and phonetic realization of narrow focus in declarative sentences by Jinan EFL learners. In *Report of Phonetic Research* (pp. 88-92). Laboratory of Phonetics and Speech Science, Institute of Linguistics. http://paslab.phonetics.org.cn/wp-content/uploads/admin/2021-13.pdf

Lian, X. R., Jia, Y., Duan, W. J., & De Roeck, A. (2013). Phonetic realization of narrow focus by Beijing EFL learners in English yes-no

sentences. In *Report of Phonetic Research* (pp. 271-275). Laboratory of Phonetics and Speech Science, Institute of Linguistics. http://paslab. phonetics.org.cn/wp-content/files/research_report/2013/2013_30.pdf

Lin, M., Li, A., & Li, Z. (2020). *Chinese Intonation and the teaching of Chinese intonation to CFL learners.* Chinese Academy of Social Science Press.

Lin, M., & Li, Z. (2011). Focus and boundary in Chinese intonation. In *Report of Phonetic Research*. Laboratory of Phonetics and Speech Science, Institute of Linguistics. http://paslab. phonetics.org.cn/wp-content/files/research_report/2011/2011_05.pdf

Liu, X., & Chen, X. (2016). The acquisition of English pitch accents by Mandarin Chinese speakers as affected by boundary tones. In *Proceeding of Speech Prosody 2016* (pp. 956-960). https://doi. org/10.21437/speechprosody.2016-196

Lu, Y., & Miran, K. (2016). Prosody transfer in second language acquisition: Tonal alignment in the production of English pitch accent by Mandarin native speakers. *Tsing Hua Journal of Chinese Studies*, *46*(4), 785-816. https://doi.org/10.6503/THJCS.2016.46(4).04

McGory, J. T. (1997). *Acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers* [Doctoral dissertation, The Ohio State University]. OhioLINK. http://rave.ohiolink.edu/etdc/view?acc_num=osu1299009149

McLarty, J. (2018). African American language and European American English intonation variation over time in the American South. *American Speech*, *93*(1), 32-78. https://doi. org/10.1215/00031283-6904032

Meng, X., & Wang, H. (2009). Boundary tone patterns in Chinese English learners' read speech. *Foreign Language Teaching and Research*, *41*(6), 447-451.

Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In J. Trouvain & U. Gut (Eds.), *Non-native prosody: Phonetic descriptions and teaching practice* (pp. 53-76). Mouton De Gruyter. https://doi. org/10.1515/9783110198751.1.53

Mennen, I. (2015). Beyond segments: Towards a L2 intonation learning theory. In E. Delais-Roussarie, Avanzi, M., & S. Herment (Eds.), *Prosody and language in contact: L2 acquisition, attrition and languages in multilingual situations* (pp. 171-188). Springer. https://doi. org/10.1007/978-3-662-45168-7_9

Mennen, I., Chen, A., & Karlsson, F. (2010). Characterising the internal structure of learner intonation and its development over time. In K. Dziubalska-Kołaczyk, M. Wrembel, & M. Kul (Eds.), *Proceedings of New Sounds 2010 (6th International Symposium on the Acquisition of Second Language Speech)* (pp. 319-324). Adam Mickiewicz University.

Mok, P., Yin, Y., Setter, J., & Nayan, N. M. (2016). Assessing knowledge of English intonation patterns by L2 speakers. In *Proceeding of Speech Prosody 2016* (pp. 543-547). https://doi. org/10.21437/SpeechProsody.2016-111

Nguyen, A., & Dao, D. (2018). The acquisition of question intonation by Vietnamese learners of English. *Asian-Pacific Journal of Second and Foreign Language Education*, *3,* Article 3. https://doi.org/10.1186/s40862-018-0044-4

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation.* MIT Press.

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 271-311). MIT Press. https:// doi.org/10.7551/mitpress/3839.003.0016

Puga, K., Fuchs, R., Hudson, T., Setter, J., & Mok, P. (2018). The perception-production link in

intonation: Evidence from German learners of English. In *Proceeding of Speech Prosody 2018* (pp. 685-689). https://doi.org/10.21437/SpeechProsody.2018-139

Puga, K., Fuchs, R., Setter, J., & Mok, P. (2017). The perception of English intonation patterns by German L2 speakers of English. In *Interspeech 2017* (pp. 3241-3245). https://doi.org/10.21437/Interspeech.2017-1279

Reed, P. E. (2020). Inter- and intra-regional variation in intonation: An analysis of rising pitch accents and rootedness. *The Journal of the Acoustical Society of America*, *147*(1), 616-626. https://doi.org/10.1121/10.0000576

Safrudiannur, S. (2020). *Measuring teachers' beliefs quantitatively: Criticizing the use of Likert scale and offering a new approach*. Springer Spektrum. https://doi.org/10.1007/978-3-658-30023-4

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C. C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). TOBI: A standard for labeling English prosody. In *The Second International Conference on Spoken Language Processing, ICSLP 1992* (pp. 867-870). https://doi.org/10.21437/ICSLP.1992-260

Truckenbrodt, H. (2012). Semantics of intonation. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (pp. 2039-2069). Mouton de Gruyter.

Veilleux, N., Barnes, J., Brugos, A., & Shattuck-Hufnagel, S. (2012). Perceptual foundations for naturalistic variability in the prosody of synthetic speech. In *Proceeding of Interspeech 2012* (pp. 2534-2537). https://doi.org/10.21437/Interspeech.2012-656

Ward, G., & Hirschberg, J. (1985). Implicating uncertainty - the pragmatics of fall-Rise intonation. *Language*, *61*(4), 747-776. https://doi.org/10.2307/414489

Ward, G., & Hirschberg, J. (1986). Reconciling uncertainty with incredulity: A unified account of the L*+H L H% intonational contour. *The Annual Meeting of the Linguistic Society of America,* New York. https://files.eric.ed.gov/fulltext/ED285390.pdf

Wells, J. C. (2006). *English intonation: An introduction*. Cambridge University Press.

Works, N. (1985). *A study of the use of English sentence stress for the comprehension of meaning by adult second language learners* [Unpublished doctoral dissertation]. The University of Illinois.

Xia, Z., & Mu, F. (2008). A study of boundary tone use at turn-taking in Chinese EFL learners's conversations. *Journal of PLA University of Foreign Languages*, *31*(1), 48-54.

Zárate-Sández, G. A. (2015). *Perception and production of intonation among English-Spanish bilingual speakers at different proficiency levels* [Doctoral dissertation, Georgetown University]. Digital Georgetown. http://hdl.handle.net/10822/761510

Zhang, G. (2010). *A handbook of English pronunciation exercises for first-year college students*. Foreign Language Teaching and Research Press.

Zhu, X. (2003). *A concise guide to contemporary English pronunciation*. Foreign Language Teaching and Research Press.

The Effect of Pre-Competition Anxiety on Tennis Accuracy Shot Performance Among Malaysian Varsity Tennis Players

Wan Ahmad Munsif Wan Pa[1]*, Mohamad Rahizam Abdul Rahim[2], Nur Shakila Mazalan[1], Mohamad Nizam Nazarudin[1], Wan Rezawana Wan Daud[3] and Mohd Firdaus Abdullah[4]

[1]*Faculty of Education, Center of Education & Community Wellbeing, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*

[2]*Faculty of Sports Science and Recreation, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*
[3]*Pusat GENIUS@Pintar Negara, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
[4]*Faculty of Sports and Exercise Science, Universiti Malaya, 50603 Kuala Lumpur, Malaysia*

ABSTRACT

Athletes with high levels of anxiety may take wrong actions and cause athletes to not be able to perform well. Specifically, when pre-competition anxiety problems occur, this situation can affect the accuracy of the shots in tennis. Therefore, this study identifies the effect of pre-competition anxiety on tennis forehand (FH) and backhand (BH) accuracy shots among Malaysian varsity tennis players. This study uses quantitative cross-sectional survey research. This research used the Competitive State Anxiety Inventory (CSAI-2) questionnaire and the tennis forehand (FH) and backhand (BH) accuracy shot instrument to survey 60 respondents utilizing a convenient sampling technique. The descriptive data showed high cognitive and physical anxiety and low self-confidence 24 hours before the competition. Researchers used the Statistical Package for the Social Sciences (SPSS) software, version 25, to perform inferential statistical analysis related to the Simple Linear Regression Test and concluded that pre-competition anxiety had a significant impact on tennis forehand (FH) and backhand (BH) accuracy shots, accounting for 10.1% of the variance. These findings imply that an excessive level of pre-competition anxiety has a significant negative impact on tennis forehand (FH) and backhand (BH) accuracy shots. The conclusion is that pre-competition anxiety poses a significant threat to the accuracy of tennis forehand (FH) and backhand (BH) accuracy shot performance. Solutions are desperately needed, such as meditation, self-talk,

Wan Ahmad Munsif Wan Pa, Mohamad Rahizam Abdul Rahim, Nur Shakila Mazalan,
Mohamad Nizam Nazarudin, Wan Rezawana Wan Daud, and Mohd Firdaus Abdullah

imagery, and sports massage intervention, which may control pre-competition anxiety and subsequently accurately perform tennis forehand (FH) and backhand (BH) shots.

Keywords: Accuracy, anxiety, pre-competition, tennis, varsity

INTRODUCTION

The impact of pre-competition anxiety on players' performance necessitates a comprehensive examination of psychological factors to facilitate effective adaptation and good responses (Kerketta, 2015). Pre-competition anxiety is a common experience among tennis players, and it can have both positive and negative effects on motivation (Juspah et al., 2023; Marih et al., 2023). By adopting effective strategies to manage anxiety and channel it into a positive competitive drive, players can enhance their motivation and perform at their best on the tennis court. Pre-competition anxiety is a serious stumbling block for players (Jastari et al., 2023). Even if a person is a professional athlete, he or she is not exempted from pre-competition anxiety difficulties (Saadan et al., 2015). Cognitive anxiety, physical anxiety, and self-problems are the three primary components of pre-competition anxiety (Saadan et al., 2016). With the advent of science, particularly psychological areas, anxiety issues have become an important aspect of mental health experts based on empirical data and scientific traditions (Misran et al., 2021).

In sports performance, it has been said that Malaysian athletes emerged as the overall champions at the 2017 Southeast Asian Games, securing 145 gold, 92 silver, and 86 bronze medals (Bernama, 2017). This claim is based on evaluating Malaysia's athletic accomplishments throughout the preceding five-year period. Nevertheless, amid the fervor, the national tennis athletes secured just a bronze medal out of the five categories contested at the 2017 Games. It raises a significant concern among researchers regarding the disparity in accomplishments between tennis players and gold medalists from the 2001 SEA Games. Based on a report by Buletin Utama (2018), it is noted that the most recent instance of a national tennis player winning a gold medal at the Southeast Asian Games occurred in 2001. Selvam Veerasingam, a former national tennis player, accomplished this noteworthy achievement. Paradoxically, despite a significant allocation of government financing towards the development of national sports, no tennis player has achieved the feat of securing a gold medal at the Southeast Asian (SEA) Games, even after the passage of 16 years. Additionally, according to statistical reports from the Association of Tennis Professionals (ATP) and Women's Tennis Associations (WTA) for 2019, no tennis players from this country have ranked among the world's top 100 tennis players. It has a significant impact on tennis practitioners and associations in Malaysia. As a result, stakeholders in the Malaysian tennis industry must find a solution to ensure that Malaysian tennis can one day compete on a global level.

Tennis players who struggle with anxiousness have a tough time making good strokes (Mullen & Hardy, 2000). The combination of ball speed and accuracy is critical in determining the quality of tennis shot techniques (Landlinger et al., 2012; Vergauwen et al., 2004). Two tennis abilities must be mastered: forehand (FH) and backhand (BH) shots to be a competent tennis player (Mavvidis et al., 2010). Muhamad et al. (2011) stated that even tennis players with competitive status struggle to grasp forehand (FH) and backhand (BH) shot techniques. As a result, researchers chose to focus on technical ability in tennis, specifically the ability to execute accurate forehand (FH) and backhand (BH) shots among Malaysian university tennis players.

The second distinction is that the effects of pre-competition anxiety on forehand (FH) and backhand (BH) shot accuracy continue to receive little attention, as previous research has focused on the relationship between pre-competition anxiety and emotions, self-confidence, cortisol hormone changes, and heart rate in athletes from a variety of sports (Ayuso-Moreno et al., 2020; Carolina-Paludo et al., 2020; Mojtahedi et al., 2023; Kang & Jang, 2018; Yang et al., 2020). Thus, researchers are greatly motivated to investigate the influence of pre-competition anxiety on the capacity of university-level tennis players in Malaysia to perform forehand (FH) and backhand (BH) shot accuracy. Therefore, this research aims to identify levels of pre-competition anxiety among university tennis players and to determine the effect of pre-competition

anxiety on the ability to perform forehand (FH) and backhand (BH) shot accuracy among university tennis players. Hence, the null hypothesis was created to provide an initial anticipation because there was limited evidence of pre-competition anxiety in tennis, specifically among Malaysian Varsity tennis players. The null hypothesis was that pre-competition anxiety does not have a significant impact on the ability to perform forehand (FH) as well as backhand (BH) shot accuracy among university tennis players.

Tennis has been identified as a fast, powerful, and extremely dynamic sport (Fernández-Fernández et al., 2014; Roetert & Kovacs, 2019). However, surveys indicate that the impact of pre-competition anxiety on forehand (FH) and backhand (BH) stroke accuracy continues to attract less attention than service shot accuracy (Landlinger et al., 2012). Accurate forehand (FH) and backhand (BH) shots are the most critical aspects that can provide an edge over performance and enable tennis players to reach a higher level of competition (Kibler, 2009). Additionally, tennis shots are classified as extremely complicated motor abilities that need the generating of force and the capacity to transfer that force throughout the body, referred to as kinetic chains (Kibler, 2014). Thus, academics have focused on gaining an additional understanding of pre-competition anxiety about the accuracy of forehand (FH) and backhand (BH) shots.

Additionally, previous researchers have indicated that pre-competition anxiety

as a 'threat' to players is more likely to arise in individual sports than in team sports (Judge et al., 2016; Sutarsyah, 2017). Although anxiousness is typically connected with an unpleasant taste and is seen negatively, some players have demonstrated a positive response and have increased their performance quality (Doron & Martinent, 2016). Players have the opportunity to compare themselves to competitors prior to the competition, as all players warm up in the same pre-competition area. If players believe they can win, such beliefs can help them perform better and control their fear (Uphill et al., 2019). On the other side, if players examine a scenario and determine that they may be unable to manage a particular stressor, this is classed as an indication of anxiety affecting athletic performance. Players react ineffectively in these situations, which has a detrimental effect on their psychological abilities, muscles, and even capabilities.

Three elements of pre-competition anxiety are explained by the Multidimensional Anxiety Theory: cognitive anxiety, physical anxiety, and performance self-confidence. Cognitive anxiety, in particular, is a mental component generated by the dread of poor self-esteem and self-esteem threats (Jarvis, 2002; Martinent et al., 2010). On the other hand, somatic anxiety is associated with the physiological features of mood anxiety that are intimately connected with physiological awakening. Increased heart rate, respiration rate, sweaty hands, abdominal discomfort, and muscle tension are all physiological responses (Cox, 1998). By contrast, self-confidence refers to a person's trust in his or her ability to overcome the difficulties inherent in the task at hand (Martens et al., 1990). Martens et al. (1990) identified three distinct types of graphs that must be understood (Woodman & Hardy, 2003). Specifically, it is hypothesized that cognitive worry has a negative linear connection with performance. In comparison, somatic anxiety is expected to have a quadratic association (U-inverted) with performance, whereas self-confidence has a positive straight-line relationship.

According to A. Kumar (2017), pre-competition anxiety refers to uncomfortable or destructive awakenings experienced the day before the competition begins. Pre-competition anxiety must be thoroughly investigated to determine the best solution or method. Additionally, this effort is necessary since players' anxiety levels can affect their performance in the tournaments they compete in (Mojtahedi et al., 2023). Athan and Sampson (2013) stated that players are frequently overcome with pre-competition anxiety prior to competing in a tournament. It results from physical difficulties, fear of defeat, feelings of inadequacy, loss of control, and guilt. However, some findings indicate that players' cognitive anxiety, somatic anxiety, and self-confidence are not significantly different (Şekeroğlu, 2019).

Numerous scholars have already investigated the relationship between time changes and anxiety using a time-to-competition perspective. According to Spielberger (2021), pre-competition anxiety is a state of mind that players

endure for a week, several hours, and a few minutes before the competition begins. Pre-competition anxiety, for instance, can be measured across three months (John et al., 2011), 30 days (Souza et al., 2019), seven days (Radzi et al., 2018), two days (Turksoy et al., 2012), one day (Saadan et al., 2015), and one hour (Muñoz et al., 2017). Other scholars performing research on a variety of sports have also reported adopting a similar approach (Carolina-Paludo et al., 2020; Mahadevan & Halankar, 2021; Mohebi et al., 2019; Şekeroğlu, 2019). According to Saadan et al. (2015), players often experience pre-competition anxiety within the 24 hours preceding the start of the tournament.

Conceptual Framework
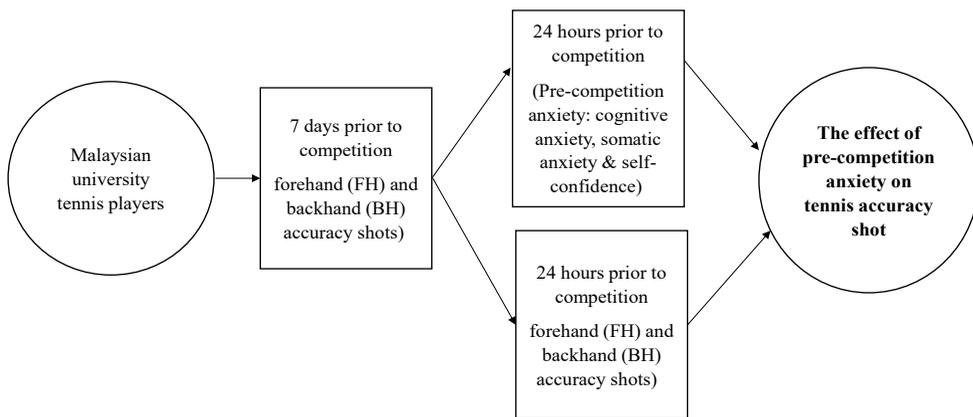
Figure 1. Conceptual framework
Source: Authors' work

The conceptual framework is a generative framework that encapsulates the research process in its essence (Adom et al., 2018). The conceptual framework depicted in Figure 1 was created for each study's processes and the underlying variables used to review the new search results. The study concept is derived from the researchers' reading of the literature.

METHODOLOGY

The researchers used a cross-sectional survey design combined with quantitative methods and a convenient sampling technique. This study surveyed 60 respondents, who were Malaysian university tennis players. All these respondents were selected from 4 public universities around the Klang Valley. Data collection on pre-competition anxiety was carried out 24 hours before the Universiti Malaysia Sports Council (MASUM) Tennis Tournament. Meanwhile, data collection on the accuracy of forehand (FH) and backhand (BH) shots was done in two sessions, seven days before and one day before the competition. Data collection on the accuracy of forehand (FH) and

backhand (BH) shots performed twice is to obtain baseline data in the first session. The researcher wants to see a change from the baseline data in the second session. To ensure that the ethics of the study are complied with, the researcher has made a letter of application together with a letter of acknowledgment of data collection to the coaches and the tournament organizers. Later, the researcher also provided consent and ethics acknowledgment forms to the respondents to be directly involved in this study. Among the data collection processes implemented are identifying study respondents, obtaining the organizer's permission and the respondent's certificate and quoting data.

The Statistical Package for the Social Sciences (SPSS) version 25 was used for data entry and analysis. A simple descriptive analysis was conducted using numbers and percentages. To discover the evidence of pre-competition anxiety on tennis accuracy shot performance, simple linear regression was implemented as a test of significance, with a threshold of less than 0.05 considered significant.

Instruments

A self-administered questionnaire known as The Competitive State Anxiety Inventory-2 (CSAI-2), developed by Martens et al. (1990), was adopted in this research and distributed to the respondents. The instrument has three components, namely cognitive anxiety, somatic anxiety, and self-confidence. Each component contained 9 question items, and the total items were 27 questions. The order of the items in this questionnaire was arranged randomly. The items for the cognitive anxiety component were 1, 4, 7, 10, 13, 16, 19, 22, and 25. The items for the somatic anxiety component were 2, 5, 8, 11, 14, 17, 20, 23, and 26. The items for the self-confidence component are 3, 6, 9, 12, 15, 18, 21, 24 and 27. All items have a Likert scale of 1 to 4, from strongly disagree to strongly agree. All items (27) underwent pilot testing, and Cronbach's Alpha results showed that the reliability value of each item was 0.65 and above. According to Chua (2014), the value of Cronbach's Alpha coefficient should be between 0.65 and 0.95, which is high and satisfactory.

Instruments for testing the accuracy of forehand (FH) and backhand (BH) shots were adapted from Wiebe (1980) as cited in Mavvidis et al. (2010) with a validity value of $r = 0.71$ and the reliability value was > 0.88. The equipment required to conduct forehand (FH) and backhand (BH) shot accuracy tests are tennis courts, tennis nets, tennis balls, tennis rackets, measuring tapes, masking tape, and cone markers. The diameter of each zone is 1.4 meters long X 1.8 meters wide. The tennis courts (no man's land area) were divided into several zones for forehand (FH) and backhand (BH) shot evaluation.

During the assessment, the tennis player stands on the baseline and hits the ball sent by the coach from the baseline on the opposite side of the court. Each player aims a shot in a specific direction from the opposite side of the court to get the highest

score. For each test, each player is only allowed to perform ten (10) forehand shots (FH) and ten (10) backhand shots (BH). Each successful shot was rated from 0 to 9 points, based on where the tennis ball landed after being hit. If the ball is stuck in the net or does not land in the specified area, the points scored are 0.

If the player fails to hit at the start of the attempt, the shot is considered a failure or rated 0. However, the player can make a repeat shot if the tennis ball gets stuck in the net (the white part) and passes the opposite court. Test scores were assessed based on the number of points accumulated after ten times of forehand shot (FH) and ten times of backhand shot (BH).

RESULTS AND DISCUSSION

Sixty university tennis players completed the study questionnaire; 40 (66.7 %) were male tennis players, and 20 (33.3%) were female tennis players. The range of years of experience in training and competing in tennis was less than five years: 15%, 5 to 10 years: 36.7%, 10 to 15 years: 38.3%, and more than 15 years: 10%. Descriptive analysis was performed to answer the first research questions.

Figure 2 shows the level of pre-competition anxiety (mean): cognitive anxiety (29.15) and somatic anxiety (28.57) are high, while self-confidence (14.30) is low 24 hours before the competition begins.

Figure 2. Pre-competition anxiety of Malaysian university tennis players

Source: Authors' work
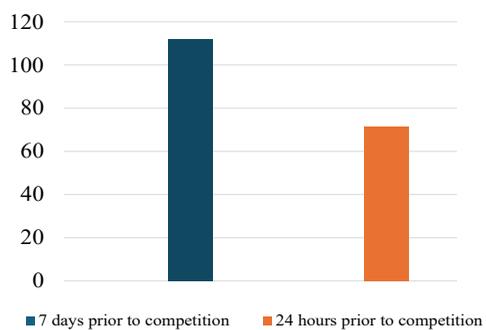
Figure 3. Tennis forehand (FH) and backhand (BH) accuracy shot

Source: Authors' work

Based on Figure 3, the graph shows the level of the tennis forehand (FH) and backhand (BH) shot accuracy among Malaysian university tennis players (mean) decreased from 111.98 to 71.50 in seven days.

As shown in Table 1, the Simple Linear Regression yielded a significant value of

$p = 0.013$ ($P < 0.05$). In conclusion, these findings indicate that the null hypothesis (Ho1) was rejected; pre-competition anxiety had significantly contributed to tennis forehand (FH) and backhand (BH) accuracy shot performance. Therefore, the researchers referred to the Coefficients and R^2 value test to determine how pre-competition

Wan Ahmad Munsif Wan Pa, Mohamad Rahizam Abdul Rahim, Nur Shakila Mazalan,
Mohamad Nizam Nazarudin, Wan Rezawana Wan Daud, and Mohd Firdaus Abdullah

Table 1

Simple linear regression test of the impact of pre-competition anxiety on tennis forehand (FH) and backhand (BH) accuracy shot

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.515	1	1.515	6.514	0.013
	Residual	13.485	58	0.233		
	Total	**15.000**	**59**			

Source: Authors' work

anxiety use affects the tennis forehand (FH) and backhand (BH) accuracy shot among Malaysian university tennis players.

As presented in Table 2, the Coefficients test showed that the correlation between the dependent and predictor variables was $R = 0.318$. Linear Regression Analysis was found to be significant [$F (1,59) = 6.514$, $p = 0.013$ ($P < 0.05$)], and R^2 equivalent to 0.101 showed that 10.1% of the change in the dependent variable of tennis forehand (FH) and backhand (BH) accuracy shot was due pre-competition anxiety. Thus, the researchers concluded that pre-competition anxiety significantly impacted tennis forehand (FH) and backhand (BH) accuracy by 10.1%. It can decrease

tennis shot accuracy performance, self-efficacy and motivation among Malaysian university tennis players. These findings suggest that pre-competition anxiety has a significant negative effect on tennis forehand (FH) and backhand (BH) accuracy shots, and solutions, such as self-talk, imagery, progressive muscle relaxation and sports massage, are urgently needed to overcome it. Interventions that can assist Malaysian university tennis player to regulate and lessen their pre-competition anxiety are also critically needed.

The present study's findings significantly contribute to the Multidimensional Theory of Anxiety by demonstrating a clear, quantifiable effect of pre-competition

Table 2

Coefficients test

	Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	R	R^2	Percentage (%)
		B	Std. Error	Beta					
1	*(predictor)*	74.756	1.277		58.530	0.001			
	Pre-competition anxiety	0.045	0.180	0.318	2.552	0.013	0.318	0.101	10.1

Note. Predictor: Pre-competition anxiety; Dependent variable: tennis forehand (FH) and backhand (BH) accuracy shot

Source: Authors' work

anxiety on the performance of tennis shots, specifically forehand (FH) and backhand (BH) accuracy. Our results, indicating that pre-competition anxiety accounted for 10.1% of the variance in shot accuracy, underscore the theory's assertion that cognitive and somatic anxiety can detrimentally affect athletic performance (G. Jones, 2000; Martinez & Scott, 2021).

In practice, managing pre-competition anxiety becomes paramount for coaches and sports psychologists working with varsity tennis players. Implementing targeted interventions, such as sports massage (Pa et al., 2021), cognitive-behavioral strategies (Liu et al., 2022) and mindfulness-based stress reduction (MBSR) techniques (B. J. Jones et al., 2020), could mitigate anxiety's negative impact. Adopting personalized anxiety management programs, as suggested by Patel and Smith (2024), can further refine athletes' preparatory routines to enhance performance.

This investigation markedly enhances our comprehension of the Multidimensional Theory of Anxiety by elucidating the quantifiable influence of pre-competition anxiety on the accuracy of tennis shots. It corroborates that cognitive and somatic anxieties detrimentally impact performance, accentuating the necessity for customized anxiety management interventions for athletes. Future studies should tackle existing limitations, such as the dependence on self-reported anxiety metrics and a restricted participant demographic, by integrating more objective measures and expanding the participant base to bolster

generalizability and thoroughly examine causality.

CONCLUSION

The present study investigated the relationship between pre-competition anxiety and performance accuracy in forehand and backhand tennis shots among university tennis players in Malaysia. Our findings reveal that both cognitive and somatic anxiety are significantly elevated 24 hours before competition, correlating strongly with a reduction in shot accuracy. This research not only supports but also quantifies the impact of pre-competition anxiety as outlined in the Multidimensional Theory of Anxiety, showing a tangible decrease in performance of 10.1% in shot accuracy.

Significantly, the study underscores the inverse relationship between self-confidence and anxiety levels, highlighting an area ripe for intervention. The decrease in tennis shot accuracy over a period leading up to competition suggests that anxiety management should be a focal point in athletic preparation, particularly in the context of university-level sports where athletes may be more susceptible to psychological pressures.

The practical implications of this study are profound. For coaches and sports psychologists, there is a clear mandate to develop and implement strategies that can help reduce anxiety levels in athletes. Such strategies could include psychological training focusing on anxiety management techniques, confidence-building exercises,

and possibly incorporating relaxation practices into the pre-competition routines. These interventions can not only enhance performance but also contribute to the overall wellbeing of the athletes.

Furthermore, the study's implications extend to the scheduling and structure of sporting events. Sports organizers and coaches are encouraged to consider the timing of anxiety peaks and structure event schedules that allow athletes adequate time to engage in psychological preparation and relaxation. This approach could help optimize performance and manage athletes' psychological health, which is crucial for their long-term development and success in sports.

In conclusion, the study provides compelling evidence of the critical impact of pre-competition anxiety on performance in sports. It also offers a blueprint for practical interventions that can help mitigate these effects. Future research should explore longitudinal interventions to confirm the effectiveness of these strategies and potentially uncover additional factors that could influence the anxiety-performance relationship in athletes. By continuing to explore these dynamics, athletic performance at all levels can be supported, fostering an environment where athletes can excel psychologically and physically.

Implication of the Study

This study contributes to the existing literature on the Multidimensional Theory of Anxiety by empirically validating the theory's application within a specific athletic context at the university level of tennis. Our findings underscore the theory's assertion that cognitive and somatic anxiety can significantly affect performance (Martens et al., 1990). The high levels of cognitive and somatic anxiety observed in players 24 hours before a competition correlate with a notable decline in shot accuracy, reinforcing the theory's distinction between these anxiety types and their unique impacts on performance. This study extends the theory by quantifying the impact (10.1% decrease in shot accuracy) and highlights the critical need to manage these anxiety types differently to optimize athletic performance.

From a practical standpoint, our results have significant implications for coaches and sports psychologists working with university tennis players. The pronounced decrease in forehand and backhand shot accuracy as a function of pre-competition anxiety suggests that interventions aimed at reducing anxiety levels could be crucial in enhancing performance. Coaches might consider implementing tailored anxiety-reduction workshops or individualized psychological support focusing on cognitive and somatic anxiety management techniques (Weinberg & Gould, 2015).

Furthermore, our study highlights the importance of fostering self-confidence among athletes, which was inversely related to anxiety levels. It could be beneficial to develop routine psychological training sessions that include confidence-building activities. Such sessions could focus on visualization techniques, positive self-talk, and setting achievable performance goals,

which may help mitigate the negative impacts of anxiety (Hanton et al., 2008).

Additionally, given the critical timing of anxiety's impact (24 hours before competition), these findings suggest that timing interventions to specifically address pre-competition anxiety could be key. It might include relaxation techniques such as deep breathing, progressive muscle relaxation, or mindfulness exercises tailored to the athletes' pre-competition routines (Gardner & Moore, 2017).

Finally, this study's implications extend to tournament preparations. Understanding that anxiety peaks close to competition time, organizers and coaches could redesign the scheduling and programming of events to include periods of mental relaxation and preparedness, potentially improving overall performance outcomes.

Limitations and Recommendations for Future Research

While the study provides valuable insights, several limitations must be acknowledged. First, the reliance on self-reported measures of anxiety may introduce bias or inaccuracies in reporting levels of pre-competition anxiety (Hamilton & White, 2022). Secondly, the sample size, limited to Malaysian varsity tennis players, may not fully represent the broader athlete population, potentially affecting the generalizability of the findings (S. Kumar & Lee, 2023). Finally, the study's cross-sectional design precludes the determination of causality between pre-competition anxiety and shot accuracy (Nguyen & Tran, 2023).

Future research should consider longitudinal designs to explore the causal relationships between pre-competition anxiety and tennis performance more comprehensively to build upon the current study's findings (Zhang & Wang, 2024). Expanding the sample to include players from various cultural backgrounds and skill levels could enhance the generalizability of the results (Roberts & Jackson, 2022). Moreover, incorporating objective measures of anxiety, such as physiological indicators, alongside self-reported data could offer a more nuanced understanding of anxiety's impact (Singh & Gupta, 2023). Investigating the efficacy of specific anxiety management interventions in improving tennis shot accuracy would also provide practical guidance for coaches and athletes (Morales & Fernandez, 2023).

ACKNOWLEDGEMENTS

REFERENCES

Adom, D., Hussain, E. K., & Joe, A. A. (2018). Theoretical and conceptual framework: Mandatory ingredients theoretical and conceptual framework: Mandatory ingredients engineering. *International Journal of Scientific Research, 7*(1), 93-98.

Athan, A. N., & Sampson, U. I. (2013). Coping with pre-competitive anxiety in sports competition. *European Journal of Natural and Applied Sciences, 1*(1), 1-9.

Ayuso-Moreno, R., Fuentes-Garcia, J. P., Collado-Mateo, D., & Villafaina, S. (2020). Heart rate variability and pre-competitive anxiety according to the demanding level of the match in female soccer athletes. *Physiology & Behavior, 222,* Article 112926. https://doi.org/10.1016/j.physbeh.2020.112926

Bernama. (2017. August 30). 145 pingat emas penyeri Merdeka 2017 [145 gold medals for Merdeka 2017]. *Malaysiakini*. https://www.malaysiakini.com/sukan/393740

Buletin Utama. (2018, December 8). Banyak kelemahan dalam arena tenis kebangsaan [Many weaknesses in the national tennis arena; Video]. Facebook. https://www.facebook.com/watch/?v=1088734417966142

Carolina-Paludo, A., Nunes-Rabelo, F., Maciel-Batista, M., Rúbila-Maciel, I., Peikriszwili-Tartaruga, M., & Simões, A. C. (2020). Game location effect on pre-competition cortisol concentration and anxiety state: A case study in a futsal team. *Revista de Psicología del Deporte, 29,* 105-112.

Chua, Y. P. (2014). *Asas statistik penyelidikan: Kaedah dan statistik penyelidikan* [Fundamentals of research statistics: Research methods and statistics]. McGraw-Hill Education (Malaysia).

Cox, R. H. (1998). *Sport psychology: Concepts and applications*. McGraw-hill.

Doron, J., & Martinent, G. (2016). Trajectories of psychological states of women elite fencers during the final stages of international matches. *Journal of Sports Sciences, 34*(9), 836-842. https://doi.org/10.1080/02640414.2015.1075056

Fernández-Fernández, J., Ulbricht, A., & Ferrauti, A. (2014). Fitness testing of tennis players: How valuable is it? *British Journal of Sports Medicine, 48*(Suppl 1), i22-i31. https://doi.org/10.1136/bjsports-2013-093152

Gardner, F. L., & Moore, Z. E. (2017). *The psychology of enhancing human performance: The mindfulness-acceptance-commitment (MAC) approach.* Springer Publishing Company.

Hamilton, P. J., & White, K. M. (2022). Challenges in measuring sports anxiety: A review. *Journal of Anxiety and Performance, 4*(1), 45-59.

Hanton, S., Neil, R., & Mellalieu, S. D. (2008). Recent developments in competitive anxiety direction and competition stress research. *International Review of Sport and Exercise Psychology, 1*(1), 45-57. https://doi.org/10.1080/17509840701827445

Jarvis, M. (2002). *Sport psychology*. Routledge.

Jastari, J., Nazarudin, M. N., & Noordin, Z. (2023). Pre-competition anxiety levels among Rahim Kajai College athletes in intercollegiate sports competitions. *Jurnal Personalia Pelajar, 26*(1), 151-159. https://doi.org/10.17576/personalia.2601.2023.04

John, S., Verma, S. K., & Khanna, G. L. (2011). The effect of mindfulness meditation on HPA-Axis in pre-competition stress in sports performance of elite shooters. *National Journal of Integrated Research in Medicine, 2*(3), 15-21.

Jones, B. J., Kaur, S., Miller, M., & Spencer, R. M. C. (2020). Mindfulness-based stress reduction benefits psychological well-being, sleep quality, and athletic performance in female collegiate rowers. *Frontiers in Psychology*, *11*(September), Article 572980. https://doi.org/10.3389/fpsyg.2020.572980

Jones, G. (2000). Stress and anxiety. In S. J. Bull (Ed.), *Sport psychology: A self-help guide* (pp. 31-51). Crowood.

Judge, L. W., Urbina, L. J Hoover, D. L., Craig, B. W Judge, L. M., Leitzelar, B. M. Pearson, D. R., Holtzclaw, K. A., & Bellar, D. M. (2016). The impact of competitive trait anxiety on collegiate powerlifting performance. *Journal of Strength and Conditioning Research, 30*(9), 2399-2405. https://doi.org/10.1519/JSC.0000000000001363

Juspah, S. N., Nazarudin, M. N., Noordin, Z., Mazalan, N. S., & Pa, W. A. M. W. (2023). Motivation of sports participation among sports and recreation students at Universiti Kebangsaan Malaysia. *International Journal of Academic Research in Progressive Education and Development, 12*(3), 89-98.

Kang, H., & Jang, S. (2018). Effects of competition anxiety on self-confidence in soccer players: Modulation effects of home and away games. *Journal of Men's Health, 14*(3), 62-68. https://doi.org/10.22374/1875-6859.14.3.9

Kerketta, I. (2015). A comparative study of sports competition anxiety between district levels male volleyball and soccer players. *International Journal of Physical Education, Sports and Health, 1*(3), 53-55.

Kibler, B. (2009). The 4000-watt tennis player: power development for tennis. *Medicine Science Tennis, 14,* 5-8.

Kibler, B. (2014). Understanding the kinetic chain in tennis performance and injury. *Aspetar: Sports Medicine Journal 3*, 492-497.

Kumar, A. (2017). Pre-competitive anxiety: An emotional experience and its pitfalls. *Health*, *3*, 303-304.

Kumar, S., & Lee, H. (2023). Cultural influences on sports anxiety: A comparative study. *International Journal of Sports Psychology, 54*(1), 75-89.

Landlinger, J., Stöggl, T., Lindinger, S., Wagner, H., & Müller, E. (2012). Differences in ball speed and accuracy of tennis groundstrokes between elite and high-performance players. *European Journal of Sport Science, 12*(4), 301-308. https://doi.org/10.1080/17461391.2011.566363

Liu, F., Zhang, X., & Wei, L. (2022). Cognitive-behavioral strategies in managing pre-competition anxiety among athletes: A meta-analysis. *Performance Psychology Journal, 29*(4), 210-225.

Mahadevan, S., & Halankar, P. (2021). A comparative study of pre-competitive anxiety level between male and female intercollegiate baseball players of Goa. *International Journal of Research Pedagogy and Technology in Education and Movement Sciences*, *10*(02), 1-7.

Marih, H., Nazarudin, M. N., & Noordin, Z. (2023). The motivation influence on level of physical activity among UKM students bachelor degree in sports and recreation education. *Jurnal Personalia Pelajar, 26*(1), 139-150. https://doi.org/10.17576/personalia.2601.2023.02

Martens, R., Vealey, R. S., & Burton, D. (1990). *Competitive anxiety in sport.* Human Kinetics.

Martinent, G., Ferrand, C., Guillet, E., & Gautheur, S. (2010). Validation of the French version of the Competitive State Anxiety Inventory-2 Revised (CSAI-2R) including frequency and direction scales. *Psychology of Sport and Exercise, 11*(1), 51-57. https://doi.org/10.1016/j.psychsport.2009.05.001

Martinez, L., & Scott, D. (2021). The relationship between cognitive and somatic anxiety and athletic performance: A review. *Journal of Sports Science and Medicine, 20*(3), 567-582. https://doi.org/10.1056/jssm.2021.20.3.567

Mavvidis, A., Stamboulis, A., Dimitriou, V., & Giampanidoy, A. (2010). Differences in forehand and backhand performance in young tennis players. *Studies in Physical Culture and Tourism, 17*(4), 315-319.

Misran, R. N., Khaiyom, J. H. A., & Razali, Z. A. (2021). The role of religiosity to address the mental health crisis of students: A study on three parameters (anxiety, depression, and stress). *Pertanika Journal of Social Sciences and Humanities, 29*(4), 2833-2851. https://doi.org/10.47836/pjssh.29.4.40

Mohebi, M., Zandi, A. G., Zarei, S., & Zandi, H. G. (2019). Self-compassion and pre-competition anxiety in martial arts student players. *Sport Sciences and Health Research, 11*(2), 97-104. https://doi.org/10.32598/JESM.11.2.4

Mojtahedi, D., Dagnall, N., Denovan, A., Clough, P., Dewhurst, S., Hillier, M., Papageorgiou, K., & Perry, J. (2023). Competition anxiety in combat sports and the importance of mental toughness. *Behavioral Sciences*, *13*(9), Article 713. https://doi.org/10.3390/bs13090713

Morales, P., & Fernandez, I. (2023). Effectiveness of anxiety management programs on tennis performance: A randomized controlled trial. *Journal of Sports Sciences, 41*(6), 678-689.

Muhamad, T. A., Rashid, A. A., Razak, M. R. A., & Salamuddin, N. (2011). A comparative study of backhand strokes in tennis among national tennis players in Malaysia. *Procedia Social and Behavioral Sciences, 15*, 3495-3499. https://doi.org/10.1016/j.sbspro.2011.04.324

Mullen, R., & Hardy, L. (2000). State anxiety and motor performance: Testing the conscious processing hypothesis. *Journal of Sports Sciences, 18*(10), 785-799. https://doi.org/10.1080/026404100419847

Muñoz, A. S., Cayetano, A. R., Calle, R. C., Blanco, J. M. F., Ramos, J. M. D. M., Vicente, R. A., & Muñoz, S. P. (2017). Female Spanish athletes face pre-competition anxiety at the highest levels of competition. *Revista de Psicología Del Deporte, 26*(4), 39-44.

Nguyen, Q., & Tran, P. (2023). The impact of cross-sectional study designs in sports psychology research: A critical analysis. *Methodological Advances in Sports Psychology, 12*(1), 22-35.

Pa, W. A. M. W., Zin, M., Salamuddin, N., & Lian, D. K. C. (2021). The effect of sports massage towards cortisol and pre-competition anxiety among Malaysian elite tennis athlete. *Journal of Contemporary Issues in Business and Government, 27*(2), 1518-1528.

Patel, R., & Smith, J. (2024). Personalizing anxiety management for athletes: Emerging trends and strategies. *Contemporary Psychology, 45*(1), 98-112.

Radzi, J. A., Yusuf, S. M., Amir, N. H., & Mansor, S. H. (2018). Relationship of pre-competition anxiety and cortisol response in individual and team sport athletes. In R. Saian & M. Abbas (Eds.), *Proceedings of the Second International Conference on the Future of ASEAN (ICoFA) 2017, 2: Science and Technology* (pp. 719-727). Springer Singapore. https://doi.org/10.1007/978-981-10-8471-3_70

Roberts, T., & Jackson, L. (2022). Expanding cultural diversity in sports: Implications for team performance and athlete development. *Journal of Sport Psychology, 34*(2), 112-128.

Roetert, E. P., & Kovacs, M. (2019). *Tennis anatomy*. Human Kinetics.

Saadan, R., Hooi, L. B., & Ali, H. M. (2015). Hubungan antara kebimbangan pra pertandingan dan orientasi matlamat dalam kalangan atlet hoki [The relationship between competitive anxiety and goal orientation among hockey athletes]. *Journal of Human Capital Development*, 8(2), 37-48.

Saadan, R., Hooi, L. B., Ali, H. M., Bokhari, M., & Abdullah, N. (2016). Perbandingan tahap kebimbangan pra-pertandingan dalam kalangan atlet hoki Majlis Sukan Sekolah Malaysia [Comparison of pre-competition anxiety levels among Malaysian School Sports Council hockey athletes]. *Jurnal ILMI*, 1(1), 43-54.

Şekeroğlu, M. Ö. (2019). An examination of the pre-competition and post-competition state anxiety levels of Russian women's national volleyball team athletes. *Journal of Education and Training Studies*, 7(5), 19-26. https://doi.org/10.11114/jets.v7i5.4185

Singh, V., & Gupta, R. (2023). Pre-competition anxiety and its impact on sports performance: A review of mediating factors. *Journal of Sports Psychology and Performance, 30*(3), 275-288.

Souza, R. A., Beltran, O. A., Zapata, D. M., Silva, E., Freitas, W. Z., Junior, R. V, da Silva, F. F., & Higino, W. P. (2019). Heart rate variability, salivary cortisol and competitive state anxiety responses during pre-competition and pre-training moments. *Biology of Sport, 36*(1), 39-46. https://doi.org/10.5114/biolsport.2018.78905

Spielberger, C. D. (2021). Stress and anxiety in sports. In D. Hackfort & C. D. Spielberger (Eds.), *Anxiety in sports* (pp. 3-17). Taylor & Francis. https://doi.org/10.4324/9781315781594-2

Sutarsyah, C. (2017). An analysis of student's speaking anxiety and its effect on speaking performance. *Indonesian Journal of English Language Teaching and Applied Linguistics, 1*(2), 143-152. https://doi.org/10.21093/ijeltal.v1i2.14

Turksoy, A., Bayansalduz, M., Altinci, E. E., & Atikir, A. (2012). Comparison of pré and post-anxiety levels of volleyball players aged 11-14. *International Journal of Academic Research, 4*(1), 143-147.

Uphill, M. A., Rossato, C. J., Swain, J., & O'Driscoll, J. (2019). Challenge and threat: A critical review of the literature and an alternative conceptualization. *Frontiers in Psychology, 10,* Article 1255. https://doi.org/10.3389/fpsyg.2019.01255

Vergauwen, L., Madou, B., & Behets, D. (2004). Authentic evaluation of forehand groundstrokes in young low- to intermediate-level tennis players. *Medicine and Science in Sports and Exercise, 36*(12), 2099-2106. https://doi.org/10.1249/01.MSS.0000147583.13209.61

Weinberg, R. S., & Gould, D. (2015). *Foundations of sport and exercise psychology* (6th ed.). Human Kinetics.

Woodman, T., & Hardy, L. (2003). The relative impact of cognitive anxiety and self-confidence upon sport performance: A meta-analysis. *Journal of Sports Sciences, 21*(6), 443-457. https://doi.org/10.1080/0264041031000101809

Yang, H., Wen, X., & Xu, F. (2020). The influence of positive emotion and sports hope on pre-competition state anxiety in martial arts players. *Frontiers in Psychology, 11,* Article 1460. https://doi.org/10.3389/fpsyg.2020.01460

Zhang, Y., & Wang, X. (2024). Longitudinal designs to explore causal relationships between pre-competition anxiety and tennis performance. *Journal of Sports Science and Medicine, 32*(2), 123-135.

Visualization Analysis of Research on Chinese Cultural Identity Based on CiteSpace (2013–2022)

Chunyuan Niu[1,2], Chwee Fang Ng[1]*, Hazlina Abdul Halim[1] and Nik Farhan Mustapha[1]

[1]*Department of Foreign Languages, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 Serdang, Malaysia*
[2]*Faculty of Foreign Languages, Suqian University, 223800 Sucheng District, Suqian City, People's Republic of China*

ABSTRACT

Though extensively examined, research on Chinese cultural identity lacks systematic categorization, leading to uncertainties about evolving research priorities. This study addresses this gap by comprehensively analyzing Chinese Cultural Identity (CCI) development trends using the Web of Science database and CiteSpace, a scientific literature visualization and analysis tool in bibliometrics and scientometrics. This study examines Chinese cultural identity development from 2013 to 2022, encompassing research trends, global distribution of research institutions, and emergent hotspots. It uncovers a two-phase progression in Chinese cultural identity research: initial gradual exploration followed by rapid advancement. While research in this field is concentrated in China, the United States, Australia, the United Kingdom, Canada, Singapore, and Malaysia, it is noteworthy that research institutions are globally distributed. Three focal points (Chinese culture examination, cultural identity investigation, and cultural identity determinants) highlight the subject's multifaceted nature. Additionally, emerging research frontiers include the empowerment of women in Chinese cultural identity, the role of information in shaping Chinese cultural identity, and the globalization of Chinese cultural identity, pushing the boundaries of understanding in this field. Further studies are encouraged to explore dissemination dynamics, adopt interdisciplinary approaches, and conduct in-depth case studies.

Keywords: Chinese cultural identity, Citespace, research hotspots, visualization analysis

INTRODUCTION

Cultural identity refers to the true consciousness of a collective that shares a common history and ancestry, which is the deepest level of identity and the basis of cohesion of a cultural community (Hall, 2015). The study of cultural identity holds immense significance in various academic disciplines, including anthropology, sociology, psychology, cultural studies, and communication (Dervin, 2012). Understanding cultural identity allows researchers and scholars to explore the complex interplay between individual experiences and broader societal contexts. Cultural identity shapes social bonds and community while triggering conflicts (Sung, 2022). Multiculturalism, globalization, and migration impact identity negotiation. It influences personal identity formation and is linked to colonialism's legacy (Sengupta, 2023). Migration prompts identity adaptation, which is relevant for integration and resilience.

While existing studies have discussed Chinese cultural identity, there is a lack of research that utilizes visual analysis tools like CiteSpace to map out and explore the interconnectedness, trends, and evolution of key themes within this complex field (Q. Wang, 2022). Few studies have delved into the visual analysis of Chinese cultural identity literature across a specific temporal scope, such as the past decade. This gap hinders a comprehensive understanding of how research hotspots and trends have evolved over this period. Moreover, some studies may focus on established themes within Chinese cultural identity. However, there is a potential gap in recognizing and analyzing emerging and emerging themes, perspectives, or controversies within the last decade.

There is a call to investigate longitudinal patterns across decades, uncovering cyclical themes within Chinese cultural identity discourse to enhance understanding. Comparative analyses with similar studies in diverse cultural contexts can illuminate cross-cultural similarities, differences, and broader global trends. Several key research questions emerge in addressing these gaps and pursuing a comprehensive analysis of Chinese cultural identity literature through visual mapping using CiteSpace. These questions delve into the multifaceted dimensions of CCI research, shedding light on its development, manifestation across borders, geographical concentration, and topical focal points. The following research questions guide our investigation:

1. What are the prominent trends and patterns in the research of CCI's development?

2. How is Chinese cultural identity manifested and distributed across various countries?

3. What is the geographical distribution and institutional concentration of research organizations studying Chinese cultural identity?

4. What are the current research hotspots within the Chinese Cultural Identity (CCI) field?

5. How can the research frontier and distribution of topics within the Chinese Cultural Identity (CCI) field be effectively identified and mapped?

The significance and potential contributions of this research are multi-faceted. Firstly, it aims to advance theoretical understanding by delving into the intricate interplay of cultural identity, nationhood, and unity within Chinese Cultural Identity (CCI; Ho, 2020). This exploration is particularly important as cultural identity forms the bedrock of individual self-perception and societal cohesion (Winden, 2023). Secondly, the study addresses research gaps in CCI literature, particularly in a comprehensive analysis of research hotspots, trends, and distribution. By bridging these gaps, the research provides a panoramic view of the current state of CCI research, enabling a more holistic understanding of the field's dynamics. Thirdly, the research seeks to uncover emerging trends within CCI by meticulously analyzing its trajectory over the past decade. This endeavor is poised to shed light on nascent perspectives, pertinent concerns, and evolving thematic foci (Žukauskas et al., 2018). Additionally, the study employs visualization tools like CiteSpace software to map the intricate tapestry of CCI research, offering insights that surpass traditional textual summaries (Song & Wang, 2020). This approach empowers researchers and policymakers with a contextual grasp of the research landscape, guiding future research directions effectively (Dacholfany et al., 2023). Lastly,

by comprehensively analyzing research hotspots, trends, and distribution, the study offers valuable guidance for future research trajectories within the expansive realm of CCI. Overall, this research not only contributes to academic discourse but also informs policy decisions and socio-cultural development by harnessing the potent role of CCI in shaping national unity, cultural preservation, and societal cohesion.

METHODOLOGY

Data Source

Web of Science is a large, comprehensive, and multidisciplinary database developed by the Institute for Scientific Information (ISI) based on the web. Compared with other databases, the WoS Core Collection database collects the references cited by articles and organizes a characteristic citation index by cited author, source and publication year (Birkle et al., 2020). The review focused on the previous studies of Chinese cultural identity in the WoS Core Collection database from 2013 to 2022. The following search strategy was conducted: Topic = "Chinese Cultural Identity". The search strategy showed 1766 publications after the exclusion of irrelevant academic articles.

Methods

The bibliometric analysis method is a research approach that primarily relies on the information contained within paper catalogs (Khudhair et al., 2021). By applying relevant mathematical and statistical knowledge, this

method enables researchers to analyze the characteristics of the number of papers published in a given field (Antons et al., 2020). In essence, bibliometric analysis provides a more objective and rigorous means of assessing the state of a particular field of research (Donthu et al., 2021).

This analysis project was established in CiteSpace 6.2.R4, updated on August 4, 2023, an information visualization software developed by Professor Chen Chaomei from the School of Information Science and Technology at Drexel University. CiteSpace is used widely in scientific knowledge mapping and literature scientometrics, which can analyze and forecast the hot research fields, research process, research evolution, and frontier of the subject (C. Chen, 2005). CiteSpace offers visual analysis and knowledge mapping capabilities, aiding researchers in understanding the structure, relationships, and emerging trends within their fields. It systematically generates accessible graphs, allowing scholars to identify hot spots and explore relationships in their research domains. It is suitable for dynamic complex network analysis, provides diversified display modes such as various cluster graphs, timeline graphs and time zone graphs, and generates various knowledge graphs (C. Chen, 2005). In the context of CiteSpace, nodes represent diverse entities, encompassing academic papers, authors, or keywords. These nodes serve as the foundation of the network, and their connections delineate relationships such as co-authorship, citation, or shared keywords.

Applying cluster graphs within CiteSpace illuminates groups or clusters of nodes characterized by shared characteristics, unveiling the structural intricacies intrinsic to academic disciplines or thematic areas. For example, a cluster may signify a distinct research topic, with each node representing a paper contributing to that area. Complementing this, timeline graphs, integral to the tool, provide a temporal dimension that facilitates the depiction of the evolution of research fields over time. This feature allows for a nuanced understanding of specific topics' emergence, growth, and decline. A timeline graph can, for instance, illustrate the fluctuations in popularity of a particular research topic over time.

Additionally, time zone graphs in CiteSpace contribute to temporal analysis by identifying bursts or periods of heightened activity within a research field. These bursts shed light on critical periods marked by significant advancements or shifts in focus. For instance, a burst might indicate intensified research following a major discovery. These functionalities empower researchers to conduct meticulous analyses of individual nodes, explore clusters of related research, and glean insights into the dynamic evolution of academic knowledge over time. Researchers can leverage CiteSpace to identify influential works, track the development of research topics, and predict future trends within a field.

This study classified and analyzed 1766 articles from the perspectives of publication year, country and region, research institution,

topic category and keywords to dig into the history, current situation, research hotspots and research contents of the research on

Chinese cultural identity. Table 1 describes the data collection and analysis protocol of this study.

Table 1
Data collection and analysis protocol

Research protocol	Retrieve results and contents
Research database	WoS Core Collection Database
Language	English
Publication type	All types
Year span	January 2013-December 2022
Retrieval criterion	Topic = Chinese Cultural Identity
Exclusion criteria	Early access, proceeding paper, book chapters
Number of samples	1766
Data analysis	CiteSpace, 6.2. R4
Analysis paths	Co-network analysis, co-occurrence analysis, a noun-term burst analysis

Source: Authors' work

RESULTS AND DISCUSSION

General Trends of CCI Research

The number of published papers is a crucial indicator for measuring the development trend of Chinese cultural identity within a specific time range. By analyzing the change in research popularity over a given period, it is possible to gain valuable insights into the dynamics of development and to make predictions about future trends. Figure 1 depicts a recent inquiry into the research outcomes of the digital economy spanning the last decade, presenting a statistical chart delineating the distribution and proportion of annually published papers. The number of published papers on Chinese cultural identity exhibits an upward trend. This trend can be divided into two distinct stages: 2013 to 2015 and 2016 to 2022. During the first stage, from 2013 to 2015,

research on Chinese cultural identity was in a slow growth phase, with fewer than 100 publications per year. However, between 2016 and 2022, there was a rapid increase in the number of articles published, rising from 109 in 2016 to a peak of 331 in 2022. It represents the highest value recorded over the past decade. The significant and relatively stable increase in published papers indicates that research on Chinese cultural identity has garnered attention from the academic community and that there is considerable scope for further exploration in this field.

Slow and Diversified Development (2013-2015)

From 2013 to 2015, research on Chinese cultural identity (CCI) exhibited a slow upward trend, with 249 published, accounting
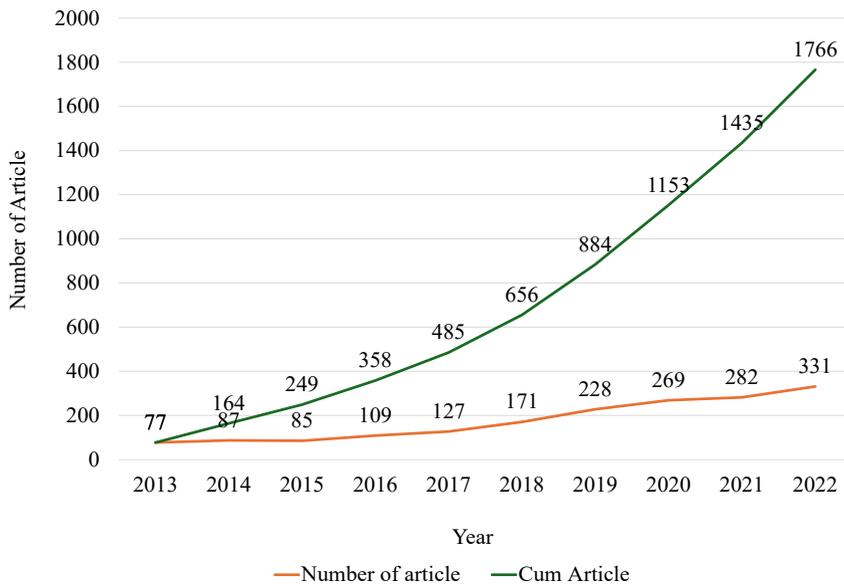
Figure 1. Map of annual publication of CCI research (2013–2022)

Source: Authors' work

for only 14.1% of the total papers. It can be attributed to the lack of a unified and clear understanding of the characteristics and importance of Chinese cultural identity. Despite this, CCI received attention from various fields during this period, resulting in diverse research directions. For instance, studies were conducted on topics such as biculturalism in cultural diversity (Shen & Liao, 2022), acculturation and attitude (Demes & Geeraert, 2013), the influence of immigration on cultural identity (Wei, 2014), self-awareness (Saad et al., 2013), mental health (Wyatt et al., 2015), and ethnic identity (D. Wang, 2016). Among these studies, speculative research was the predominant approach, with practical research being relatively rare (Telzer et al., 2014).

Fast and In-Depth Development (2016-2022)

Since 2016, when the number of published papers on Chinese cultural identity research exceeded 100, there has been a steady increase in the number of papers published on this topic. The average annual growth has been around 50, indicating a significant and steady increase in published papers. By December 31, 2022, there were 1517 papers, accounting for 85.9% of the total. Additionally, the scope of research has expanded compared to the period from 2013 to 2015. It suggests that Chinese cultural identity research has become an increasingly popular and important study area. The main research topics in this stage include cultural heritage, cultural integration, cultural products and their consequences. As seen

from the above keywords, the research on CCI begins to shift from abstract concepts and classifications to concrete models and practices. Cultural heritage is an important part of Chinese cultural identity (Lee & Wong, 2017), including philosophy, literature, history, calligraphy, fine arts, music, dance, martial arts, technology, and architecture. With economic globalization, cultural integration has become a hot spot in state, enterprise and education (L. A. Liu et al., 2018). With the development of tourism, cultural products as a culture carrier have also attracted scholars' attention (T. Li et al., 2020). In the research process, scholars also analyze the consequences of Chinese cultural identity (Wells et al., 2022). With the deepening of the research on Chinese cultural identity, the research results in this stage increased rapidly in quantity and improved in quality.

Productive Countries

The analysis also unveils the countries or regions that exhibit a significant concentration of research focused on Chinese cultural identity and their macro, meso, and micro levels of collaborative relationships. The classification of nodes is based on country, with a predefined threshold value of 45. Figure 2 illustrates a network representation of countries by employing CiteSpace for visualization analysis. According to the outcomes of quantitative analysis conducted via CiteSpace, 79 countries have engaged in research concerning Chinese cultural identity. Notable participants include

China, the United States, Australia, the United Kingdom, Canada, Singapore, and Malaysia. These seven countries collectively account for 1,526 published articles on Chinese cultural identity. In Figure 2, each circular node corresponds to a country; its size is proportional to the frequency of published articles on Chinese cultural identity. The node's color corresponds to the time of article publication, while the linkage strength between nodes reflects the intensity of collaborative relationships. Closer proximity of links denotes stronger cooperation.

From 2013 to 2022, China produced 636 research papers, accounting for 41.7% of the total scholarly output, while the United States contributed 423 papers, representing 27.7% of the corpus. Figure 2 illustrates that both countries have relatively autonomous research affiliations with some interconnections. However, the United States has a more extensive and interconnected network of collaborations with various countries worldwide, indicating a higher degree of scholarly interconnectedness. In contrast, China maintains more independent research avenues with a comparatively smaller collaborative network. Meanwhile, examining the publication landscape on Chinese cultural identity from 2013 to 2022 reveals notable disparities among Hong Kong, Taiwan, and Macao, reflecting diverse research environments. With its international prominence, Hong Kong exhibited a substantial output of 200 publications, indicative of a robust and prolific research landscape.
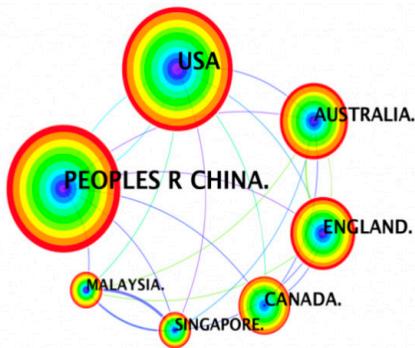
Size of the Circle: Numerical Value
Thickness of the Line: Relation Closeness
Figure 2. International cooperative network atlas
Source: Authors' work

This higher count is likely influenced by extensive collaboration networks and a cosmopolitan research environment, fostering domestic and international partnerships. While demonstrating a moderate publication count of 69, Taiwan may reflect a nuanced mix of regional and international collaborations influenced by its distinct geopolitical context. In contrast, Macao's lower count of 10 publications suggests a more localized and potentially specialized approach to Chinese cultural identity research. Additionally, factors such as language preferences, political contexts, and research resources may further elucidate the reasons for the divergent publication counts in these regions. Notably, Malaysia and Singapore demonstrate significant research synergy, highlighting the power of research to foster collaborative bonds even within a regional context, transcending geopolitical boundaries for shared academic exploration and Southeast Asia is defined by cultural diversity and openness (King, 2016).

Research Organizations

Examining institutional collaboration brings pertinent organizational insights within Chinese cultural identity research. Specific parameters were employed to facilitate the analysis and categorization of research institutions. The time span was defined as 2013–2022, while each publication year was assigned a weight of 1. The top N criterion was set at 5 for selection. Figure 3 shows a network visualization of the literature culled from the Web of Science (WoS) database. This visual representation encompasses 318 nodes and 206 connecting lines, yielding a network density of 0.0041. This density value indicates a comparatively low degree of collaboration among institutions.

Upon more nuanced scrutiny of Figure 3, an intricate tapestry of research endeavors unfolds, underscoring the scholarly prominence of institutions within Hong Kong, China, in exploring the contours of Chinese cultural identity. The Chinese University of Hong Kong, The Hong Kong Polytechnic University, and the City University of Hong Kong emerge as key players, leaving an indelible imprint on the scholarly landscape with their respective contributions of 88, 32, and 28 publications. This robust participation points to the academic vitality of Hong Kong's research milieu in delving into the nuanced facets of Chinese cultural identity. The substantial contributions from The Chinese University of Hong Kong, The Hong Kong Polytechnic University, and the City University of Hong Kong underscore Hong Kong's academic prominence in Chinese cultural identity

research. These institutions, particularly The Chinese University of Hong Kong, renowned globally for research excellence, attract scholars making significant contributions to the discourse.

Similarly, The Hong Kong Polytechnic University and the City University of Hong Kong are esteemed institutions that draw researchers across diverse academic disciplines, enriching the overall research landscape. Hong Kong's unique geopolitical position as a Special Administrative Region (SAR), situated at the crossroads of East and West, provides a conducive environment for exploring multifaceted dimensions of Chinese cultural identity. This context is vital for these institutions, allowing nuanced investigations. Collaboration and interdisciplinary approaches within these institutions significantly contribute to robust research output, fostering comprehensive examinations of Chinese cultural identity from various perspectives. Moving forward, fostering further interdisciplinary collaborations to explore intersections between Chinese cultural identity and contemporary issues is imperative. Additionally, investigating how Chinese cultural identity manifests in Hong Kong's local context, considering its unique history and sociopolitical dynamics, and exploring the impact of global trends on its evolution are crucial areas for future research. The academic vitality of these institutions, coupled with their interdisciplinary strengths, positions researchers to address emerging themes and contribute to ongoing scholarly discourse on Chinese cultural identity in Hong Kong.

Size of the Circle: Numerical Value   Thickness of the Line: Relation Closeness
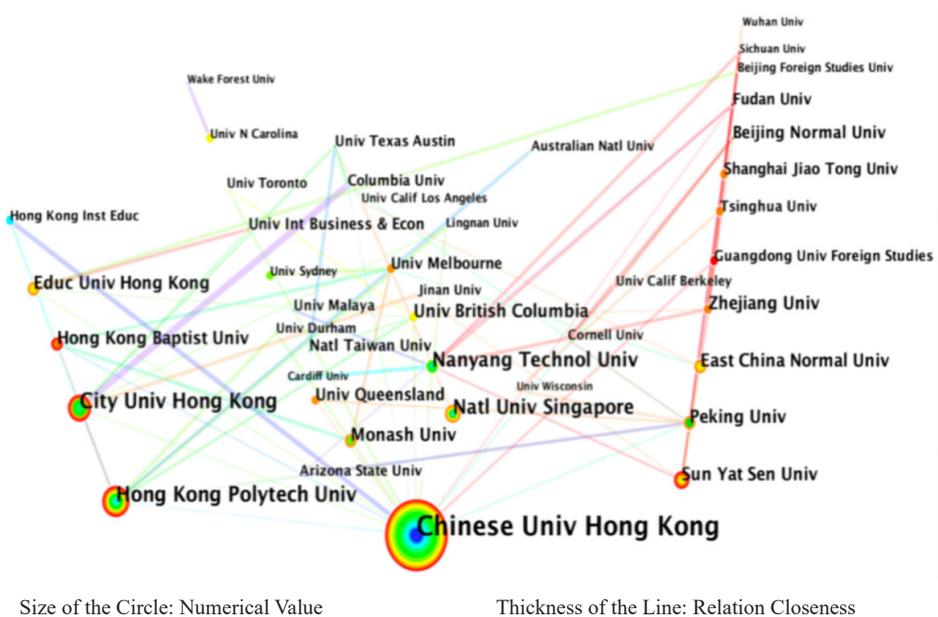
Figure 3. WOS Chinese cultural identity research institution network map

Source: Authors' work

Beyond China, global research institutions, like the National University of Singapore and Nanyang Technological University, are actively engaged in studying Chinese cultural identity (CCI), each contributing 24 and 26 research papers, respectively. This global scholarly interest highlights CCI's transnational appeal, drawing diverse scholars into a shared discourse. Similarly, Monash University and the University of Queensland have produced 17 and 13 research papers on CCI in Australia, while the University of British Columbia has contributed 16. These institutions collectively showcase the global significance of CCI research, fostering international collaboration and discourse. This interplay between institutions, both inside and outside China, reflects CCI's universal importance and the commitment of the academic community to comprehensively explore its multifaceted dimensions.

Research Hotspots

Using CiteSpace's analytical capabilities, this study delves into a detailed examination, producing a clear visualization of keyword connections. Figure 4 showcases this intricate web of scholarly associations, providing valuable insights into the core themes in Chinese cultural identity research. In this carefully crafted network diagram, node sizes correspond to keyword frequency, visually representing the importance of different concepts. The connections between nodes reveal relationships, with denser networks indicating stronger thematic links. This analytical method offers a deep understanding of the research landscape, uncovering key focal points and revealing the intricate interplay of ideas. The result is a comprehensive overview that captures the main themes and identifies the central lines of inquiry in Chinese cultural identity research. This visual representation is a crucial tool, shedding light on the various dimensions, underlying patterns, and significant research directions within this complex field.

Keyword frequency and centrality are quantitative measures used to gauge the significance of specific terms within a given research domain. Nodes with elevated frequency and centrality values indicate prevalent research focal points during a particular time frame. Table 2 compiles the top 20 keywords with the highest frequency using cluster analysis. This assortment provides a comprehensive overview of the predominant thematic elements that have garnered substantial attention within the research context.

Based on the analysis of the CCI keyword co-occurrence network diagram presented in Figure 4 and the list of keywords with high frequency and centrality provided in Table 2, the salient trends in Chinese cultural identity research can be categorized into the following three thematic aspects:

Chinese Culture Examination

This research category encompasses a broad and in-depth examination of Chinese culture, delving into its tangible and intangible aspects. It includes exploring China's

material culture, such as its art, architecture, and artifacts, as well as its spiritual facets, such as its beliefs, values, and customs. This scholarly exploration of Chinese culture traverses various studies, beginning with examining the social contract among young Chinese citizens and the ensuing study pressures. Zhao (2020) widens the scope by applying queer China studies and global media studies to analyze a show's portrayal of linguistic, geocultural, gender, and sexual variations in the context of China's media and cultural globalization. Kathina et al. (2020) shift the focus to Kenya, emphasizing the complexities of integrating Chinese Language and Culture into second language education.

A. Chen (2021) addresses localization concerns in Montessori education, while Lou (2021) extends the discourse to postcolonial contexts, highlighting the role of national identity and language proficiency in cultural
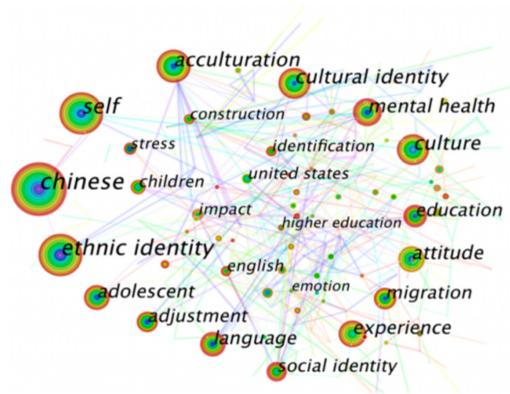
Size of the Circle: Numerical Value

Thickness of the Line: Relation Closeness

Figure 4. CCI keywords co-occurrence network graph
Source: Authors' work

adaptation. Dai (2021) takes an economic perspective, investigating local cultural capital as a motivator for entrepreneurial behavior in China. In the regional context, M. Liu and Zhong (2020) discuss British influence in shaping Hong Kong's identity. Karim et al. (2021) report on acculturative

Table 2
Keywords list of CCI

No.	Keywords	Frequency	Centrality	No.	Keywords	Frequency	Centrality
1	Chinese	145	0.04	11	adolescent	49	0.04
2	ethnic identity	104	0.07	12	education	45	0.03
3	self	91	0.06	13	migration	44	0.05
4	acculturation	68	0.05	14	adjustment	43	0.03
5	culture	65	0.01	15	social identity	40	0.14
6	culture identity	64	0.03	16	children	33	0.09
7	attitude	58	0.04	17	impact	29	0.10
8	mental health	55	0.11	18	English	29	0.07
9	experience	53	0.03	19	identification	27	0.08
10	language	51	0.09	20	construction	24	0.07

Source: Authors' work

challenges faced by Pakistani students in Hong Kong. Tang (2021) explores teaching traditional Chinese science in Hong Kong universities and delves into the urban experiences of mainland Chinese youth in Hong Kong. Concluding the examination, Richards and King (2022) employ the Event Experience Scale to gauge tourist experiences at cultural festivals in Hong Kong, emphasizing their role in protecting the city's identity as a hub of Chinese and traditional culture. Researchers in this field employ a cross-disciplinary approach, drawing on the areas of China Studies, Communication, and Cultural Studies to analyze, communicate, perceive and represent Chinese culture both within China and in the world.

Cultural Identity Investigation

This area of research involves in-depth inquiries into cultural identity, encompassing discussions related to national cultural identity and cultural attitudes and the impact on people's identity or their self-conception and self-perception. It seeks to comprehend how individuals perceive and associate themselves with their cultural backgrounds. This area of research aims to understand how these factors interact and influence an individual's perception of their cultural background. This comprehensive exploration of cultural identity within China examined through a series of scholarly works, delves into nuanced perspectives and diverse geographical contexts. A. Cheng and Szeto (2019) scrutinize the effects of short-term study abroad on Hong

Kong university students' national identity, revealing multifaceted perceptions shaped by the region's unique geopolitical status. Wu et al. (2020) extend the exploration, investigating how social capital influences cultural identities among Chinese youths.

V. Cheng (2020) critically evaluates Hong Kong's revamped museum, emphasizing its reliance on obsolete narratives and the evolving dynamics of politics and art. Lam and Ryan (2020) propose that sociological factors in Macau outweigh micro-level variables in shaping perceptions of place. Ng (2021) explores identity incompatibility among Hong Kong citizens, Chung (2021) investigates translation issues, and Lowe (2021) explores the role of the cultural commons in Hong Kong's pro-democracy movement. W. Lin and Yang (2019) focus on the cultural lens to examine the development of higher education in Taiwan. McIntyre (2019) explores the role of visual arts in relations between Taiwan and mainland China. Loewenberg (2020) highlights traditional imperatives that contribute to the responsiveness to psychoanalysis in China. The collective research offers a concise yet comprehensive understanding of Chinese cultural identity.

Cultural Identity Determinant

Within the realm of Chinese cultural identity, a range of determinants influence how individuals perceive and associate themselves with their cultural backgrounds. These factors include self-awareness, educational experiences, personal encounters, migration, and the acquisition

of English language skills and shape how individuals perceive and associate themselves with their cultural background. Especially in examining cultural identity determinants within China, Chan (2020) strategically leverages Wing Sang Law's conceptualization of colonial power as a network of relations to analyze the dynamic shifts in power relations between Hongkongers and mainlanders. This analytical framework proves particularly pertinent in elucidating the impact of China's rapid economic growth and cultural influence on the intricate fabric of identity dynamics. Chan's inquiry prompts essential questions concerning the nuanced differentiation of "Chinese" from "China" in contemporary Hong Kong, adding depth to our understanding amid the ongoing evolution of coloniality.

Transitioning seamlessly from this exploration of power dynamics, Luo and Xiao (2022) pivot the focus toward the concept of "cultural governance" (p. 418). Grounded in a meticulous tracing of its evolution within the broader context of China's agenda for the modernization of state governing capacity, the authors meticulously unravel the term-shaping processes within the political discourse that birthed these governance concepts. Drawing from case studies in key mainland Chinese urban centers—Shenzhen, Foshan, and Dongguan in the Guangdong-Hong Kong-Macau Greater Bay Area—This research methodically explores urban planning, infrastructure development, community reformation, and the adoption of a cultural-sustainable development strategy by Chinese authorities. This scholarly exploration underscores the adaptability and flexibility of the state's involvement. It offers nuanced insights into models and challenges associated with pioneering cultural governance programs, thus contributing to a comprehensive comprehension of what may be perceived as a consequential "cultural turn" in urban governance. In tandem, these studies collectively illuminate the intricate intersections of cultural identity, power relations, and governance strategies within the dynamic context of China.

Research Frontier Identification

The mutation detection algorithm in word analysis systematically identifies significant but infrequently occurring terms in literature. Frequent mutations in keywords often indicate future topic evolution. Chen Chaomei's team integrated Kleinberg's burst detection into CiteSpace, creating a mutation word detection function. Figure 5 shows the intensity and trend of mutant words, offering insights into future trends in Chinese cultural identity research. It calculates and extracts focal words with high mutation frequency, revealing the cutting-edge landscape of the discipline. Our study used CiteSpace to analyze keyword mutations in exploring Chinese cultural identity from 2013 to 2022, resulting in 25 keywords with mutations.

Empowerment of Women in CCI

The keyword "women" displayed a significant intensity with a strength value

Top 25 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2013 - 2022
social support	2013	2.83	2013	2017	
threat	2013	2.69	2013	2014	
personality	2013	2.46	2013	2015	
space	2013	2.89	2014	2016	
ethnicity	2013	2.62	2014	2015	
geography	2013	2.62	2014	2015	
chinese american	2013	2.9	2015	2016	
soft power	2013	3.34	2016	2017	
history	2013	2.86	2016	2017	
conflict	2013	2.86	2016	2017	
bicultural identity integration	2013	2.43	2016	2017	
context	2013	3.06	2017	2019	
emotion	2013	2.67	2017	2019	
information	2013	4.46	2018	2019	
attitude	2013	3.46	2018	2019	
collectivism	2013	3.07	2018	2019	
african american	2013	2.67	2018	2019	
consumption	2013	2.45	2018	2020	
women	2013	5.89	2019	2020	
learner	2013	2.88	2019	2020	
community	2013	2.67	2019	2020	
strategy	2013	2.67	2019	2022	
impact	2013	4.07	2020	2022	
place	2013	3.19	2020	2022	
socialization	2013	2.48	2020	2022	

Figure 5. Top 25 keywords with the strongest citation bursts
Source: Authors' work

of 5.89, ranking first in Table 2. Despite not extending into 2022, "women" exhibited the highest intensity. Scholars often integrate this topic with feminism, suggesting that research on women's Chinese cultural identity remains a prominent future trend. It indicates an increased focus on understanding the role of women in the context of Chinese cultural identity. Some studies posit that Chinese women's self-empowerment via social media stems from a gentle, rational, and resolute approach that integrates a new female identity into the Confucian harmonious society, resulting in a unique digital feminism with Chinese characteristics (Chang et al., 2016). Others advocate for an institutional-level approach to empowering Chinese women and seek to raise awareness of similar strategies for promoting women's empowerment in non-western contexts (Z. Lin & Yang, 2019).

Role of Information in Chinese Cultural Identity

As indicated in Table 2, the keyword "information" associated with Chinese cultural identity holds substantial strength with a value of 4.46, ranking second among the 25 keywords. Recognizing that information is a vital carrier of culture; this trend aligns with the rapid advancement of information technology and China's ongoing support for its cultural heritage. This proliferation of channels for disseminating Chinese culture results in an ever-widening scope of influence. Some scholars point out that the advancement of digital technology has facilitated the creation of databases for intangible Chinese cultural heritage (Fan & Wang, 2022). Some studies examine the efficacy of an innovative teaching model for massive open online courses to evaluate its potential impact on Chinese folk culture and identity within higher education (R. Li, 2022).

Globalization of Chinese Cultural Identity

As the world becomes increasingly interconnected, the dissemination of Chinese culture extends beyond national boundaries. Moreover, the cultural diversity brought about by economic globalization is progressively reshaping the fabric of global civilization. Given the present era and the ongoing shifts in society, the importance of preserving ethnic cultures gains heightened prominence amid the

context of globalization. Some studies leverage the traits of educational evolution toward diversity to scrutinize the correlation between education and preserving ethnic culture (Zhang, 2019). Nevertheless, this transformation gives rise to novel challenges concerning the interplay of conflict, integration, and coexistence among international students from diverse cultural backgrounds (Hu & Dai, 2021). Research in this direction may explore how Chinese cultural identity adapts and interacts with global cultural trends, contributing to the increasing influence of Chinese culture on the international stage.

The discerned trends serve as crucial indicators, shedding light on domains of persistent significance and nascent intrigue in Chinese cultural identity research. These identified patterns underscore the ongoing pertinence of certain themes and illuminate burgeoning areas of interest that are gaining prominence within academic discourse. These insights assume a paramount role for scholars and researchers committed to probing the intricate nuances of Chinese cultural identity. They furnish a compass to navigate the complex landscape of evolving dynamics. They offer valuable guidance for those poised to unravel the intricate threads of Chinese cultural identity's evolution in the forthcoming years.

CONCLUSION

The study has illuminated two distinct phases characterizing Chinese cultural identity research: an initial phase (2013–2015) marked by diverse exploration and a subsequent phase (2016–2022) distinguished by rapid progress. The temporal progression delineated in this study signifies a dynamic evolution within Chinese cultural identity research. This evolution prompts a critical reevaluation of existing theoretical frameworks, necessitating a deeper understanding of the developmental trajectory of cultural identity over time. The identified key research areas, notably Chinese culture examination, cultural identity investigation, and determinants of identity, serve as established focal points that have significantly contributed to shaping the scholarly discourse. However, the study also illuminates notable research gaps and introduces emerging themes that warrant further exploration. Specifically, the investigation into women's empowerment within Chinese cultural identity introduces a novel and deserving dimension, meriting thorough examination for a nuanced understanding of gender dynamics in this cultural context.

Moreover, the acknowledgment of information dissemination as a burgeoning area of interest underscores the necessity for research elucidating the mechanisms and impact of information flows on cultural identity formation. Furthermore, the recognition of globalizing Chinese cultural identity as an emerging frontier emphasizes the need for research that delves into the dynamics of cultural identity in the context of globalization, elucidating its implications for individuals and societies. In summary, this analysis provides valuable insights for scholars and practitioners. It

underscores the imperative for focused exploration into these emerging themes to address existing research gaps and propel theoretical frameworks forward in the continually evolving landscape of Chinese cultural identity research. Nevertheless, it is imperative to acknowledge the study's limitations, primarily stemming from its reliance on data from the Web of Science database and the CiteSpace software. While these sources have provided substantial insights, they may not encompass the relevant research in the field. In-depth case studies, as well as comparative and interdisciplinary approaches, will undoubtedly contribute to a more nuanced and holistic understanding of Chinese cultural identity.

Furthermore, an in-depth study of Chinese cultural identity in Hong Kong, Macau, and Taiwan would offer valuable insights and contribute to a comprehensive field examination. In light of these findings, this study underscores the significance of continued research efforts. A pronounced need exists for deeper investigations into the dissemination of this identity and the intricate dynamics surrounding it. The study compellingly asserts the importance of future research endeavors to delve into the nuanced interplay and multifaceted propagation of Chinese cultural identity. Such endeavors will undoubtedly contribute to a more comprehensive understanding of this intricate construct, enriching the scholarly discourse in a pivotal area of inquiry.

Implication for Theory and Practice

The research findings bear noteworthy implications for theoretical development and practical application in Chinese cultural identity studies. The discernment of a two-phase progression in Chinese cultural identity research underscores the dynamic nature of this field, prompting a call for adaptable theoretical frameworks. Scholars and practitioners can leverage these insights to refine existing theories and formulate novel conceptual models that encapsulate the intricate dynamics of Chinese cultural identity over time. Moreover, identifying concentration in specific countries, juxtaposed with the global distribution of research institutions, underscores the need for cultural identity theories to encompass localized and global perspectives. This recognition is instrumental for practitioners navigating multicultural contexts, providing nuanced insights into the varied manifestations and study approaches of Chinese cultural identity across diverse regions.

The delineation of emerging research frontiers, particularly in areas like women's empowerment, the role of information, and globalization, presents compelling opportunities for theoretical advancement and practical application. Delving into these nascent areas allows researchers to expand existing theoretical frameworks and derive practical insights that can inform policies, interventions, and cultural practices pertaining to Chinese cultural identity. In summary, this research furnishes valuable insights that enrich theoretical

comprehension and practical applicability within Chinese cultural identity studies. By integrating these implications into subsequent research endeavors and practical undertakings, scholars and practitioners can contribute to a more nuanced and comprehensive grasp of Chinese cultural identity and its ramifications for individuals, communities, and societies.

Limitation and Recommendations

Despite the valuable insights gleaned from this study, it is crucial to acknowledge certain limitations that warrant consideration in future research endeavors. Firstly, the reliance on data solely from the Web of Science database and CiteSpace software may restrict the comprehensiveness of the findings. Alternative databases and analytical tools should be explored to ensure a more exhaustive exploration of the Chinese cultural identity research landscape. Additionally, while the study offers a broad overview of global trends and patterns in Chinese cultural identity research, it may overlook nuanced regional variations and localized contexts. Future research could address this limitation by conducting in-depth case studies focusing on specific regions or cultural contexts, thereby providing a richer understanding of the complexities inherent in Chinese cultural identity dynamics. Furthermore, the study primarily focuses on bibliometric and scientometric analyses, which may overlook qualitative insights and perspectives equally valuable in understanding cultural identity

phenomena. Future research could employ mixed methods approaches that integrate quantitative analyses with qualitative methodologies, allowing for a more holistic exploration of Chinese cultural identity.

In light of these limitations, several recommendations can be proposed for future research endeavors in Chinese cultural identity. Firstly, scholars should explore alternative databases and analytical tools to complement existing data sources and enhance the robustness of their findings. Secondly, efforts should be made to incorporate regional perspectives and localized contexts into research methodologies, thereby comprehensively capturing the diversity and complexity of Chinese cultural identity dynamics. By addressing these limitations and embracing these recommendations, future research endeavors can contribute significantly to advancing our understanding of Chinese cultural identity and its implications for individuals, communities, and societies.

ACKNOWLEDGEMENT

REFERENCES

Antons, D., Grünwald, E., Cichy, P., & Salge, T. O. (2020). The application of text mining methods in innovation research: Current state, evolution patterns, and development priorities. *R&D Management*, *50*(3), 329-351. https://doi.org/10.1111/radm.12408

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363-376. https://doi.org/10.1162/qss_a_00018

Chan, C. K. K. (2020). Chinese migrants' spatial politics of belonging, identity, and citizenship in Santiago de Chile. *Citizenship Studies*, *25*(1), 106-123. https://doi.org/10.1080/13621025.2020.1825624

Chang, M. X., Jetten, J., Cruwys, T., & Haslam, C. (2016). Cultural identity and the expression of depression: A social identity perspective. *Journal of Community and Applied Social Psychology*, *27*(1), 16-34. https://doi.org/10.1002/casp.2291

Chen, A. (2021). Exploration of implementation practices of Montessori education in mainland China. *Humanities and Social Sciences Communications*, *8*(1), Article 250. https://doi.org/10.1057/s41599-021-00934-3

Chen, C. (2005). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, *57*(3), 359-377. https://doi.org/10.1002/asi.20317

Cheng, A., & Szeto, E. (2019). Changing Hong Kong University students' national identity through studying abroad. *Asian Education and Development Studies*, 8(2), 233-247. https://doi.org/10.1108/AEDS-05-2017-0045

Cheng, V. (2020). The misrepresentation of Hong Kongness. *Museum Worlds*, *8*(1), 149-167. https://doi.org/10.3167/armw.2020.080111

Chung, F. M. Y. (2021). Translating culture-bound elements: A case study of traditional Chinese theatre in the socio-cultural context of Hong Kong. *Fudan Journal of the Humanities and Social Sciences*, *14*(3), 393-415. https://doi.org/10.1007/s40647-021-00322-w

Dacholfany, M. I., Matiala, T. F., Mokodenseho, S., & Hanum, G. K. (2023). Analyzing the landscape of education management research: A bibliometric analysis of scholarly publications on education management. *The Eastasouth Journal of Learning and Educations*, *1*(2), 44-53. https://doi.org/10.58812/esle.v1i02.106

Dai, Y. (2021). A behavioral cultural-based development analysis of entrepreneurship in China. *Administrative Sciences*, *11*(3), Article 91. https://doi.org/10.3390/admsci11030091

Demes, K. A., & Geeraert, N. (2013). Measures matter. *Journal of Cross-Cultural Psychology*, *45*(1), 91-109. https://doi.org/10.1177/0022022113487590

Dervin, F. (2012). Cultural identity, representation and othering. In J. Jackson (Ed.), *The Routledge handbook of language and intercultural communication* (pp. 181-194). Routledge. https://doi.org/10.4324/9780203805640.ch11

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, *133*, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

Fan, T., & Wang, H. (2022). Research of Chinese intangible cultural heritage knowledge graph construction and attribute value extraction with graph attention network. *Information Processing & Management*, *59*(1), Article 102753. https://doi.org/10.1016/j.ipm.2021.102753

Hall, S. (2015). Cultural identity and diaspora. In P. Williams & L. Chrisman (Eds.), *Colonial discourse and post-colonial theory* (pp. 392-403). Routledge. https://doi.org/10.4324/9781315656496

Ho, E. L. (2020). Leveraging connectivities: Comparative diaspora strategies and evolving cultural pluralities in China and Singapore. *American Behavioral Scientist*, *64*(10), 1415-1429. https://doi.org/10.1177/0002764220947754

Hu, Y., & Dai, K. (2021). Foreign-born Chinese students learning in China: (Re)shaping intercultural identity in higher education institution. *International Journal of Intercultural Relations*, *80*, 89-98. https://doi.org/10.1016/j.ijintrel.2020.11.010

Karim, S., Hue, M. T., & Ullah, R. (2021). Acculturative challenges among Pakistani secondary school students in Hong Kong. *International Journal of Intercultural Relations, 82,* 1-11. https://doi.org/10.1016/j.ijintrel.2021.02.006

Kathina, M. M., Otieno, F. O., Gikonyo, P., Khayeka-Wandabwa, C., & Owaki, M. F. (2020). Chinese language and culture interplay in invigorating social development: A context of Kenya. *International Journal of Academic Research in Business and Social Science, 10*(4), 295-303. https://doi.org/10.6007/IJARBSS/v10-i4/7130

Khudhair, A., Li, H., Ren, G., & Liu, S. (2021). Towards future BIM technology innovations: A bibliometric analysis of the literature. *Applied Sciences*, *11*(3), Article 1232. https://doi.org/10.3390/app11031232

King, V. T. (2016). Conceptualising culture, identity and region: Recent reflections on Southeast Asia. *Pertanika Journal of Social Sciences and Humanities, 24*(1), 25-42.

Lam, I. K. V., & Ryan, C. (2020). Can marketing create destination personality? The case of Macau. *Tourism Critiques: Practice and Theory*, *1*(1), 5-20. https://doi.org/10.1108/TRC-03-2020-0006

Lee, K. H., & Wong, D. T. K. (2017). Chinese popular music as a musical heritage and cultural marker of the Malaysian Chinese. *International Journal of Heritage Studies*, *23*(10), 989-1001. https://doi.org/10.1080/13527258.2017.1362577

Li, R. (2022). Chinese folk music: Study and dissemination through online learning. courses. *Education and Information Technologies*, *27*(7), 8997-9013. https://doi.org/10.1007/s10639-022-11003-w

Li, T., Tang, J., Karuppiah, V., Li, Y., Xu, N., & Chen, J. (2020). Co-culture of Trichoderma atroviride SG3403 and Bacillus subtilis 22 improves the production of antifungal secondary metabolites. *Biological Control, 140*, Article 104122. https://doi.org/10.1016/j.biocontrol.2019.104122

Lin, W., & Yang, R. (2019). To borrow or to mix? A cultural approach to observing Taiwan's Higher Education. *Higher Education Forum,16*(1), 143-159.

Lin, Z., & Yang, L. (2019). Individual and collective empowerment: Women's voices in the MeToo movement in China. *Asian Journal of Women's Studies*, *25*(1), 117-131. https://doi.org/10.1080/12259276.2019.1573002

Liu, L. A., Adair, W. L., Tjosvold, D., & Poliakova, E. (2018). Understanding intercultural dynamics: Insights from competition and cooperation in complex contexts. *Cross Cultural & Strategic Management*, *25*(1), 2-31. https://doi.org/10.1108/CCSM-11-2017-0151

Liu, M., & Zhong, J. (2020). Between national and local: Identity representations of post-colonial Hong Kong in a local English newspaper. *Discourse, Context & Media, 36*, Article 100401. https://doi.org/10.1016/j.dcm.2020.100401

Loewenberg, D. (2020). A digital path to a diploma. *Education Next, 20*(1), 50-56. https://www.educationnext.org/digital-path-to-diploma-online-credit-recovery-classes/

Lou, N. M. (2021). Acculturation in a postcolonial context: Language, identity, cultural adaptation, and academic achievement of Macao students in Mainland China. *International Journal of Intercultural Relations*, *85*, 213-225. https://doi.org/10.1016/j.ijintrel.2021.10.004

Lowe, J. (2021). The affective cultural commons of Hong Kong's pro-democracy movement and figurations of the anonymous protestor. *Continuum*, *35*(4), 614-633. https://doi.org/10.1080/10304312.2021.1933905

Luo, M., & Xiao, J. (2022). Institutional policing of Western heritage: The case of Shanghai Symphony Museum. *Museum Management and Curatorship*, *37*(4), 418-436. https://doi.org/10.1080/09647775.2021.2023903

McIntyre, S. (2019). Soft power and the role of art in the development of Taiwan–mainland China relations 1. In J. Beattie, R. Bullen & M. Galikowski (Eds.), *China in Australasia* (pp. 151-169). Routledge. https://doi.org/10.4324/9781351203470-10

Ng, V. S. (2021). Dual-identity incompatibility as a cause of radicalization: A case study of Hong Kong. *Issues & Studies*, *57*(01), Article 2150004. https://doi.org/10.1142/S1013251121500041

Richards, G., & King, B. (2022). The experience of cultural festivals: Evidence from Hong Kong. *Journal of Policy Research in Tourism, Leisure and Events, 14*(3), 296-309. https://doi.org/10.1080/19407963.2022.2033249

Saad, C. S., Damian, R. I., Benet-Martínez, V., Moons, W. G., & Robins, R. W. (2013). Multiculturalism and creativity: Effects of cultural context, bicultural identity, and ideational fluency. *Social Psychological and Personality Science, 4*(3), 369-375. https://doi.org/10.1177/1948550612456560

Sengupta, R. (2023). Negotiating minority identities in Europe through cultural preservation: Music as heritage among the Dutch Hindustani diaspora in the Netherlands. *Journal of Intercultural Studies, 45*(1), 111-126. https://doi.org/10.1080/07256868.2023.2244899

Shen, F. C., & Liao, K. Y. H. (2022). Educational channeling, internalized stereotyping, expectations fulfillment and major choice on Asian Americans' major satisfaction: A moderated mediation model. *Journal of Vocational Behavior, 132*, Article 103647. https://doi.org/10.1016/j.jvb.2021.103647

Song, P., & Wang, X. (2020). A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years. *Asia Pacific Education Review*, *21*, 473-486. https://doi.org/10.1007/s12564-020-09640-2

Sung, C. C. M. (2022). Identity conflicts and negotiations: Narratives of Asian international students' experiences in Hong Kong. *Journal of Intercultural Studies*, *43*(5), 639-656. https://doi.org/10.1080/07256868.2022.2063818

Tang, D. T. (2021). The production of contemporary sociological knowledge in Hong Kong. *Journal of Historical Sociology, 34*(1), 120-133. https://doi.org/10.1111/johs.12310

Telzer, E. H., Fuligni, A. J., Lieberman, M. D., Miernicki, M. E., & Galván, A. (2014). The quality of adolescents' peer relationships modulates neural sensitivity to risk taking. *Social Cognitive and Affective Neuroscienc*e, *10*(3), 389-398. https://doi.org/10.1093/scan/nsu064

Wang, D. (2016). Learning or becoming: Ideology and national identity in textbooks. for international learners of Chinese. *Cogent Education, 3*(1), Article 1140361. https://doi.org/10.1080/2331186X.2016.1140361

Wang, Q. (2022). The digitisation of intangible cultural heritage oriented to inheritance and dissemination under the threshold of neural network vision. *Mobile Information Systems*, Article 6323811. https://doi.org/10.1155/2022/6323811

Wei, L. (2014). Translanguaging knowledge and identity in complementary classrooms for multilingual minority ethnic children. *Classroom Discourse, 5*(2), 158-175. https://doi.org/10.1080/19463014.2014.893896

Wells, G. A., Hsieh, S. C., Zheng, C., Peterson, J., Tugwell, P., & Liu, W. (2022). Risedronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD004523.pub4

Winden, K. P. (2023). *"Do you feel comfortable existing in this space?": A qualitative exploration of college students' sense of belonging, connectedness, and campus climate* (Publication No. 30486010) [Doctoral dissertation, State University of New York at Buffalo]. State University of New York at Buffalo ProQuest Dissertations & Theses. https://www.proquest.com/dissertations-theses/do-you-feel-comfortable-existing-this-space/docview/2828298764/se-2

Wu, Q., Ou, Y., & Jordan, L. P. (2020). Mapping the cultural identities of youths in Hong Kong from a social capital perspective. *Social Sciences, 9*(11), Article 205. https://doi.org/10.3390/socsci9110205

Wyatt, L. C., Ung, T., Park, R., Kwon, S. C., & Trinh-Shevrin, C. (2015). Risk factors of suicide and depression among Asian American, Native Hawaiian, and Pacific Islander youth: A systematic literature review. *Journal of Health Care for the Poor and Underserved, 26*(2), 191-237. https://doi.org/10.1353/hpu.2015.0059

Zhang, J. (2019). Educational diversity and ethnic cultural heritage in the process of globalization. *International Journal of Anthropology and Ethnology*, *3*, Article 7. https://doi.org/10.1186/s41257-019-0022-x

Zhao, K. (2020). Made in contemporary China: Exploring the national identity of Chinese international undergraduate students in the US. *Studies in Higher Education, 45*(12), 2451-2463. https://doi.org/10.1080/03075079.2019.1615046

Žukauskas, P., Vveinhardt, J., & Andriukaitienė, R. (2018). Philosophy and paradigm of scientific research. In P. Žukauskas, J. Vveinhardt, & R. Andriukaitienė (Eds.), *Management Culture and Corporate Social Responsibility* (pp. 506-518). InTech. https://doi.org/10.5772/intechopen.70628

SOCIAL SCIENCES & HUMANITIES

Review Article

A Systematic Literature Review of Design Considerations, Challenges and Guidelines in Primary School Physical Learning Space Design

Rongrong Sun and Muhammad Firzan Abdul Aziz*

School of Housing, Building and Planning, Universiti Sains Malaysia, 11800 Gelugor, Pulau Pinang, Malaysia

ABSTRACT

Many countries are keen to enhance existing learning spaces beyond the status quo, as non-traditional learning spaces can be leveraged to cultivate talent and ability in the 21st century. Recently, many primary schools have begun to practice planning and constructing non-traditional learning. This review highlights the available evidence on the considerations, challenges, and existing learning space design guidelines based on primary-school research conducted from 2000 to January 2024. The Web of Science (WoS) and Scopus databases are intensively searched for research conducted in primary school settings in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The finding shows optimism regarding non-traditional learning spaces fostering more flexible, innovative, and open learning environments that support and assist student-centred pedagogical approaches, and it summarises the three results from the seven aspects. The primary considerations are physical space and pedagogical organisational design, challenges from users and designers and current research and guidelines for users and designers. Based on the three study results, this research proposes suggestions for physical learning spaces. There is an urgent need to design guidelines to promote primary school learning efficiency and create an environment that students and teachers like.

Keywords: Educational building, learning spaces, physical design, primary schools, systematic literature review (SLR)

INTRODUCTION

Job disruptions, the need for new skills, and growing socio-economic polarisation exert increased pressure on primary school systems to educate future workers and global citizens (World Economic Forum, 2020). Global educational approaches

have undergone a substantial revolution in the past two decades to support a student-centred approach to teaching and learning (T&L; Li et al., 2005). Many countries are keen to enhance existing learning spaces beyond the status quo, where non-traditional learning spaces can be leveraged to cultivate abilities and talents in the 21st century, which include cooperation, empathy, social awareness and global citizenship to foster quality education (World Economic Forum, 2020). Following the COVID-19 epidemic, students need to acquire skills like resilience, adaptability, and critical thinking in addition to academic knowledge (Abdollahi et al., 2020), and the teacher has more options for digital technology (Brown et al., 2021).

Learning space design has always attracted much attention and is influenced by the teaching mode and curriculum. Physical learning spaces are the most crucial modern infrastructural necessity for 21st-century education (Uduku, 2015). Physical design is crucial in classroom management (Gremmen et al., 2016) and in promoting the transformation of the teaching and learning model (Szpytma & Szpytma, 2022). Cleveland and Fisher (2014) found that physical learning space design significantly impacts students' academic performance. Schools have begun redesigning classrooms to better utilise the physical environment (Attai et al., 2021) to create adaptable and frequently modifiable learning settings (Gremmen et al., 2016). Comfortable, safe, and flexible learning spaces are more conducive to stimulating students' learning interests and improving learning efficiency (Cardellino & Woolner, 2019; Cleveland & Fisher, 2014; Kariippanon et al., 2017; Vijapur et al., 2021).

Despite these obstacles, the early 21st century is marked by the return of innovative learning spaces (Cardellino & Woolner, 2019). The concepts of "mobile," "agile," and "flexible" learning environments from the twenty-first century, along with conscious pedagogical innovations, are what propel innovative learning spaces (Cleveland & Fisher, 2014). However, these environments can also be difficult for teachers and students as they become used to the new learning and working settings. These difficulties include loudness, visual distractions, and collaborative teaching methods, which can be difficult to implement (Mulcahy & Morrison, 2017). The lack of guidelines also makes designing and building learning spaces a rather open question for those involved (Rönnlund et al., 2021).

There are only two review articles (Jagust et al., 2018; Vijapur et al., 2021) related to the topic of "Primary School Physical Learning Space Design" in the Web of Science (WoS). Vijapur et al. (2021) mainly focused on the interior design and IEQ of non-traditional learning spaces and found a lack of reporting on the specific zoning, layout, and usage patterns. The literature indicates that problems concerning all the considerations of occupants within non-traditional learning spaces have hardly been addressed (Zhang et al., 2019). Jagust et al. (2018) mainly presented the

technology-enhanced learning between learning in formal and informal contexts. The two reviews were limited to ten years: 2007–2016 (Jagust et al., 2018) and 2010–2020 (Vijapur et al., 2021), and there is a limitation to the challenges and guidelines of 21st-century learning space.

This paper aims to highlight the available evidence on the considerations, challenges, and existing learning space design guidelines based on primary-school research conducted from 2000 to 2022 to bridge this gap. The systematic literature review was performed to address the following research questions (RQs):

- **RQ1:** What are the main learning space design considerations that affect teaching and learning activities?

- **RQ2:** What are the primary school learning space design challenges?

- **RQ3:** What existing design guidelines optimise primary school learning spaces?

METHODOLOGY

The systematic literature review methodology was used in this study to provide a thorough overview of the relevant literature and synthesise the findings (Liu et al., 2022). Systematic literature review (SLR) follows certain processes to collect substantial data from various databases, which will be carefully evaluated and categorised either qualitatively or quantitatively (Ghadwan et al., 2022). This study uses the SLR

procedure that Denyer and Tranfield (2006) and Dash and Thilagam used to standardise SLR and writing style (2022). This protocol is known as "The Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA)" (Moher et al., 2010). Figure 1 depicts an overview of the PRISMA statement-based search and selection procedure. Relevant literature was identified using the appropriate inclusion and exclusion criteria (Shaffril et al., 2020; Xiao & Watson, 2019).

Search Strategy

The main databases searched in this study were the Web of Science and Scopus databases, which are prominent scholarly research databases (Gusenbauer & Haddaway, 2020; Martín-Martín et al., 2018). Using an electronic database search, literature from peer-reviewed journals was identified from the reference lists of all relevant articles. Keywords were identified in both databases with full search strings using Boolean operators, phrase searching, and truncation capabilities (Table 1). Only studies that met every one of the following inclusion criteria elements were considered for inclusion:

1. Peer-reviewed journal article in English

2. Published between 2000 and January 2024

3. Empirical paper

Research that focused on virtual space and outdoor areas was not included because

it was outside the scope of this paper. The following keywords were searched in combination with two databases. The search results are shown in Table 1.

- [primary school OR elementary school] +

- [design]+

- [traditional classroom (OR) physical learning space (OR) active learning space (OR) flexible learning space (OR) innovation learning space]

Table 1
Search strings

Database	Search string	Results
Scopus	"Primary school" OR "elementary school" AND "design" AND "traditional classroom" OR "physical learning space" OR " active learning space" OR "flexible learning space" OR "innovation learning space"	756
Web of Science	TS = (primary school OR elementary school) AND TS = (design) AND TS = (traditional classroom OR physical learning space OR active learning space OR flexible learning space OR innovation learning space)	355

Source: Authors' work

According to the PRISMA flowchart summary (Figure 1), a preliminary database, which included 1111 articles, was produced by reviewing the keywords across two databases. The initial screening step involved scanning the abstracts and titles and removing duplicates. A total of 29 duplicates were removed, and 871 articles were excluded to eliminate non-relevant articles that focus on teaching and learning, such as online learning, flipped classrooms, and blended learning. In a second stage review of the remaining 211 articles, 113 were excluded as the full text was unavailable and only focused on technology and the physical activities of spaces. Of the remaining 98 articles, 65 were excluded as they were teaching and learning, not primary school, and not conducted in a physical learning space when the researcher read the articles. Thus, 33 articles that had high-quality appraisals were included in the review.

Quality Appraisal

Quality appraisal is essential for determining bias in systematic review studies. The authors evaluated the 24 qualitative articles using the Critical Appraisal Skills Programme (CASP) qualitative studies checklist, and eight quantitative articles used observational cohort and cross-sectional studies developed by the National Institute of Health (National Heart, Lung, and Blood Institute, 2021). CASP contains ten appraisal questions to assess the risk of bias in qualitative studies. The evaluation procedure was recorded in Microsoft Excel, employing indicators such
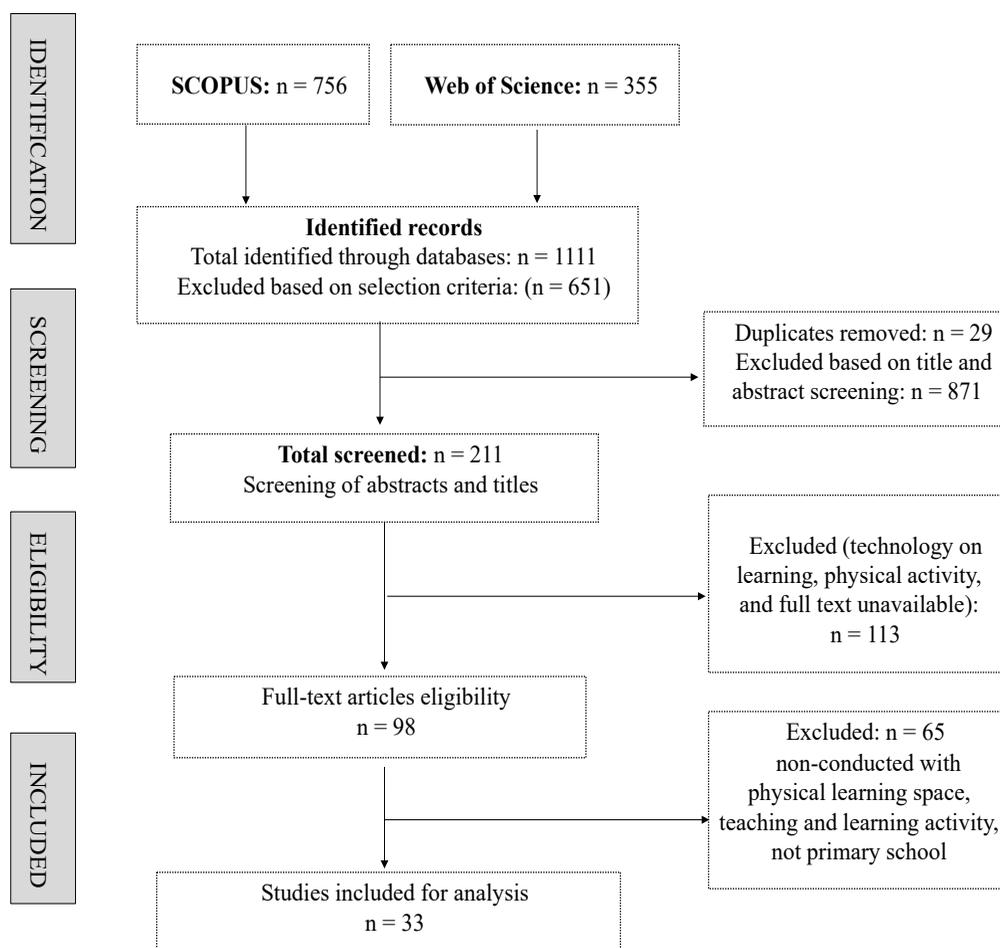
Figure 1. The PRISMA flow diagram summarising the study selection

Source: Authors' work

as Yes (Y), No (N), and Can't Tell (CT). The evaluation method revealed that the chosen studies were fair (scoring seven to nine out of 10 items). Observational Cohort and Cross-Sectional Studies consist of 14 items to which the evaluators would respond with a "yes," "no," or other ("CD: cannot determine", "NA: not applicable", or "NR: not reported",). As a result, the evaluation method showed that the chosen studies had a fair quality (scoring eight and 12 out of 14 items).

Data Analysis

The 34 articles were analysed using the qualitative thematic analysis method, which followed six steps (Braun & Clarke, 2006), as shown in Figure 2. As introduced by Zairul et al. (2023), it is imperative to verify the metadata of the chosen articles to ensure they fulfil the requirements and suit. We have double-checked the publication dates to ensure the articles are published during the evaluation period. A procedure was

followed to make sure that all relevant data was provided. The included studies' abstract, findings, discussion, and conclusion sections were examined carefully. Data that could address the RQs was collated and analysed thematically. The thematic analysis was conducted by grouping, condensing, and examining parallels and correlations to identify recurring themes. Three themes and seven subthemes were summarised from the 28 initial codes in this study. The authors' primary reason for creating these codes was the articles' heavy emphasis and repetition. It is common for the researcher to code the article abstract, introduction, results, discussion and conclusions, then to expand and summarise those codes into themes and subthemes. The themes, sub-themes and codes are shown in Table 2.
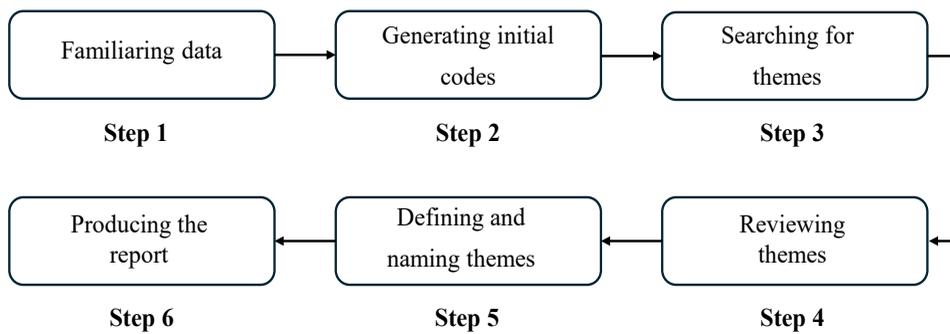
Figure 2. Six steps of thematic analysis
Source: Braun and Clarke (2006)

Table 2
Themes, subthemes and initial codes

Themes	Sub-themes	Initial codes
Physical space design and pedagogical organisation	Physical space design	Naturalness (visual -windows, door, nature; acoustic; temperature)
		Comfort
		Ownership
		Flexibility
		Complexity (colour)
		Child-friendly and learning needs
		Layout
		Furniture (desk, seat)
	Pedagogical organisation	Organisation (scheduling and curriculum)
		Staff culture
		Student milieu (motivation, social climate)
		Principals and senior leaders support

Table 2 *(Continue)*

Themes	Sub-themes	Initial codes
Users, designers and research challenges	Users challenge	Leader: Insufficient mechanisms Teacher: New forms of leadership and collaboration Greater empathy and appreciation Teacher educated decisions Students: Student skills and coping strategy
	Designer challenge	Combine design factors Formed decisions
	Research challenge	Small sample size Fewer theories Investigate many socio-material components, instructional strategies, ergonomic health concerns, and student and teacher control levels
Design guidelines for users and designers	User guidelines	User participant design process Professional learning mechanisms Pedagogy consistency
	Designer guidelines	Flexible classroom supplies Social interaction Students' desires and expectations

Source: Authors' work

RESULTS AND DISCUSSION

Each article was examined to determine the (1) research technique, (2) study purpose, (3) learning space (traditional or non-traditional), and (4) number of participants. Table 3 summarises the analysis findings. Among the 34 articles in the review, 17 focused on non-traditional learning spaces, innovative learning environments (ILEs), open learning spaces, and flexible learning environments. Five articles focused on traditional learning spaces, one focused on mixed learning spaces, and others did not mention space type. One article reported a comparative study of traditional and non-traditional spaces. Twenty-four studies used qualitative research methods, 8 used quantitative, and 2 reported mixed-method studies.

We categorised and assessed the papers. The original codes underwent multiple rounds of recoding, merging, and categorisation. Codes that were used infrequently and not related to the questions were removed. Three themes emerged in the end, as shown in Table 4: (1) physical space design and pedagogical organisation, (2) users, designers and research challenges, and (3) design guidelines for users and designers. Results outside the topics or the evaluated publications will be provided for explanatory purposes when necessary. Each theme comprises two or three sub-themes, which were further studied below (Figure 3).

Table 3
The analysis results

No.	Author	Country	Space	Participant	Method	Study purpose
1.	Niemi et al. (2022)	Finland	NT	N = 684 S (8–13 years)	Quantitative	Investigate students' opinions on flexible and open learning environments and the degree to which these views relate to student satisfaction with their education.
2.	Reinius et al. (2021)	Finland	NT	N = 17 S (second-grade)	Qualitative	Examine the activities that students and teachers engage in at 'deskless schools' (flexible physical classroom spaces)
3.	Campbell et al. (2013)	Australia	NT	N = 3 schools	Qualitative	Determine how non-traditional settings influence instructional practices and the challenges created for professional development.
4.	Saltmarsh et al. (2015)	Australia	NT	N = 3 schools	Qualitative	Determine the influence of teachers' ideas and methods of applying pedagogy theories on open-plan spaces.
5.	Starkey et al. (2021)	New Zealand	NT	N = 1 school	Qualitative	Investigate furniture usage in a flexible learning environment.
6.	Carvalho et al. (2020)	New Zealand	NT	N = 222 primary school teachers, 126 secondary school teachers, and 163 school leaders	Mixed method	Investigate the relationship between learning environments and academic courses and the influence of learning environments on primary school students' performance in mathematics and the arts.
7.	Cardellino and Woolner (2019)	UK	NT	N = One school	Qualitative	Analyse the relationships between teaching, learning, and space to support education changes.
8.	Reh et al. (2011)	Germany	NT	N = One school	Qualitative	Examine the relationship between educational practices, power, and space.

Table 3 *(Continue)*

No.	Author	Country	Space	Participant	Method	Study purpose
9.	Oliveras-Ortiz et al. (2021)	USA	NT	N = 35 S	Qualitative	Explore students' perceptions of the influence of learning environment design on student engagement.
10.	Baloğlu (2019)	Turkey	NT	N = 142 T from 15 primary schools	Qualitative	Explore the close relation of educational philosophies and methods with the built environment design that accommodates them.
11.	Rönnlund et al. (2021)	Sweden	T, NT	N = 20 principals, school managers, and architects	Qualitative	How stakeholders perceive social, pedagogical, and physical components of learning settings.
12.	Harouni (2013)	USA	T	N = 3 T and 3 S	Qualitative	Explore the recycled uniformity relationship between a creative public school teacher and the school settings
13.	Herman and Tondeur (2021)	Belgium	*	N = 9 T	Qualitative	Explore the space-human activity relationship.
14.	Barrett et al. (2015)	UK	*	N = 3766 S	Quantitative	Determine how physical classroom elements affect students' academic development.
15.	Wang and Wang (2020)	China	T	N = 17 T	Qualitative	Determine teachers' willingness to engage in cooperative learning and the challenges of cooperative learning in space.
16.	López-Chao et al. (2020)	Spain	T	N = 583 S	Quantitative	Investigate the relationship between learning environments and academic courses and the influence of learning environments on primary school students' performance in mathematics and the arts.
17.	Deppeler et al. (2022)	Australia	NT	N = 1 school (architect, principals, 4T, 98 S)	Qualitative	Examine how the architects and school administrators comprehended the risks, values, and educational goals that influenced the design of the new school building and gauge how the community views it.

Table 3 *(Continue)*

No.	Author	Country	Space	Participant	Method	Study purpose
18.	Mackey et al. (2018)	Australia	NT	N = 28 T and 16 school leaders from a total of 17 schools	Mixed method	Examine the experiences of principals and teachers in moving to a co-teaching environment, highlighting some of the new opportunities and difficulties that leaders have in helping teachers function well in these cooperative settings.
19.	Sigurdardottir and Hjartarson (2016)	Iceland	NT	N = 1 school	Qualitative	Expose the educational principles underlying a comparatively innovative school design and assess their performance in classroom settings.
20.	Bluteau et al. (2022)	France	NT	N = 107 S	Quantitative	Analyse how flexible seating affects elementary school children's mental and general well-being.
21.	Attai et al. (2021)	USA	*	N = 206 S from 10 classrooms	Qualitative	Explored the different effects of adaptable furniture in conjunction with professional development (PD) for teachers in primary schools.
22.	Tokarek et al. (2022)	USA	*	N = 99 S	Quantitative	Examine whether students use stand-biased desks during the school day affects their after-school sedentary and physical activity habits.
23.	Swartz et al. (2020)	USA	*	N = 97 S	Qualitative	Assess the effects of a stand-biased desk on fidgeting and attention.
24.	Wallace et al. (2022)	USA	*	N = 99 S	Quantitative	Stand-biased desks showed moderate improvements in cognition and no deleterious effects, suggesting that they may be a helpful classroom intervention for elementary school children. There were no significant effects on cognition or behavioural ratings from the standing-biased desk intervention.
25.	Yao et al. (2024)	China	T	N = 40 S	Quantitative	Investigate the influence of electronic screen viewing distance on the visual perception comfort of students.
26.	To and Grierson (2019)	Scotland	*	N = 3 schools	Qualitative	Measure the natural experience of children within three primary school Spaces at different locations in Glasgow, Scotland.

Table 3 *(Continue)*

No	Author	Country	Space	Participant	Method	Study purpose
27.	Szpytma and Szpytma (2022)	Poland	T	N = 1 school	Qualitative	Explore Elementary school buildings in Poland, tracing their evolution and asking whether their problems are universally applicable.
28.	Nyabando and Evanshen (2022)	USA	NT	N = 16 S from 4 classrooms and 3 distracts	Qualitative	Examine how second-grade students view the physical learning environment in their classrooms.
29.	Mokhtarmanesh and Ghomeishi (2019)	Iran	*	N = 1 female elementary schools	Qualitative	Investigate the influence of the environment on users' decisions on school preference.
30.	S. Lee (2019)	Korea	*	N = School space innovation projects	Qualitative	Examine numerous recent initiatives to reorganise classrooms to identify their concrete features.
31.	Whitehouse (2009)	USA	NT	N = 1 school	Qualitative	Test the current iterations of transformational school design.
32.	Dolan et al. (2006)	USA	*	N = 1 school	Qualitative	Recognise as essential to a student's comfort as well as the particular tactics employed to preserve and improve the environment, a feeling of place, and a sense of neighbourhood.
33.	Killeen et al. (2003)	USA	*	N = 77 S	Quantitative	Examine whether children's ownership of their educational experience would increase if the artwork were permanently displayed in a school and integrated into the building's architecture.

Note. * = not mentioned; In space line: "NT" means non-traditional space; "T" means traditional space; In participant line: "S" means students; "T" means teacher.

Table 4
The author and theme originated

	Author	Physical space design and pedagogical organisation		Users, designers and research challenges			Design guidelines for user and designer	
		Physical space design	Pedagogical organisation	Users' challenges	Designers' challenges	Research challenges	User' Guideline	Designer Guideline
1.	Niemi et al. (2022)	-	-	-	-	✓	✓	✓
2.	Reinius et al. (2021)	-	-	✓	-	✓	-	-
3.	Campbell et al. (2013)	-	✓	✓	-	-	-	-
4.	Saltmarsh et al. (2015)	✓	-	-	-	✓	✓	-
5.	Starkey et al. (2021)	-	-	✓	-	✓	-	-
6.	Carvalho et al. (2020)	-	-	✓	-	✓	-	-
7.	Cardellino and Woolner (2019)	✓	✓	-	-	-	✓	-
8.	Reh et al. (2011)	-	-	✓	-	✓	-	-
9.	Oliveras-Ortiz et al. (2021)	-	✓	-	-	-	-	-
10.	Baloğlu (2019)	✓	-	-	-	✓	-	-
11.	Rönnlund et al. (2021)	-	-	-	-	✓	-	-
12.	Harouni (2013)	-	-	-	-	✓	-	-
13.	Herman and Tondeur (2021)	✓	-	-	-	-	-	-
14.	Barrett et al. (2015)	✓	-	-	✓	-	-	✓
15.	Wang and Wang (2020)	-	✓	✓	-	-	-	-
16.	López-Chao et al. (2020)	-	-	-	-	-	-	✓

Table 4 *(Continue)*

	Author	Physical space design and pedagogical organisation		Users, designers and research challenges			Design guidelines for user and designer	
		Physical space design	Pedagogical organisation	Users' challenges	Designers' challenge	Research challenges	User' Guideline	Designer Guideline
17.	Deppeler et al. (2022)	-	-	-	/	-	-	-
18.	Mackey et al. (2018)	-	/	/	-	-	-	-
19.	Sigurdardottir and Hjartarson (2016)	/	-	-	-	-	-	-
20.	Bluteau et al. (2022)	/	-	/	-	-	-	-
21.	Attai et al. (2021)	/	-	-	-	-	-	-
22.	Tokarek et al. (2022)	/	-	-	-	-	-	-
23.	Swartz et al. (2020)	/	-	-	-	-	-	-
24.	Wallace et al. (2022)	/	-	-	-	-	-	-
25.	Yao et al. (2024)	/	-	-	-	-	-	-
26.	To and Grierson (2019)	/	-	-	-	-	-	-
27.	Szpytma and Szpytma (2022)	-	-	-	-	/	-	/
28.	Nyabando and Evanshen (2022)	-	-	-	-	/	-	-
29.	Mokhtarmanesh and Ghomeishi (2019)	/	-	-	-	-	-	-
30.	S. Lee (2019)	/	-	-	-	-	-	-
31.	Whitehouse (2009)	-	-	-	-	-	/	-
32.	Dolan et al. (2006)	/	-	-	-	-	-	-
33.	Killeen et al. (2003)	/	-	-	-	-	-	-

Note. "-" = not mentioned; "/" = mentioned
Source: Authors' work

Theme 1 Physical space design and Pedagogical Organisation	Theme 2 Users, designer and research Challenges	Theme 3 Design guideline for user and designer
Sub-theme 1: Physical space design	Sub-theme 3: Users Challenge	
Sub-theme 2: Pedagogical Organisation	Sub-theme 4: Designer Challenge	Sub-theme 6: User guideline
	Sub-theme 5: Research Challenge	Sub-theme 7: Designer guideline
RQ1 What are the main learning space design considerations that affect teaching and learning activities?	RQ2 What are the primary school learning space design challenges?	RQ3 What are the primary school learning space design challenges?

Figure 3. Overall network and how the themes answer the research question
Source: Authors' work

Physical Space Design and Pedagogical Organisation

To address RQ1, the 34 articles were inductively examined to produce categories of recurrent considerations, as shown in Table 5. The data were read and re-read to identify and categorise the derived factors into one theme: physical space design and pedagogical organisation.

Table 5
Design consideration

Design consideration		Citation
Physical space design	visual (windows, door, nature)	Barrett et al. (2015), To and Grierson (2019)
	acoustic	Sato and Bradley (2008)
	temperature	Barrett et al. (2015)
	furniture (desk, seat)	Attai et al. (2021), Bluteau et al. (2022), Swartz et al. (2020), Tokarek et al. (2022), Wallace et al. (2022)
	colour	Barrett et al. (2015), Mokhtarmanesh and Ghomeishi (2019)
	comfort	Dolan et al. (2006)
	ownership	Barrett et al. (2015), Herman and Tondeur (2021), Killeen et al. (2003)
	flexibility	Barrett et al. (2015), Bluteau et al. (2022), Gultekin and Ira (2022)
	complexity	Barrett et al. (2015)
	layout	Baloğlu (2019), Cardellino and Woolner (2019), K. S. Lee et al. (2019)
	child-friendly design and design for learning needs	Gultekin and Ira (2022)

Table 5 *(Continue)*

Design consideration		Citation
Pedagogical organisational design	Organisation (scheduling and curriculum) Staff culture Student milieu (motivation, social climate) Student engagement	Cardellino and Woolner (2019) Oliveras-Ortiz et al. (2021)
	Principals and senior leaders support	Mackey et al. (2018)

Source: Authors' work

Physical Space Design

The transition to Education 4.0 shifted learning content and experiences, which imposed new demands on learning spaces. Educational methods must be modified to confer students with the knowledge and abilities to construct a more diverse, harmonious, and productive world (World Economic Forum, 2020). Furthermore, global educational policies emphasise teaching strategy modernisation and creating innovative learning environments that support 21st-century learning, such as the English and Welsh Building Schools for the Future initiative, the Canadian, Peruvian, and New Zealand ILE projects and the Australian Building the Education Revolution government programme. These new open-plan school designs were driven by the modern themes of mobile, agile, and flexible learning spaces (Dovey & Fisher, 2014).

The included articles demonstrated that physical learning spaces have become more open and flexible. Non-traditional physical learning spaces aided the improvement of cooperative and collaborative learning skills (Campbell et al., 2013), which align with the evolving demands of 21st-century education. Generally, modern learning spaces include technologically advanced areas without assigned student desks or teacher podiums. Larger spaces function as versatile working places given the presence of flexible furniture, different seating arrangements, removable dividers, and acoustic curtains or movable walls that encourage grouping (Campbell et al., 2013; Niemi et al., 2022; Saltmarsh et al., 2015). Students can occupy different positions in various locations, which enables more flexible physical activity and movement throughout the space. Flexible furniture and materials of non-traditional spaces require students to have self-control, problem-solving, self-reliance, cooperation, and soft skills such as working together to make the right decision (Bluteau et al., 2022). The non-traditional space is more flexible to adapt different teaching methods to align with the evolving demands of 21st-century education.

Expanding traditional classrooms provided students with more options on how

time, space, working groups, and working arrangements could be used while also conferring greater autonomy according to the pedagogical approach. Students in open and flexible learning environments frequently have more options for personalised learning, planning, decision-making, and taking ownership of supervising their work (Bøjer, 2019; Yeoman, 2018). Nevertheless, many students preferred classrooms with desks, single teachers, and walls to open school buildings. Additionally, several students complained that their schools were 'too modern' or 'too open'.

Many learning technologies, such as interactive whiteboards, personal learning environments, wireless networks, mobile devices, the internet, and high-quality digital learning resources, can be accessed from home and the workplace. Accordingly, such technologies are changing how students experience learning and what they hope to achieve (Joint Information Systems Committee, 2006). The literature review revealed that applying technology (mobile devices and their software) significantly improved students' learning interest and efficiency.

Digital technology is not only interactive but also provides teachers with more options than those they are accustomed to in the traditional physical learning environment (Brown et al., 2021). Digital technology also gains students' attention and interest, which enhances their learning motivation (Hoon & Shaharuddin, 2019). An examination of cognitive outcomes to determine whether children learned more with technology than by traditional classroom learning participation revealed that children achieved similar knowledge improvements with a self-directed learning method (iPhone game) and a guided method (traditional lesson; Furió et al., 2015). children can learn regardless of location and time without being present in a formal learning environment and do not require personal supervision.

Ahmad et al. (2017) reported that pupils demonstrated a modest understanding of the concept of technology, which suggested insufficient computers in the classroom. Osman et al. (2011) reported that no computers were available for students to use in the classroom. The classroom should feature technology, such as laptops and an internet connection, to facilitate obtaining T&L materials. Classroom technological equipment must, therefore, undergo periodic upgrading to satisfy pupils' needs.

Understanding how the created environment is used and perceived is crucial to physical space design. Woolner (2014) proposed a theory of user participation for effective design to enhance learning effectiveness, where user participation in school design is encouraged to ensure design alignment and comprehension among users. Recent studies confirmed that end users (students and teachers) should participate in designing and implementing new learning environments where students frequently prefer soft furnishings, vibrant colours, and technology-rich learning spaces in open learning environments and institutions (Niemi et al., 2022). Nonetheless, few designers considered users' wishes and

expectations (Niemi et al., 2022) and lacked a deeper understanding of classroom educational practices and how people, material, and space interact (Herman & Tondeur, 2021). Bruce Jilk successfully designed an innovative school in Iceland that is still being used as intended, where he included school officials and architects in the building factors of the organisation, pedagogy, and school space (Sigurdardottir & Hjartarson, 2016). Nevertheless, at the policy level, England's user participation design process was time-consuming, inefficient, and expensive (James, 2011). The design phase was considered time-consuming, but the users' long-term happiness disproved James's charges of inefficiency regarding the participatory design.

Pedagogical Organisation

A successful design is related to staff culture, student dynamics, and school organisation, which influence the infrastructure performance in a T&L environment (Gislason, 2010). Gislason's (2010) school climate model comprised four interconnected dimensions: organisation, staff culture, student milieu, and ecology. Gislason highlighted that these four elements need to interact with each other. This consistency is even more important when implementing non-traditional spaces involving new practices, organisation, and teaching. Organisation is vital to any learning environment (Cardellino & Woolner, 2019). Effective organisation and successful design are essential whenever

non-traditional spaces are implemented (Gislason, 2010) and traditional spaces. Space design can foster student engagement (Oliveras-Ortiz et al., 2021). Time and investment are two essential aspects of effectively using non-traditional learning spaces in successfully transitioning teacher practice and pedagogies (Campbell et al., 2013). Thus, non-traditional learning spaces might be rapidly abandoned if teacher and administrator support is inadequate or the government does not allocate the necessary time, budget, and training.

Users, Designer and Research Challenges

The 34 articles were inductively examined to identify the key challenges to addressing RQ2. The data were analysed and summarised in terms of the themes of users, designers, and research challenges, as shown in Table 6.

Users' Challenges

Successful learning space design is significant in promoting new teaching methods (Niemi et al., 2022; Reinius et al., 2021; Saltmarsh et al., 2015; Starkey et al., 2021). The rapid evolution of non-traditional learning spaces has resulted in insufficient mechanisms for professional learning assistance (Campbell et al., 2013). New learning spaces require users to possess new teamwork skills (Campbell et al., 2013). Specifically, teachers must adopt new forms of leadership and collaboration among their colleagues and students. Teachers must be environmentally competent to make wise

Table 6
The summary of challenges

Theme	Category	Finding	Author
Users, designers and research challenge	User	**Leader:** Insufficient mechanisms **Teacher:** New forms of leadership and collaboration Greater empathy and appreciation Teacher educated decisions Co-teaching **Students:** Student skills and coping strategy	Bluteau et al. (2022), Campbell et al. (2013), Mackey et al. (2018), Starkey et al. (2021)
	Designer	Combine design factors Formed decisions	Barrett et al. (2015, 2016), Gislason (2010)
	Research	Small sample size Investigate many socio-material components, instructional strategies, ergonomic health concerns, and student and teacher control levels One-item indicator measurement error describes a trend toward innovative school design	Harouni, (2013), Niemi et al. (2022), Nyabando and Evanshen (2022), Rönnlund et al. (2021), Starkey et al. (2021)

decisions about utilising the classroom and help students make furniture use decisions (Carvalho et al., 2020). The teachers also need to establish expectations, plan the curriculum, arrange the area, and help the kids learn how to use it (Carvalho et al., 2020). Given that new learning spaces disrupt traditional teaching approaches, teachers are required to demonstrate greater empathy and appreciation for others practising in the space (Campbell et al., 2013). It also challenges how students can use, alter, and compromise their visibility and audibility in a school that has been "opened," as well as who, what, and how they can exhibit themselves (Reh et al., 2011). Overall, users should be able to make educated decisions about their environmental effects when they use the classroom and advise students on how to do the same (Starkey et al., 2021). Nonetheless, few studies examined the effects of new settings from the users' viewpoints and behaviours (Grannäs & Stavem, 2021).

Designer' Challenges

The learning space is a dynamic ecosystem that should be designed to meet educational practices and is considered a constantly shifting meshwork of interacting people, places, and things (Ingold, 2011). Policymakers and designers should understand classroom occurrences and users' practices.

López-Chao et al. (2020) stated that academic performance is directly related to room size, views, ventilation, and place attachment. Focus on space design enabled the identification and typifying of three design principles and seven environmental parameters (naturalness: light, temperature and air quality; individualisation: ownership and flexibility; simulation: complexity and colour), all of which form optimal learning spaces for students and improve pupils' academic performance (Barrett et al., 2015, 2016). According to Barrett et al. (2015), the naturalness design principle accounts for approximately 50% of the impact on learning, while the other two principles comprise approximately one-quarter. Nevertheless, several elements are designer issues and aid user adaptation of spaces to better support learning. Thus, addressing these factors in combination constitutes a design challenge (Barrett et al., 2015, 2016). The lack of educational building studies represents a substantial gap in education studies and indicates that designers and users rarely make informed decisions about school space design and use (Gislason, 2010).

Research Challenges

There are fewer theories in this field, which include only the school climate model (Gislason, 2010) and the SIN model (Barrett et al., 2015). More advanced theoretical and analytical techniques are also required to comprehend the essential components of learning environments, how designed environments shape interactions, and to develop instruments for assessing learning environments (Carvalho & Goodyear, 2014). There is also less sample size (Johler et al., 2022; Nyabando & Evanshen, 2022; Yao et al., 2024) and limited schools (Mackey et al., 2018). Research on the ongoing interactions of material, spatial, and human actors was scarce (Herman & Tondeur, 2021). Research has examined how changing school architecture affects users' viewpoints and behaviours (Reinius et al., 2021). Further study is required to examine furniture used in various modern classroom scenarios to create comprehensive literature about many facets of environmental competence. It entails looking into various socio-material components, pedagogical strategies, ergonomic health concerns, and student and teacher control levels (Carvalho et al., 2020).

More time and more research are needed to determine how innovative practices and educational changes will play out in the long run (Sigurdardottir & Hjartarson, 2016). Numerous theoretical debates centre on social structures, often known as "structurally moulded conditions," (Archer, 2003, p. 13) and how they affect people's ability to act in an agentive manner (Saltmarsh et al., 2015). It is difficult to investigate many socio-material components, instructional strategies, ergonomic health concerns, and student and teacher control levels (Starkey et al., 2021). Harouni (2013) found that educational environments cannot just describe socialising and control mechanisms. It needs to take into account people's attempts to fight back against the limits

within these systems. It is challenging when the stakeholders describe a trend towards innovative school design, and teachers' points have been ignored (Rönnlund et al., 2021).

Design Guideline for User and Designer

This paper summarised some guidelines to answer RQ3 (Table 7).

User' Guideline

The absence of design guidelines and standard definitions renders it challenging for teachers to manage spaces that are difficult to define (Dovey & Fisher, 2014). Consequently, the government should formulate policies to support users in managing and using their spaces. Policies supporting new learning space transformations and applications for users should be explored. The government and education department should review its regulations and promote the most recent teaching techniques, educational technologies, and equipment to users to establish efficient learning environments (Whitehouse, 2009).

Additional studies should examine stakeholders' influences on policy in relation to modern pedagogical approaches and the school environment (Rönnlund et al., 2021). It is also important to consider how teachers use learning spaces, including using user feedback (allowing teachers to collaborate with designers to create school buildings) from the design process to enhance the end design (Bøjer, 2019; Cardellino & Woolner, 2019).

Designer Guideline

School building quality affects academic performance (Barrett et al., 2015), as pupils share an apparent relationship with the environment (López-Chao et al., 2020). Sensory stimulation can both enhance and diminish a young child's learning process. Most primary school students are six to 12 and active, with unyielding curiosity and poor concentration. Therefore, the school design should be combined with students' age characteristics. The physical learning environment, furnishings, and flexible learning space resources should support and enable student-centric instructional practices.

Table 7
Design guidelines

Guideline
• Time and investment in change management should be considered for a successful shift of pedagogies and teacher practice to use agile and flexible learning spaces (Campbell et al., 2013).
• Professional learning mechanisms and support should be offered to meet this potential (Campbell et al., 2013).
• Students' desires and expectations for their schools and school days should be considered when redesigning a learning environment (Gultekin & Ira, 2022; Niemi et al., 2022).
• End users should be involved in the phases and processes of designing and implementing new learning environments (Gultekin & Ira, 2022; Niemi et al., 2022).

Table 7 *(Continue)*

Guideline
• A physical learning environment should be adaptable, open, interactive, aesthetically pleasing (or tailored to the kids' preferences), flexible, and integrated into the community. It should also be meant to serve as an educational tool (Gultekin & Ira, 2022).
• Worldwide, more multidisciplinary studies in architecture, education, and psychology should focus on purposefully designed school environments (Szpytma & Szpytma, 2022).
• The pedagogy should be aligned with the physical learning spaces (Gultekin & Ira, 2022).
• There should be interesting focal points in classrooms and enough room for group work. As a result, classroom supplies (such as chairs, tables, and technology) need to be flexible (Gultekin & Ira, 2022).
• Physical learning environments should be effective and extend learning beyond traditional classroom settings (Gultekin & Ira, 2022).
• Classrooms should have focal points and physical layouts that facilitate group work. For this reason, the classroom furniture, technology, and presenting tools need to be flexible (Gultekin & Ira, 2022).
• Social interaction should be encouraged in the design learning settings process (Gultekin & Ira, 2022).

Flexibility	Learning spaces should respond to diverse concurrent instructional activities.
	• Shared spaces between teachers, content areas, and communities can provide opportunities for ideas and socialisation.
	• There should be individualised, personal, or small-group inquiry-based learning spaces.
	• Open-plan informal learning spaces should be built from unused areas and provide personalised learning environments.
	• Learning place adaptability must be maximised.
	• Classrooms must allow differing group sizes and activities.
	Furniture
	• Must be flexible and sufficiently adaptable to be rearranged throughout the day.
	Wall
	• Wall system use should be flexible.
	• Walls should be movable to increase the effectiveness of multipurpose large spaces, such as gymnasiums and libraries.
Complexity	Learning spaces should respond to multidisciplinary teaching.
	• Interdisciplinary learning, thinking, and working should be used to identify issues and devise fresh creative methods to resolve them.
Ownership	• The main entry should be inclusive, exciting, welcoming, open, bright, safe, well-maintained, and demonstrate the school community values.
	• All school spaces should encourage learning, collaboration, and socialisation.
	• Space design should consider the requirements of disabled administrative staff, teachers, and pupils.

Table 7 *(Continue)*

	Guideline
Colour	The space should feature cold, soft colours.
	• Cold colours increase awareness levels and enhance concentration and memory.
	•A space should not contain more than three significant brightly coloured areas.
Light	• Daylighting (control of natural light entering the space) should be a priority. Glare and hot spots reduce learning effectiveness.
	• Controls, high-efficiency artificial lighting, and daylighting should be incorporated to maximise visual comfort.
Air quality	• The space must ensure superior indoor air quality by using natural ventilation.
	• Natural ventilation should improve engagement. The indoor space should be connected to the outdoors by providing easy access from classrooms to gardens and view windows in classrooms and other outdoor areas that can be utilised in the curriculum. Windows that face classrooms or other active areas (teacher prep areas or break rooms) should feature opening sashes.
Temperature	• The space must ensure thermal comfort.
	• Teachers can control the temperature of individual classrooms.
	• Appropriately sized heating, ventilation, and air conditioning (HVAC) systems should maintain comfortable humidity.

Source: Authors' work

CONCLUSION

This systematic review obtained the best available evidence globally to summarise the state of the art of physical learning spaces, wherein the considerations, challenges, design guidelines for spaces and pedagogical support were examined and analysed. The finding shows that the innovative learning space is easy to align with the evolving demands of 21st-century education, and technology can improve pedagogy and space transformation. From the review, there are two key findings:

First, compared to traditional learning spaces, new learning spaces are compatible with the most recent learning approaches and create a more flexible and engaging learning environment. The physical learning environment, furniture, and resources of flexible learning spaces can support and assist student-centred pedagogical approaches.

Second, researchers were concerned about the influence of technology on learning efficiency. Most research has demonstrated that technology can capture students' interest and attention as a learning tool. Recent studies focused on technology to promote T&L activity and suggested that learning space design should enhance technology use. However, there are insufficient computers in the classroom, and the level of student technology knowledge is low.

In the future, the government should support the new learning space by allocating the required time, money, and training. It also needs to provide the necessary professional learning assistance for users with sufficient mechanisms. Sufficient computers should be provided in the classroom, and student technology knowledge is low. Users are required to have new collaborative abilities and greater empathy and appreciation for others practising in the new space. The effects of new settings from the users' viewpoints and behaviours should be explored for the research.

This review did not identify a detailed guide for designing new learning spaces. Nevertheless, some considerations and directions were determined regarding the expected environment of schools, students, and teachers. Therefore, there is an urgent need to design detailed guidelines for promoting primary school learning efficiency and creating an environment that students and teachers like.

ACKNOWLEDGEMENT

REFERENCES

Abdollahi, A., Panahipour, S., Tafti, M. A., & Allen, K. A. (2020). Academic hardiness as a mediator for the relationship between school belonging and academic stress. *Psychology in the Schools*, *57*(5), 823-832. https://doi.org/10.1002/pits.22339

Ahmad, C. N. C., Shaharim, S. A., & Abdullah, M. F. N. L. (2017). Teacher-student interactions, learning commitment, learning environment and their relationship with student learning comfort. *Journal of Turkish Science Education*, *14*(1), 57-72.

Archer, M. S. (2003). *Structure, agency and the internal conversation*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087315

Attai, S. L., Reyes, J. C., Davis, J. L., York, J., Ranney, K., & Hyde, T. W. (2021). Investigating the impact of flexible furniture in the elementary classroom. *Learning Environments Research*, *24*(2), 153-167. https://doi.org/10.1007/s10984-020-09322-1

Baloğlu, Y. B. (2019). Re-defining the boundaries at schools: Perspectives from teachers' interpretations of sources of spatial change. *International Journal of Architectural Research*, *14*(1), 5-17. https://doi.org/10.1108/ARCH-04-2019-0088

Barrett, P., Davies, F., Zhang, Y., & Barrett, L. (2015). The impact of classroom design on pupils' learning: Final results of a holistic, multi-level analysis. *Building and Environment*, *89*, 118-133. https://doi.org/10.1016/j.buildenv.2015.02.013

Barrett, P., Davies, F., Zhang, Y., & Barrett, L. (2016). The holistic impact of classroom spaces on learning in specific subjects. *Environment and Behavior*, *49*(4), 425-451. https://doi.org/10.1177/0013916516648735

Bluteau, J., Aubenas, S., & Dufour, F. (2022). Influence of flexible classroom seating on the wellbeing and mental health of upper elementary school students: A gender analysis. *Frontiers in Psychology*, *13*, Article 821227. https://doi.org/10.3389/fpsyg.2022.821227

Bøjer, B. (2019). *Unlocking learning spaces: An examination of the interplay between the design of learning spaces a nd pedagogical practices.* Rune Fjord Studio & The Royal Danish Academy of Fine Arts. https://adk.elsevierpure.com/ws/portalfiles/portal/63453689/Bodil_Bojer_PhDthesis_Unlocking_Learning_Spaces.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Brown, J., McLennan, C., Mercieca, D., Mercieca, D. P., Robertson, D. P., & Valentine, E. (2021). Technology as thirdspace: Teachers in Scottish schools engaging with and being challenged by digital technology in first COVID-19 lockdown. *Education Sciences*, *11*(3), Article 136. https://doi.org/10.3390/educsci11030136

Campbell, M., Saltmarsh, S., Chapman, A., & Drew, C. (2013). Issues of teacher professional learning within 'non-traditional' classroom environments. *Improving Schools*, *16*(3), 209-222. https://doi.org/10.1177/1365480213501057

Cardellino, P., & Woolner, P. (2019). Designing for transformation – a case study of open learning spaces and educational change. *Pedagogy, Culture & Society*, *28*(3), 383-402. https://doi.org/10.1080/14681366.2019.1649297

Carvalho, L., & Goodyear, P. (2014). *The architecture of productive learning networks* (1st ed.). Routledge. https://doi.org/10.4324/9780203591093

Carvalho, L., Nicholson, T., Yeoman, P., & Thibaut, P. (2020). Space matters: Framing the New Zealand learning landscape. *Learning Environments Research*, *23*(3), 307-329. https://doi.org/10.1007/s10984-020-09311-4

Cleveland, B., & Fisher, K. (2014). The evaluation of physical learning environments: A critical review of the literature. *Learning Environments Research*, *17*(1), 1-28. https://doi.org/10.1007/s10984-013-9149-3

Dash, S. P., & Thilagam, N. L. (2022). A systematic review on inter-relationship of residential neighborhood characteristics on quality of life of elderly. *Pertanika Journal of Social Sciences and Humanities*, *30*(4), 1533-1566. https://doi.org/10.47836/pjssh.30.4.05

Denyer, D., & Tranfield, D. (2006). Using qualitative research synthesis to build an actionable knowledge base. *Management Decision*, *44*(2), 213-227. https://doi.org/10.1108/00251740610650201

Deppeler, J., Corrigan, D., Macaulay, L., & Aikens, K. (2022). Innovation and risk in an innovative learning environment: A private public partnership in Australia. *European Educational Research Journal*, *21*(4), 602-626. https://doi.org/10.1177/14749041211030400

Dolan, P., Sturm, B., & Wollmuth, C. (2006). Prairie crossing charter school: Comfort as a principal component of high performance school design. *Journal of Green Building*, *1*(3), 17-25. https://doi.org/10.3992/jgb.1.3.17

Dovey, K., & Fisher, K. (2014). Designing for adaptation: The school as socio-spatial assemblage. *The Journal of Architecture*, *19*(1), 43-63. https://doi.org/10.1080/13602365.2014.882376

Furió, D., Juan, M. C., Seguí, I., & Vivó, R. (2015). Mobile learning vs. traditional classroom lessons: A comparative study. *Journal of Computer Assisted Learning*, *31*(3), 189-201. https://doi.org/10.1111/jcal.12071

Ghadwan, A. S., Ahmad, W. M. W., & Hanifa, M. H. (2022). Financial planning for retirement models: An integrative systematic review. *Pertanika Journal of Social Sciences and Humanities*, *30*(2), 879-900. https://doi.org/10.47836/pjssh.30.2.24

Gislason, N. (2010). Architectural design and the learning environment: A framework for school design research. *Learning Environments Research*, *13*(2), 127-145. https://doi.org/10.1007/s10984-010-9071-x

Grannäs, J., & Stavem, S. M. (2021). Transitions through remodelling teaching and learning environments. *Education Inquiry*, *12*(3), 266-281. https://doi.org/10.1080/20004508.2020.1856564

Gremmen, M. C., van den Berg, Y. H. M., Segers, E., & Cillessen, A. H. N. (2016). Considerations for classroom seating arrangements and the role of teacher characteristics and beliefs. *Social*

Psychology of Education, *19*(4), 749-774. https://doi.org/10.1007/s11218-016-9353-y

Gultekin, M., & Ira, G. O. (2022). Pedagogy-driven design fundamentals of 21st century primary schools' physical learning environments. *Journal of Education and Future,* (21), 99-110. https://doi.org/10.30786/jef.805905

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*(2), 181-217. https://doi.org/10.1002/jrsm.1378

Harouni, H. (2013). Lived-in room: Classroom space as teacher. *Berkeley Review of Education*, *4*(2). https://doi.org/10.5070/b84110012

Herman, F., & Tondeur, J. (2021). Untangling the sociomateriality of the classroom: Biographies of school spaces (c. 1960–2014). *Oxford Review of Education*, *47*(5), 681-695. https://doi.org/10.1080/03054985.2021.1924654

Hoon, L. N., & Shaharuddin, S. S. (2019). Learning effectiveness of 3D hologram animation on primary school learners. *Journal of Visual Art and Design*, *11*(2), 93-104. https://doi.org/10.5614/j.vad.2019.11.2.2

Ingold, T. (2011). *Being alive: Essays on movement, knowledge and description.* Routledge. https://doi.org/10.4324/9780203818336

Jagust, T., Boticki, I., & So, H. J. (2018). A review of research on bridging the gap between formal and informal learning with technology in primary school contexts. *Journal of Computer Assisted Learning*, *34*(4), 417-428. https://doi.org/10.1111/jcal.12252

James, S. (2011). *Independent report: Review of education capital* (Report No. DFE-00073-2011). U.K. Department for Education. https://www.gov.uk/government/publications/review-of-education-capital

Johler, M., Krumsvik, R. J., Bugge, H. E., & Helgevold, N. (2022). Teachers' perceptions of

their role and classroom management practices in a technology rich primary school classroom. *Frontiers in Education*, *7*, Article 841385. https://doi.org/10.3389/feduc.2022.841385

Joint Information Systems Committee. (2006). *Designing spaces for effective learning: A guide to 21st century learning space design*. https://www.d41.org/cms/lib/IL01904672/Centricity/Domain/422/learningspaces.pdf

Kariippanon, K. E., Cliff, D. P., Lancaster, S. L., Okely, A. D., & Parrish, A.-M. (2017). Perceived interplay between flexible learning spaces and teaching, learning and student wellbeing. *Learning Environments Research*, *21*(3), 301-320. https://doi.org/10.1007/s10984-017-9254-9

Killeen, J. P., Evans, G. W., & Danko, S. (2003). The role of permanent student artwork in students sense of ownership in an elementary school. *Environment and Behavior*, *35*(2), 250-263. https://doi.org/10.1177/0013916502250133

Lee, K. S., Kim, H. J., & Kang, J. (2019). From uniformity to sustainable diversity: Exploring the design attributes of renovating standardized classrooms in Korea. *Sustainability*, *11*(20), Article 5669. https://doi.org/10.3390/su11205669

Lee, S. (2019). Restructuring types of classroom space. *Journal of Korea Intitute of Spatial Design*, *14*(5), 53-64. https://doi.org/10.35216/kisd.2019.14.5.53

Li, P. P., Locke, J., Nair, P., & Bunting, A. (2005). Creating 21st century learning environments. In *PEB Exchange, Programme on Educational Building* (Vol. 10; pp. 15-26)*.* OECD iLibrary. https://doi.org/10.1787/558676471016

Liu, Q. Y., Ang, L. H., Waheed, M., & Kasim, Z. M. (2022). Appraisal theory in translation studies—a systematic literature review. *Pertanika Journal of Social Sciences and Humanities*, *30*(4), 1589-1605. https://doi.org/10.47836/pjssh.30.4.07

López-Chao, V., Amado Lorenzo, A., Saorín, J. L., De La Torre-Cantero, J., & Melián-Díaz, D. (2020).

Classroom indoor environment assessment through architectural analysis for the design of efficient schools. *Sustainability*, *12*(5), Article 2020. https://doi.org/10.3390/su12052020

Mackey, J., O'Reilly, N., Jansen, C., & Fletcher, J. (2018). Leading change to co-teaching in primary schools: A "down under" experience. *Educational Review*, *70*(4), 465-485. https://doi.org/10.1080/00131911.2017.1345859

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, *12*(4), 1160-1177. https://doi.org/10.1016/j.joi.2018.09.002

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336-341. https://doi.org/10.1016/j.ijsu.2010.02.007

Mokhtarmanesh, S., & Ghomeishi, M. (2019). Participatory design for a sustainable environment: Integrating school design using students' preferences. *Sustainable Cities and Society*, *51*, Article 101762. https://doi.org/10.1016/j.scs.2019.101762

Mulcahy, D., & Morrison, C. (2017). Re/assembling 'innovative' learning environments: Affective practice and its politics. *Educational Philosophy and Theory*, *49*(8), 749-758. https://doi.org/10.1080/00131857.2016.1278354

National Heart, Lung, and Blood Institute. (2021). *Study quality assessment tools: Quality assessment tool for observational cohort and cross-sectional studies.* https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Niemi, K., Minkkinen, J., & Poikkeus, A.-M. (2022). Opening up learning environments: Liking school among students in reformed learning spaces. *Educational Review*, 1-18. https://doi.org/10.1080/00131911.2022.2098927

Nyabando, T., & Evanshen, P. (2022). Second grade students' perspectives of their classrooms'

physical learning environment: A multiple case study. *Early Childhood Education Journal*, *50*(5), 709-720. https://doi.org/10.1007/s10643-021-01183-4

Oliveras-Ortiz, Y., Bouillion, D. E., & Asbury, L. (2021). Learning spaces matter: Student engagement in new learning environments. *Journal of Education*, *201*(3), 174-182. https://doi.org/10.1177/0022057420908062

Osman, K., Ahmad, C. N. C., & Halim, L. (2011). Students' perception of the physical and psychosocial science laboratory environment in Malaysia: Comparison across subject and school location. *Procedia-Social and Behavioral Sciences*, *15*, 1650-1655. https://doi.org/10.1016/j.sbspro.2011.03.347

Reh, S., Rabenstein, K., & Fritzsche, B. (2011). Learning spaces without boundaries? Territories, power and how schools regulate learning. *Social & Cultural Geography*, *12*(1), 83-98. https://doi.org/10.1080/14649365.2011.542482

Reinius, H., Korhonen, T., & Hakkarainen, K. (2021). The design of learning spaces matters: Perceived impact of the deskless school on learning and teaching. *Learning Environments Research*, *24*(3), 339-354. https://doi.org/10.1007/s10984-020-09345-8

Rönnlund, M., Bergström, P., & Tieva, Å. (2021). Tradition and innovation. Representations of a "good" learning environment among Swedish stakeholders involved in planning, (re)construction and renovation of school buildings. *Education Inquiry*, *12*(3), 249-265. https://doi.org/10.1080/20004508.2020.1774239

Saltmarsh, S., Chapman, A., Campbell, M., & Drew, C. (2015). Putting "structure within the space": Spatially un/responsive pedagogic practices in open-plan learning environments. *Educational Review*, *67*(3), 315-327. https://doi.org/10.1080/00131911.2014.924482

Sato, H., & Bradley, J. S. (2008). Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *The Journal of the Acoustical Society of*

America, *123*(4), 2064-2077. https://doi.org/10.1121/1.2839283

Shaffril, H. A. M., Ahmad, N., Samsuddin, S. F., Samah, A. A., & Hamdan, M. E. (2020). Systematic literature review on adaptation towards climate change impacts among indigenous people in the Asia Pacific regions. *Journal of Cleaner Production*, *258*, Article 120595. https://doi.org/10.1016/j.jclepro.2020.120595

Sigurdardottir, A. K., & Hjartarson, T. (2016). The idea and reality of an innovative school: From inventive design to established practice in a new school building. *Improving Schools*, *19*(1), 62-79. https://doi.org/10.1177/1365480215612173

Starkey, L., Leggett, V., Anslow, C., & Ackley, A. (2021). The use of furniture in a student-centred primary school learning environment. *New Zealand Journal of Educational Studies*, *56*, 61-79. https://doi.org/10.1007/s40841-020-00187-9

Swartz, A. M., Tokarek, N. R., Strath, S. J., Lisdahl, K. M., & Cho, C. C. (2020). Attentiveness and fidgeting while using a stand-biased desk in elementary school children. *International Journal of Environmental Research and Public Health*, *17*(11), Article 3976. https://doi.org/10.3390/ijerph17113976

Szpytma, C., & Szpytma, M. (2022). School architecture for primary education in a post-socialist country: A case study of Poland. *Compare: A Journal of Comparative and International Education*, *52*(4), 519-542. https://doi.org/10.1080/03057925.2020.1777843

To, P. T., & Grierson, D. (2019). An application of measuring visual and non-visual sensorial experiences of nature for children within primary school spaces Child-nature-distance case studies in Glasgow, Scotland. *Archnet-IJAR*, *14*(2), 167-186. https://doi.org/10.1108/ARCH-05-2019-0139

Tokarek, N. R., Cho, C. C., Strath, S. J., & Swartz, A. M. (2022). The impact of stand-biased desks on afterschool physical activity behaviors of elementary school children. *International Journal of Environmental Research and Public*

Health, *19*(13), Article 7689. https://doi.org/10.3390/ijerph19137689

Uduku, O. (2015). Spaces for 21st-century learning. In *Routledge Handbook of International Education and Development* (pp. 196-209). Routledge.

Vijapur, D., Candido, C., Göçer, Ö., & Wyver, S. (2021). A ten-year review of primary school flexible learning environments: Interior design and IEQ performance. *Buildings, 11*(5), Article 183. https://doi.org/10.3390/buildings11050183

Wallace, A. L., Swartz, A. M., Cho, C. C., Kaiver, C. M., Sullivan, R. M., & Lisdahl, K. M. (2022). Stand-biased desks impact on cognition in elementary students using a within-classroom crossover design. *International Journal of Environmental Research and Public Health*, *19*(9), Article 5684. https://doi.org/10.3390/ijerph19095684

Wang, X., & Wang, T. (2020). The mutability of pedagogical practice and space use: A case study of collaborative learning and classroom space in a Chinese primary school. *Compare: A Journal of Comparative and International Education*, *52*(5), 729-747. https://doi.org/10.1080/03057925.2020.1811640

Whitehouse, D. (2009). Designing learning spaces that work: A case for the importance of history. *History of Education Review*, *38*, 94-108. https://doi.org/10.1108/08198691200900016

Woolner, P. (2014). *School design together*. Routledge. https://doi.org/10.4324/9781315774107

World Economic Forum. (2020). *Schools of the future: Defining new models of education for the fourth industrial revolution.* https://www3.weforum.org/docs/WEF_Schools_of_the_Future_Report_2019.pdf

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, *39*(1), 93-112. https://doi.org/10.1177/0739456X17723971

Yao, F., Bo, G., & Zitong, W. (2024). Association between visual distance and students' visual perception comfort in primary and secondary

classrooms. *Journal of Asian Architecture and Building Engineering*, *23*(2), 569-581. https://doi.org/10.1080/13467581.2023.2245019

Yeoman, P. (2018). The material correspondence of learning. In R. A. Ellis & P. Goodyear (Eds.), *Spaces of teaching and learning: Integrating perspectives on research and practice* (pp. 81-103). Springer Singapore. https://doi.org/10.1007/978-981-10-7155-3_6

Zairul, M., Azli, M., & Azlan, A. (2023). Defying tradition or maintaining the status quo? Moving towards a new hybrid architecture studio education to support blended learning post-COVID-19. *Archnet-IJAR*, *17*(3), 554-573. https://doi.org/10.1108/ARCH-11-2022-0251

Zhang, D., Ortiz, M. A., & Bluyssen, P. M. (2019). Clustering of Dutch school children based on their preferences and needs of the IEQ in classrooms. *Building and Environment*, *147*, 258-266. https://doi.org/10.1016/j.buildenv.2018.10.014

SOCIAL SCIENCES & HUMANITIES

Review Article

A Systematic Literature Review: The Efficacy of Cognitive Behavioral Therapy in Addressing Stress Among University Students

Nurul Shuhada Sharif Mudin*, Aslina Ahmad, Pau Kee, Muhammad Najib Husman @ Othman and Suhaila Abdul Wahab

Department of Psychology and Counselling, Faculty of Human Development, Universiti Pendidikan Sultan Idris, 35900 Tanjung Malim, Perak, Malaysia

ABSTRACT

Despite the growing recognition of mental health challenges among university students, the effectiveness and adaptability of Cognitive Behavioral Therapy (CBT) in these settings, particularly in terms of diverse delivery methods and their impact on outcomes, remains underexplored. This Systematic Literature Review (SLR) rigorously evaluates the efficacy of CBT in reducing stress among university students, utilizing the PRISMA 2020 framework. The review analyzes 27 carefully selected articles, focusing on objective analysis, diverse delivery methods, session and treatment duration, various CBT interventions, and the applied skills and strategies. The findings confirm CBT as an exceptionally effective stress management tool, noted for its adaptability and versatility. The study underscores the importance of online CBT and emphasizes the need for interventions tailored to the diverse needs of students, enhancing their mental well-being. In light of these findings, future research should investigate the mediators and moderators that affect the effectiveness of CBT, assess the long-term impact on academic and mental health outcomes, and explore the sustainability of these interventions. This study contributes to CBT's theoretical and practical understanding in higher education settings. It lays the groundwork for future explorations to optimize mental health interventions for university students.

Keywords: Cognitive Behavioral Therapy (CBT), stress, systematic literature review (SLR), university students

INTRODUCTION

Stress profoundly influences individuals, particularly university students, precipitating various emotional and behavioral challenges (Joompathong et al., 2022; Kaplan et al., 2023). While empirical evidence underscores the correlation between heightened stress levels and subsequent emotional and behavioral disturbances, including psychological distress and potential psychopathologies (Kaplan et al., 2023; Modrego-Alarcón et al., 2021), a critical evaluation of existing literature is essential to justify the necessity of conducting a systematic literature review (SLR).

University life presents a milieu replete with incessant demands and challenges, significantly shaping education interpretation among early adult students (Ahmad & Meriç, 2021; Pasaribu & Zarfiel, 2019). Pasaribu and Zarfiel (2019) highlight early adult students' challenges in navigating educational demands. Moreover, the stressors encapsulated within academic, familial, financial, and personal domains collectively contribute to the stress experienced during this developmental phase (Cozzolino et al., 2020). However, while existing literature elucidates the multifaceted nature of stress among university students, a critical analysis of its underlying assumptions and methodological rigor is warranted.

Psychological distress can lead to conditional self-acceptance thinking in individuals (Pasaribu & Zarfiel, 2019). Stress, which can produce negative emotions and thoughts, is not confined to the psychological realm but can also lead to various physical health problems like heart diseases, hypertension, and lung diseases (Ashraf et al., 2020; Murad, 2021). Recognizing the importance of evidence-based practices, universities must effectively address student mental health needs (Newcombe et al., 2023).

Given the detrimental impact of stress on students' academic performance, mental health, and overall well-being, it is crucial to identify effective interventions. Cognitive behavioral therapy (CBT) has emerged as a promising approach to managing stress and promoting coping skills among university students (Shearin & Brewer-Mixon, 2020). CBT, training, and combined psychological and study skills interventions can effectively modify negative emotional effects (Hamdan et al., 2021). Implementing preventive interventions, encouraging unconditional self-acceptance, and providing accessible mental health services on campuses are vital steps universities must take to support the well-being of their students (Newcombe et al., 2023; Pasaribu & Zarfiel, 2019).

This systematic literature review (SLR) aims to assess the efficacy of CBT interventions in reducing stress symptoms and improving coping mechanisms, specifically within the university student population. By understanding the impact of stress on students and offering appropriate interventions, universities can play a crucial role in helping students cope better with stress and improve their overall mental health. The systematic literature

review (SLR) is a structured approach that thoroughly identifies and integrates relevant research, employing organized, transparent, and reproducible procedures at every stage (Higgins et al., 2011). This study's SLR specifically utilizes an integrative review analysis combining quantitative, qualitative, and mixed methods (Shaffril et al., 2021). The proposed SLR on the efficacy of CBT in addressing stress among university students is distinct from existing SLRs, as it focuses specifically on this unique population. While there may be previous SLRs on CBT and stress management, this review aims to provide targeted insights and recommendations tailored to university students' challenges.

The primary research question of this SLR is: *"What is the efficacy of cognitive behavioral therapy in addressing stress among university students?"* The objectives include identifying relevant peer-reviewed articles on CBT interventions for stress among university students, analyzing the effectiveness of CBT in reducing stress symptoms and improving coping skills, and assessing its impact on academic performance, mental well-being, and overall quality of life. The review will also explore variations in treatment protocols and outcomes across different studies, identify research gaps, and provide evidence-based insights and recommendations for university counseling services and mental health professionals.

METHODOLOGY

Review Protocol PRISMA

An SLR methodology was the most suitable approach to meet the research objectives. This method ensures a comprehensive and systematic review of the literature on CBT's effectiveness in reducing academic stress among university students, facilitating a thorough collection and critical analysis of data. It aligns directly with assessing the literature and identifying evidence-based interventions that can inform practice and policy.

Following this justification, the current investigation emerged from using the PRISMA 2020 approach (Page et al., 2021) for conducting a review. The PRISMA model, an acronym for Preferred Reporting Items for Systematic Reviews and Meta-Analyses furnishes guidelines to ensure the thorough and transparent execution of systematic reviews and meta-analyses. PRISMA's central objective is to aid authors in achieving clear and comprehensive reporting of their systematic reviews and meta-analyses (Liberati et al., 2009). The researchers initiated their systematic literature review (SLR) following the PRISMA guidelines, starting with formulating relevant research queries. Subsequently, the authors elaborate on the three fundamental stages of the systematic search process: identification, screening involving inclusion and exclusion criteria, and determining eligibility. The procedure of obtaining, organizing, and analyzing data is then detailed, emphasizing its readiness for integration into the research endeavor.

Furthermore, the PRISMA checklist was developed to establish protocols for identifying, selecting, assessing, and summarizing research in social sciences and education (Page et al., 2021). Within the realm of social science inquiries, the application of the PRISMA literature review methodology holds several advantages. According to Okoli's (2015) research, these benefits encompass the capacity for all researchers to access data from credible sources, the strategic use of keywords facilitating effective search result filtering for relevant articles, and the potential for researchers to employ collected data to craft foundational studies addressing identified gaps. Moreover, PRISMA is a valuable guideline for highlighting literature, yielding more systematic and detailed outcomes (Page et al., 2021). The PRISMA guidelines outline four processes: (1) identification, (2) screening, (3) eligibility, and (4) inclusion (Shaffril et al., 2020).

Formulation of the Research Questions

The study's research subject was formulated using the PICo method. Researchers can construct a robust research topic for their literature review by employing the PICo approach. The foundational elements uphold PICo are the population problem, interest, and context. Based on these concepts, the study's authors concentrated on three key areas: university students (the study's population), stress (the interest), and the effectiveness of Cognitive Behavioral Therapy (CBT; the context).

"What is the effectiveness of Cognitive Behavioral Therapy (CBT) in reducing academic stress among university students?"

These questions led to the authors' major study topic. This research question evaluates CBT's efficacy in addressing academic stress and its impact on university students. It prompts the systematic review to analyze and synthesize existing literature to provide a comprehensive understanding of the effectiveness of CBT interventions in this context.

Systematic Searching Strategies

The systematic searching methods procedure consists of four primary steps: identification, screening, eligibility, and inclusion.

Identification

The identification process begins with recognizing keywords and subsequent endeavors to identify analogous or identical terms using thesauri, dictionaries, and preexisting research. This procedure involves locating akin words, expressions, and keywords, collectively constructing the principal means of entry with various lexical variations. The aim is to furnish specific databases with supplementary avenues for unearthing and assessing pertinent material (Shaffril et al., 2020). Nonetheless, frequently generating keywords responds to research inquiries (Okoli, 2015).

For this study, the authors employed an online thesaurus, phrases from preceding research, keywords derived from data

analysis, and keywords provided by subject experts to discover synonyms (Shaffril et al., 2020). Another strategy for enhancing keyword richness involves the utilization of Boolean operators, truncation, phrase inquiries, manipulations of database field codes, and wildcard functions, thereby unlocking the complete potential of a database search string (Table 1). The applicable combination of keywords that have been used, such as ("university students" OR "college student" OR "undergraduate student"). Boolean operators, truncation and field code were employed to improve current phrases and generate a comprehensive search string. Moreover, the Scopus database yielded 99 relevant articles as a result of the search. The results from Science Direct are approximately 16272, Google Scholar is 175, and Wiley is 720. A manual search was also conducted in the Proquest database using the keywords "efficacy," "Cognitive Behavioral Therapy," "stress," and "university students." As a result, a total of 100,047 relevant articles were obtained. The current research successfully obtained 117313 articles from five databases during the first part of the systematic review procedure.

Table 1

The search string utilized for the systematic review process

Database	Search string
Scopus	TITLE-ABS-KEY(("stress" OR "depressed" OR "distress" OR "academic stress") AND ("effectiveness" OR "efficiency" OR "efficacy") AND ("cognitive behavioral therapy" OR "cognitive behavioural therapy" OR "cognitive therapy") AND ("university students" OR "college student" OR "undergraduate student" OR "higher education student"))
Science Direct	(("cognitive behavioral therapy" OR "cognitive therapy") AND ("stress" OR "distress" OR "academic pressure") AND ("university students" OR "college student" OR "undergraduate student") AND ("effectiveness"))
Google Scholar	allintitle: stress "cognitive behavioral therapy" OR "cognitive behavioural therapy" OR effectiveness OR efficiency OR efficacy "university students" OR "college student" OR "undergraduate student"
Wiley	(("cognitive behavioral therapy" OR "cognitive therapy") AND ("stress" OR "distress" OR "academic pressure") AND ("university students" OR "college student" OR "undergraduate student") AND ("effectiveness"))

Source: Authors' work

Screening

In this research, articles underwent an automated filtration process by applying predetermined standards for article inclusion, leveraging the database's sorting functionality. As Kitchenham and Charters (2007) proposed, the research subject should underpin the basis for these selection

criteria. Okoli (2015) recommended that researchers strategically choose specific timeframes for their reviews due to the pragmatic challenges academics face in attempting to appraise the entirety of presently accessible published material.

Based on the standards researchers gave, a database sorting function automatically identified all 1002 papers in this investigation. The authors have established the exclusion and inclusion criteria. First, only an article journal containing empirical data is selected, meaning a thesis, proceeding procedure, books, or conference are all rejected. Second, the search attempts ruled out non-English publishing and concentrated only on publications published in English to minimize misunderstanding and problems translating. Third, regarding timeliness, five years (between 2019 and 2023) has been chosen as an appropriate time to observe the evolution of research and related articles. Table 2 lists the criteria for exclusion and inclusion:

Table 2
The inclusion and exclusion criteria

Criteria	Inclusion	Exclusion
Timeline	2019–2023	<2019
Document type	Article journal (research article)	Article review, chapter in the book, book series, chapters in the book, series book, conference proceeding, thesis
Language	English	Non-English
Geographical areas	All around the world	-
Subject area	Psychology. Social Sciences	Besides Psychology. Social Sciences

Source: Authors' work

This process reduced one duplicated article. It excluded 116311 articles (Scopus: 58, Science Direct: 16213, Google Scholar: 110 and Wiley: 718, Proquest: 99,212) that did not satisfy the inclusion criteria. The other 1002 articles (Scopus: 41, Science Direct: 59, Google Scholar: 65 and Wiley, 2; Proquest, 835) were employed in the third process, entitled eligibility.

Eligibility

In determining eligibility, the 1002 articles that successfully cleared the initial screening phase underwent a meticulous evaluation by the researchers to ascertain their alignment with the predetermined criteria. The authors rigorously examined the titles and abstracts of these articles, identifying 59 articles that were by the study's research objectives. The exclusion of 943 articles stemmed from their deficiency in addressing the effectiveness

of Cognitive Behavioral Therapy (CBT) on university students and their failure to concentrate on academic stress within this demographic. Figure 1 presents a visual representation outlining the comprehensive process of article selection.
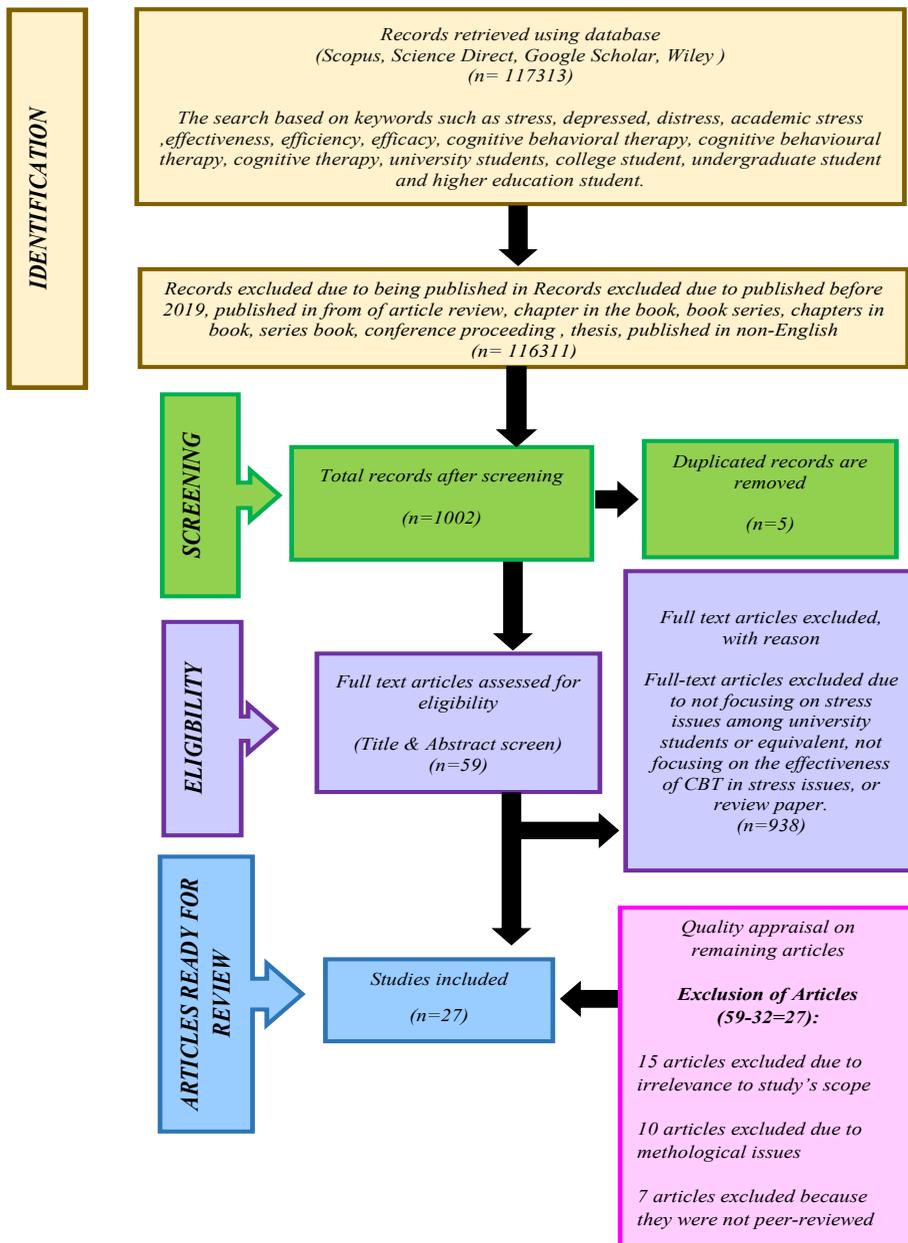

Figure 1. Flow diagram of the search process

Source: Authors' work

Quality Appraisal

The quality assessment phase was designed to verify the methodological and analytical rigor of the studies selected for our review. We adopted a systematic literature review approach following the Kitchenham guidelines (Asl et al., 2022; Kitchenham et al., 2009). Our main objective was to gather evidence on researchers' practices in secondary and tertiary educational settings, focusing on tertiary literature reviews.

Employing Barbara Kitchenham's framework, which outlines a systematic process for conducting comprehensive reviews, the corresponding author and two co-authors closely examined each article's methodology and analyses. We applied five assessment criteria, accepting responses of "yes," "no," or "don't know/can't tell."

Articles meeting at least three of these criteria were included in our review, and all decisions were made by consensus.

As shown in Figure 1, we initially considered 59 articles. However, we excluded 32 based on the following criteria: 15 for being outside the study's scope, 10 for methodological flaws, and seven because they lacked peer review. It ensured that the 27 remaining articles met our stringent standards for inclusion.

Additionally, by following the procedures exemplified in Table 3, as Abouzahra et al. (2020) recommended, we ensured that the study preserved and adequately addressed all five identified themes or topics throughout the review process.

Table 3

The measures used to evaluate the effectiveness of the methodology and analysis in the selected articles

Item	Question
QA1	Is the purpose of the study clearly stated?
QA2	Is the interest and the usefulness of the work clearly presented?
QA3	Is the study methodology clearly established?
QA4	Are the concepts of approach clearly established?
QA5	Is the work compared and measured with other similar work?
QA6	Are the limitations of the work clearly mentioned?

Note. The scoring procedures used to evaluate each QA was: Yes (Y) = 1, Partly (P) = 0.5 or No (N) = 0
Source: Authors' work

The current research investigates the research inquiries, reaching a consensus while promptly resolving any differences through author discussions. Following this approach, all authors unanimously determined that the selected articles met the minimum quality standards for methodology and analysis. In entirety, 11 articles satisfied all criteria. In comparison, 11 met a minimum of five criteria, and an additional four achieved at least four (Table 4).

Table 4
The inclusion and exclusion criteria

Study	Research design	QA1	QA2	QA3	QA4	QA5	QA6	Number of criteria fulfilled	Inclusion in the review
Hanani et al. (2022)	QN	0.5	1	1	1	1	1	6/6	✓
Cozzolino et al. (2020)	QN	1	1	1	1	1	1	6/6	✓
Newcombe et al. (2023)	QN	0.5	1	1	1	1	1	5.5/6	✓
El Morr et al. (2020)	QN	1	1	1	1	1	1	6/6	✓
Tomoiaga and David (2022)	QN	1	1	1	1	1	1	6/6	✓
Rackoff et al. (2022)	QN	1	1	1	1	1	1	6/6	✓
Sharma et al. (2022)	QN	0.5	1	1	1	1	1	5.5/6	✓
Hwang et al. (2022)	QN	0	1	1	0.5	1	1	4.5/6	✓
Rodriguez et al. (2021)	QN	0	1	1	1	0.5	1	4.5/6	✓
Ashraf et al. (2020)	QN	1	1	1	1	1	1	6/6	✓
Azevedo and Menezes (2020)	QN	1	1	1	1	1	1	6/6	✓
Mohamed (2022)	QN	1	1	1	1	0	0	4/6	✓
Modrego-Alarcón et al. (2021)	QN	0.5	1	1	0.5	1	1	5/6	✓
Joompathong et al. (2022)	QN	0.5	1	1	0.5	1	1	5/6	✓
Murad (2021)	QN	1	1	1	1	0.5	0	4.5/6	✓
Pasaribu and Zarfiel (2019)	QL	1	1	1	1	1	0.5	5.5/6	✓
Ritvo et al. (2021)	QN	1	1	1	0.5	1	0.5	5/6	✓
Recabarren et al. (2019)	QN	0.5	1	1	0.5	1	1	5/6	✓

Table 4 *(Continue)*

Study	Research design	QA1	QA2	QA3	QA4	QA5	QA6	Number of criteria fulfilled	Inclusion in the review
Shearin and Brewer-Mixon (2020)	QN	0	1	1	1	1	1	5/6	✓
Lee et al. (2019)	QN	0.5	1	1	0	1	0.5	4/6	✓
Kaplan et al. (2023)	QN	1	1	1	0.5	1	0.5	5/6	✓
Hamdan et al. (2021)	QN	1	1	1	1	1	1	6/6	✓
Changklang and Ranteh (2023)	QN	1	1	1	1	1	1	6/6	✓
Shahrokhian et al. (2022)	QN	0.5	1	1	0.5	1	1	5/6	✓
Huberty et al. (2019)	QN	1	1	1	1	1	2	6/6	✓
Agah et al. (2023)	QN	1	1	1	1	1	1	6/6	✓
Eneogu et al. (2023)	QN	0.5	1	1	1	1	1	5.5/6	✓

Note. Quality assessment; QN (DC) = Quantitative descriptive; QN (NR) = Quantitative non-randomized; QL = Qualitative; MX = Mixed-Method; C = Cannot tell

Source: Authors' work

Data Extraction and Analysis

The articles underwent thematic analysis to manage the diverse research designs in the review and effectively blend these variations using qualitative synthesis methods (Whittemore & Knafl, 2005). Among the available qualitative synthesis approaches, this study adopted the method recommended by Flemming et al. (2019), which emphasized the adaptability of thematic synthesis for merging data from different research designs. The thematic analysis involves recognizing patterns in existing studies by revealing shared characteristics or connections within the accessible data (Braun & Clarke, 2019).

During this review, the process of thematic synthesis adhered to the steps suggested by Kiger and Varpio (2020). Initially, researchers acquainted themselves with the complete dataset through repeated readings, gaining crucial insights into the raw data that laid the foundation for subsequent stages. The subsequent step

involved crafting initial codes, where data was organized at a detailed level. The researchers scrutinized selected articles in this phase, extracting relevant data tied to the main research question. Following this, themes were developed as researchers employed inductive coding techniques to identify links and resemblances among the coded data. The synthesis procedure leaned on an inductive coding framework, leading to the emergence of themes from the coded data. These themes were intertwined with the original data and encapsulated the entire dataset (Braun & Clarke, 2019).

During our study's data extraction and analysis phase, the research team undertook a comprehensive evaluation of all the themes and sub-themes that emerged during the initial coding phase. A structured approach to thematic synthesis guided this critical step. Our objective was to rigorously assess the relevance and contribution of each theme towards fulfilling the research goals. This systematic evaluation ensured that only the most pertinent themes were selected for further analysis, providing a robust foundation for the synthesis process.

We decided to group the themes into five primary categories to consolidate the vast data into a coherent structure. This decision was based on several key criteria: thematic similarity, which ensures consistency and cohesiveness among grouped themes; direct relevance to the core questions of our research, which ensures that each theme significantly contributes to addressing the overarching research questions; and the potential of the themes to collectively provide a comprehensive overview of the efficacy of Cognitive Behavioral Therapy (CBT).

The primary themes defined were as follows:

1. Objective analysis: This theme focused on the measurable outcomes derived from various CBT interventions, providing a quantitative foundation for assessing the effectiveness of these treatments.

2. Delivery method: This encompassed a range of mechanisms through which CBT is delivered. A noteworthy aspect of this theme was the 'Effectiveness of online CBT for stress,' chosen for its growing significance in contemporary research and practice settings, reflecting the shift towards digital health solutions.

3. Session/Treatment duration: This theme captured the variations in the duration and frequency of CBT sessions, which are crucial factors affecting the treatment's outcomes. Understanding these variables helps determine the optimal session parameters for effective stress management.

4. Diversity of CBT interventions for stress: This theme was particularly broad, encompassing a variety of CBT techniques used in managing stress. It includes detailed

explorations of specific programs, psychoeducational components, and their use alongside other therapeutic methods, illustrating the versatile application of CBT in stress management.

5. CBT skills/strategies for stress: Synthesizing the specific strategies and skills imparted through CBT, this theme focuses on the practical applications of CBT skills that are vital for effectively managing stress.

Following this, the researchers presented the final array of themes and sub-themes to two experts who were well-versed in qualitative synthesis and community development. These experts were asked to validate the themes and sub-themes, confirming their relevance and alignment with the research inquiries.

RESULTS

Background of the Selected Studies

Out of the 27 articles included in the study, three papers focused on research conducted in the United States (Huberty et al., 2019; Sharma et al., 2022; Shearin & Brewer-Mixon, 2020), and three articles centered their studies in Canada (El Morr et al., 2020; Newcombe et al., 2023; Ritvo et al., 2021). Additionally, two papers focused on Thailand (Changklang & Ranteh, 2023; Joompathong et al., 2022) and Nigeria (Agah et al., 2023; Eneogu et al., 2023). Meanwhile, each research study also focused on Romania (Tomoiaga & David,

2022), California (Rackoff et al., 2022), Republic of Korea (Hwang et al., 2022), Brazil (Azevedo & Menezes, 2020), Spain (Modrego-Alarcón et al., 2021), Egypt (Mohamed, 2022), Saudi Arabia (Hamdan et al., 2021), Jordan (Murad, 2021), Indonesia (Pasaribu & Zarfiel, 2019), Turkey (Kaplan et al., 2023), Switzerland (Recabarren et al., 2019), Taiwan (Lee et al., 2019), Italy (Cozzolino et al., 2020), Palestine (Hanani et al., 2022), China (Rodriguez et al., 2021), Pakistan (Ashraf et al., 2020), and a study conducted in the Iran (Shahrokhian et al., 2022; Figure 2).

It has been documented that 26 studies were centered around quantitative analyses (Agah et al., 2023; Ashraf et al., 2020; Azevedo & Menezes, 2020; Changklang & Ranteh, 2023; Cozzolino et al., 2020; El Morr et al., 2020; Eneogu et al., 2023; Hamdan et al., 2021; Hanani et al., 2022; Huberty et al., 2019; Hwang et al., 2022; Joompathong et al., 2022; Kaplan et al., 2023; Lee et al., 2019; Modrego-Alarcón et al., 2021; Mohamed, 2022; Murad, 2021; Newcombe et al., 2023; Rackoff et al., 2022; Recabarren et al., 2019; Ritvo et al., 2021; Rodriguez et al., 2021; Shahrokhian et al., 2022; Sharma et al., 2022; Shearin & Brewer-Mixon, 2020; Tomoiaga & David, 2022), while only one study was centered around qualitative analyses (Pasaribu & Zarfiel, 2019; Figure 3).

Regarding the publication dates, four were three articles published in 2019 (Huberty et al., 2019; Lee et al., 2019; Pasaribu & Zarfiel, 2019; Recabarren et al., 2019), five studies were published in 2020
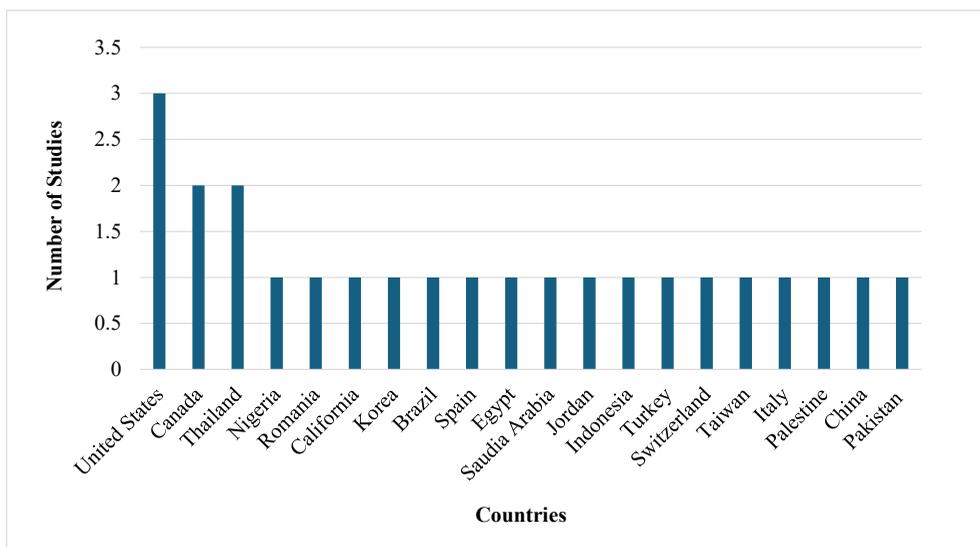
Figure 2. Countries where the selected studies were conducted
Source: Authors' work
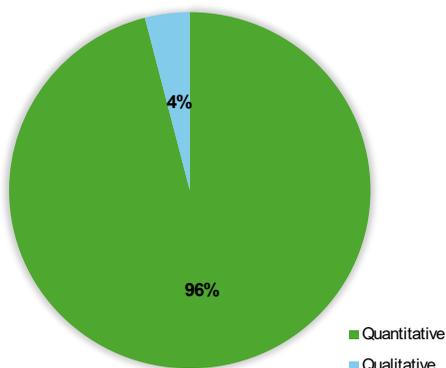
Figure 3. Research design of selected studies
Source: Authors' work

(Ashraf et al., 2020; Azevedo & Menezes, 2020; Cozzolino et al., 2020; El Morr et al., 2020; Shearin & Brewer-Mixon, 2020), five studies were published in 2021 (El Morr et al., 2020; Hamdan et al., 2021; Modrego-Alarcón et al., 2021; Murad, 2021; Rodriguez et al., 2021), eight papers were published in 2022 (Hanani et al., 2022; Hwang et al., 2022; Joompathong et al.,

2022; Mohamed, 2022; Rackoff et al., 2022; Shahrokhian et al., 2022; Sharma et al., 2022; Tomoiaga & David, 2022), and five were published in 2023 (Agah et al., 2023; Changklang & Ranteh, 2023; Eneogu et al., 2023; Kaplan et al., 2023; Newcombe et al., 2023; Figure 4).

Furthermore, the review revealed that two articles were published in the JMIR Mental Health (El Morr et al., 2020; Ritvo et al., 2021). In contrast, only one article was published in the following journals: BMC Psychiatry (Hanani et al., 2022), Brain and Behavior (Cozzolino et al., 2020), Journal of College Student Psychotherapy (Newcombe et al., 2023), Games for Health Journal: Research, Development, and Clinical Applications (Tomoiaga & David, 2022), Journal of Adolescent Health (Rackoff et al., 2022), Healthcare (Sharma et al., 2022), Journal of Medical Internet Research

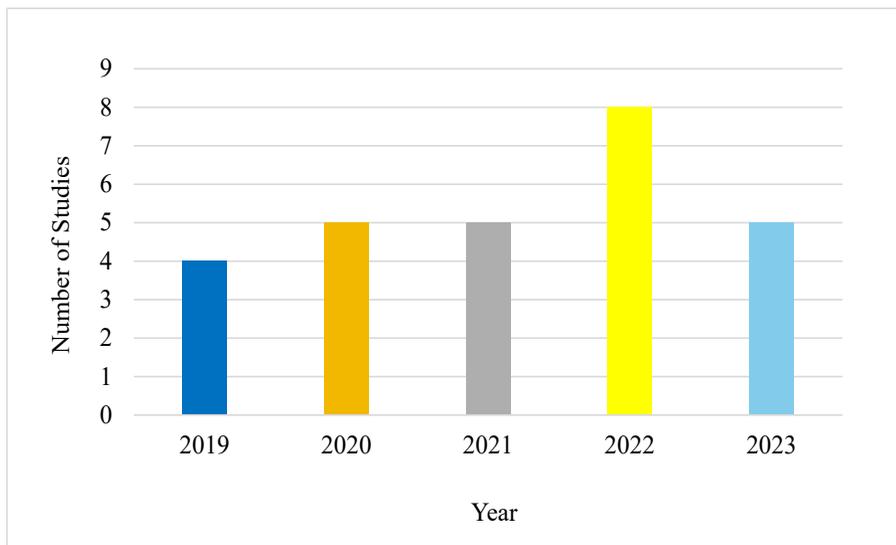*Figure 4.* Countries where the selected studies were conducted
Source: Authors' work

(Hwang et al., 2022), JMIR Formative Research (Rodriguez et al., 2021), Pakistan Journal of Psychology (Ashraf et al., 2020), SMAD, Rev Eletrônica Saúde Mental Álcool Drog (Azevedo & Menezes, 2020), Behavior Research and Therapy (Modrego-Alarcón et al., 2021), Research Journal Specific Education (Mohamed, 2022), The Open Public Health Journal (Joompathong et al., 2022), Universal Journal of Educational Research (Murad, 2021), Advances in Social Science, Education and Humanities Research (Pasaribu & Zarfiel, 2019), Frontiers in Psychiatry (Recabarren et al., 2019), Journal of Physical Therapy Education (Shearin & Brewer-Mixon, 2020), International Journal of Environmental Research and Public Health (Lee et al., 2019), Journal of Research and Health (Kaplan et al., 2023), Health Information and Libraries Journal (Hamdan et al., 2021), Journal of Education and Health Promotion (Changklang &

Ranteh, 2023), Journal of Rational-Emotive and Cognitive-Behavior Therapy (Eneogu et al., 2023), International Journal of Sport Studies for Health (Shahrokhian et al., 2022), JMIR Mhealth Uhealth (Huberty et al., 2019), and Current Psychology (Agah et al., 2023).

The Developed Themes

Thematic analysis was conducted on 27 selected articles, identifying five main themes: objective analysis, delivery methods, session/treatment duration, diversity of CBT interventions for stress issues, and CBT skills/strategies for stress.

Objective Analysis

The primary objective of these studies was to investigate the prevalence of stress among university students and evaluate the effectiveness of cognitive-behavioral

therapy (CBT) in reducing stress and improving mental health outcomes. Additionally, the studies explored the association of stress with other variables, such as academic performance (Ashraf et al., 2020; Mohamed, 2022), academic stress (Ashraf et al., 2020; Eneogu et al., 2023), achievement motivation (Murad, 2021), self-esteem (Changklang & Ranteh, 2023), self-efficacy (Azevedo & Menezes, 2020), self-acceptance (Pasaribu & Zarfiel, 2019), depression (Changklang & Ranteh, 2023; Ritvo et al., 2021), anxiety (Changklang & Ranteh, 2023; Purkayastha et al., 2020; Ritvo et al., 2021), and well-being (Shahrokhian et al., 2022).

Delivery Method

The psychotherapy interventions in the studies utilized both online (El Morr et al., 2020; Hanani et al., 2022; Hwang et al., 2022; Rackoff et al., 2022; Recabarren et al., 2019; Rodriguez et al., 2021; Tomoiaga & David, 2022) and face-to-face sessions (Changklang & Ranteh, 2023; Cozzolino et al., 2020; Kaplan et al., 2023; Newcombe et al., 2023; Shearin & Brewer-Mixon, 2020). Online sessions were conducted through web-based platforms (El Morr et al., 2020; Rodriguez et al., 2021), smartphone apps (Hwang et al., 2022), online games (Tomoiaga & David, 2022) and virtual reality (VR) experiences (Modrego-Alarcón et al., 2021). These virtual sessions allowed for flexibility and accessibility, with participants engaging in group discussions, lectures, and training remotely, often via Zoom or specific mobile applications. On the other hand, face-to-face sessions were also employed, held in person at counseling centers or university campuses. These sessions typically lasted between 60 to 90 minutes, and participants attended weekly meetings over several weeks.

Effectiveness of Online Cognitive Behavioral Therapy for Stress.
Online interventions, specifically self-help and guided online interventions, have remarkably effectively reduced stress, depression, and psychological distress across various populations (Rackoff et al., 2022). These interventions often incorporate a range of beneficial components, including video-based modules (El Morr et al., 2020), mindfulness practices (El Morr et al., 2020; Rodriguez et al., 2021), mindfulness meditation mobile apps (Huberty et al., 2019), discussion forums (El Morr et al., 2020), smartphone-based programs (Hwang et al., 2022), and videoconferencing with a moderator (El Morr et al., 2020). Moreover, online self-help program interventions have demonstrated noteworthy success in reducing stress and enhancing mental health among college students. The evidence strongly supports the idea that online CBT has become a practical and cost-effective alternative to face-to-face CBT, likely due to technological advancements and our increasing familiarity with digital platforms (Hwang et al., 2022). Additionally, computer-based interventions represent a promising approach to bridging the gap between the demand for global mental health treatment and the services received (Rodriguez et al., 2021).

Duration of the Session/Treatment

The duration of treatment sessions in the studies included in the systematic literature review (SLR) varied considerably. Treatment interventions typically lasted six weeks, with follow-up assessments conducted at 6- and 12-month intervals. Most studies conducted treatment for eight weeks (Azevedo & Menezes, 2020; Changklang & Ranteh, 2023; El Morr et al., 2020; Hanani et al., 2022; Huberty et al., 2019; Joompathong et al., 2022; Recabarren et al., 2019).

Some studies conducted only with single sessions (Cozzolino et al., 2020), 4-week treatments (Kaplan et al., 2023; Shearin & Brewer-Mixon, 2020), 6-week treatments (Lee et al., 2019; Pasaribu & Zarfiel, 2019), seven-week treatments (Newcombe et al., 2023) and 10-week treatments (Ashraf et al.,

2020; Hwang et al., 2022; Shahrokhian et al., 2022). Additionally, some interventions had a greater number of treatment sessions, such as 14 sessions (Mohamed, 2022; Murad, 2021) and 16 sessions (Agah et al., 2023; Table 5).

The duration of each session ranged from 45 minutes to 2 hours, depending on the specific program. Participants were encouraged to spend at least 20 minutes per visit on the website or engage in activities related to the intervention. Some interventions included weekly sessions, while others had sessions twice a week. The interventions were conducted in person or online, and each session lasted between 50 and 120 minutes. The interventions sometimes included additional counseling or therapy sessions to support participants' progress.

Table 5
Treatment schedule

Study	Treatment duration	Session length
Hanani et al. (2022)	8 weeks	60 minutes/weekly
Cozzolino et al. (2020)	Single sessions	30 minutes
Newcombe et al. (2023)	7 weeks	30 minutes–45 minutes/twice weekly
El Morr et al. (2020)	8 weeks	N/A
Tomoiaga and David (2022)	7 levels (The REThink game)	The average playtime of a level is between 10 and 20 minutes
Hwang et al. (2022)	10 weeks	50 minutes
Ashraf et al. (2020)	10 sessions	-
Azevedo and Menezes (2020)	8 weeks	2.5 hours each/one per week
Mohamed (2022)	14 sessions	50 minutes each/2 session per week
Joompathong et al. (2022)	8 weeks	90 minutes/per week
Murad (2021)	14 sessions	50 minutes each/twice per week
Pasaribu and Zarfiel (2019)	6 sessions spread over 3 months	Each session lasted approximately 120 min

Table 5 *(Continue)*

Study	Treatment duration	Session length
Recabarren et al. (2019)	8 sessions (2 months)	N/A
Shearin and Brewer-Mixon (2020)	4 sessions	N/A
Lee et al. (2019)	13 goal-directed sessions (6 weeks)	N/A
Kaplan et al. (2023)	4 sessions	45 min sessions
Changklang and Ranteh (2023)	8 weeks	60–90 minutes
Shahrokhian et al. (2022)	10 sessions	120 minutes
Huberty et al. (2019)	8 weeks	10 minutes to 30 minutes/per week
Agah et al. (2023)	16 session	60 minutes per session/twice per week

Source: Authors' work

Diversity of CBT Interventions in Stress Issues

CBT Programs. Cognitive-behavioral therapy (CBT) is a comprehensive program that has demonstrated its effectiveness in reducing psychological Stress (Hanani et al., 2022; Murad, 2021) and enhancing achievement motivation (Mohamed, 2022) among participants. The program thoughtfully incorporates various techniques to help individuals cope with stress and manage its related physical symptoms and emotions. Notably, mindfulness exercises are thoughtfully integrated to promote self-awareness and present-moment focus (Joompathong et al., 2022), while stress management techniques are skillfully taught to equip individuals with practical tools for confidently handling stressors (Joompathong et al., 2022).

Moreover, CBT is a highly effective intervention for improving mental health by skillfully changing cognitive distortions and behaviors (Hanani et al., 2022). Research highlights its remarkable efficacy in reducing depressive symptoms and stress among adolescents (Shahrokhian et al., 2022) and university students (Changklang & Ranteh, 2023). Furthermore, CBT sessions actively empower individuals to identify high-risk situations and thoughtfully develop coping strategies (Agah et al., 2023). Group sessions provide valuable social support and profoundly influence individuals, helping them confidently shape their personalities and behaviors (Changklang & Ranteh, 2023). In a specific study, students in the treatment group experienced an exceptional reduction in examination-induced stress thanks to the impactful intervention sessions (Agah et al., 2023).

Remarkably, CBT is a cost-effective and versatile solution seamlessly applicable in educational environments to address many psychological challenges (Changklang & Ranteh, 2023). Overall, the extensive body

of research consistently underscores the profound effectiveness of CBT in skillfully managing stress and wholeheartedly enhancing overall well-being.

CBT Psychoeducation. Recent research has demonstrated the effectiveness of Cognitive Behavioral Therapy (CBT) psychoeducation in improving mental well-being in students facing psychosocial problems (Kaplan et al., 2023). The stress management program developed by researchers utilized a grouping process within a psychoeducational group, providing participants with valuable skills in managing stress effectively (Joompathong et al., 2022). This approach encouraged the experimental group to exchange their stress-related experiences with one another, allowing them to learn from each other's coping strategies and achievements. Through recording and using worksheets, the experimental group learned and implemented stress management techniques gained from the program in their daily lives (Joompathong et al., 2022). The experimental group reported that exposure to various learning techniques enabled them to confront and manage stress more efficiently (Joompathong et al., 2022). The exchange of learning experiences within the group fostered a supportive environment, enabling members to acquire new coping methods and enhance their stress management skills even further (Joompathong et al., 2022).

In conclusion, CBT psychoeducation offers a promising and beneficial approach to stress management in students. Providing them with the necessary tools and a supportive learning environment equips them with effective strategies to handle stress, fostering improved mental well-being and coping abilities.

CBT, Along with Other Methods. The research findings indicate that integrating CBT with various complementary methods enhances its effectiveness in addressing stress-related issues. These supplementary approaches encompass bibliotherapy (Hamdan et al., 2021), Brain wave modulation (BWM; Cozzolino et al., 2020), video-based programs (Ritvo et al., 2021), mindfulness techniques (Azevedo & Menezes, 2020; Modrego-Alarcón et al., 2021), online game-based cognitive-behavioral therapy (Tomoiaga & David, 2022) and exercise programs (Newcombe et al., 2023).

The study adopts the multi-theory model of health behavior change as its theoretical framework (Sharma et al., 2022). It recognizes that while CBT can be a valuable component in stress management, its implementation alone may not fully address sick leave-related challenges, emphasizing the necessity for supplementary interventions (Hwang et al., 2022). By combining CBT with diverse methods, researchers and practitioners can create comprehensive and holistic approaches to support individuals in managing stress and achieving better well-being.

CBT Skills/Strategies for Stress

CBT sessions involve educating individuals about stress, identifying negative thoughts,

developing coping strategies, and practicing relaxation techniques. Sessions may include understanding stress, identifying automatic thoughts, coping, planning, and developing healthy habits. The intervention includes pre-assessment, goal setting, application of CBT techniques, and post-assessment.

The intervention incorporates various techniques, such as:

1. Psychoeducation (Ashraf et al., 2020; Changklang & Ranteh, 2023; Hanani et al., 2022; Joompathong et al., 2022; Newcombe et al., 2023).

2. Mindfulness (Recabarren et al., 2019; Shearin & Brewer-Mixon, 2020).

3. Cognitive restructuring (though challenging; Ashraf et al., 2020; Hamdan et al., 2021; Newcombe et al., 2023; Recabarren et al., 2019; Tomoiaga & David, 2022).

4. Educating about automatic thought (Agah et al., 2023; Changklang & Ranteh, 2023; Joompathong et al., 2022).

5. Addressing dysfunctional thought (Agah et al., 2023; Eneogu et al., 2023; Pasaribu & Zarfiel, 2019; Shahrokhian et al., 2022).

6. Managing cognitive distortions (Lee et al., 2019; Shearin & Brewer-Mixon, 2020).

7. Learn relaxation exercises (such as deep breathing and progressive muscle relaxation; Ashraf et al., 2020; Hamdan et al., 2021; Hanani et al., 2022; Joompathong et al., 2022; Kaplan et al., 2023; Lee et al., 2019; Mohamed, 2022; Pasaribu & Zarfiel, 2019; Recabarren et al., 2019; Shahrokhian et al., 2022; Tomoiaga & David, 2022).

8. Engage in problem-solving (Agah et al., 2023; Changklang & Ranteh, 2023; Kaplan et al., 2023; Pasaribu & Zarfiel, 2019; Tomoiaga & David, 2022).

9. Modeling technique (Shahrokhian et al., 2022).

10. Engage in goal setting (Newcombe et al., 2023).

11. Engage in time management (Hamdan et al., 2021).

12. Communication skills training (Shahrokhian et al., 2022) and;

13. Assertiveness training (Shahrokhian et al., 2022).

Participants are educated about automatic negative thoughts and taught to recognize and transform them through activities like journaling (Pasaribu & Zarfiel, 2019), completing activity sheets (Pasaribu & Zarfiel, 2019), and engaging in daily homework (Agah et al., 2023; Ashraf et al., 2020; Mohamed, 2022; Recabarren et al., 2019; Shearin & Brewer-Mixon, 2020).

DISCUSSION

The findings of the thematic analysis underscore the positive effects of Cognitive Behavioral Therapy (CBT) on stress management among university students. The primary objective of these studies was to explore stress prevalence and assess the efficacy of CBT interventions in reducing

stress and improving mental well-being in this group.

The range of delivery methods, including online and face-to-face sessions, highlights the evolving field of psychological interventions. Online CBT interventions, facilitated through web-based platforms, mobile apps, and virtual reality experiences, leverage technology to reach a wider audience, offering flexibility and accessibility. Meanwhile, traditional in-person sessions conducted on campuses or counseling centers remain effective. Integrating online platforms acknowledges the preference for digital solutions while recognizing the value of in-person interactions.

CBT interventions offer various delivery options, with online approaches proving particularly effective in stress reduction (Naini et al., 2021; Weiner et al., 2020). Online interventions provide flexibility and accessibility, allowing clients to engage with digital modules independently. Conversely, face-to-face sessions establish a traditional support system, fostering connections between therapists and group members (Maloney et al., 2022; Marins et al., 2020). Group counseling, whether conducted online or in person, enables individuals to communicate, interact, and address shared concerns (Yusop et al., 2020).

The findings strongly support the efficacy of online CBT interventions in reducing stress, depression, and psychological distress. These interventions incorporate various components like video modules (Cozzolino et al., 2020; El Morr et al., 2020; Ritvo et al., 2021; Rodriguez et al., 2021), mindfulness practices (Cozzolino et al., 2020; El Morr et al., 2020; Newcombe et al., 2023; Rodriguez et al., 2021; Shearin & Brewer-Mixon, 2020), and discussion forums (El Morr et al., 2020). Online platforms offer convenience and cost-effectiveness as a practical alternative to traditional face-to-face interventions. Embracing online CBT reflects a shift towards utilizing technology for improved accessibility to mental health resources, addressing the global demand for mental health treatment.

Notably, online group counseling demonstrates effects similar to face-to-face group counseling, underscoring the adaptability of CBT interventions (Oh & Cho, 2020). Leveraging digital media for initial contact and support encourages engagement from hesitant clients (Wong et al., 2018). Service levels can be adjusted from online to in-person interventions, accommodating individual preferences (Wong et al., 2018). Blended interventions, combining both face-to-face and online elements, have proven more effective in stress reduction than either method alone (Benoit, 2023; Borjalilu et al., 2019), broadening the reach of CBT.

Another notable aspect of CBT interventions is their flexibility in duration, ranging from brief single sessions to extended programs (Terp et al., 2022). This adaptability caters to diverse student needs, offering interventions ranging from single sessions to 16-week programs. Session durations, varying from 45 minutes to 2

hours, accommodate different attention spans. It ensures interventions can be tailored to students' schedules and availability, enhancing successful implementation. By aligning stress management interventions with regular activities, universities can increase student engagement and benefits (Parsons et al., 2022; Terp et al., 2022).

The diverse range of CBT interventions highlights their versatility in addressing stress-related challenges. CBT programs effectively enhance mental well-being and motivation by integrating mindfulness exercises and stress management techniques. Integrating psychoeducation and supplementary methods such as bibliotherapy, video-based programs, and exercise regimens further enriches the intervention landscape. This multifaceted approach aligns with the various facets of stress, offering personalized solutions.

Additionally, CBT interventions display versatility in comprehensively addressing stress-related issues. Mindfulness-based interventions and stress management techniques benefit health students. Group sessions provide valuable social support and a nurturing environment for stress management (Zhang et al., 2022). Integrating CBT with complementary methods, including bibliotherapy (Hamdan et al., 2021), brain wave modulation (BWM; Cozzolino et al., 2020), video-based programs (Ritvo et al., 2021), mindfulness techniques (Azevedo & Menezes, 2020; Modrego-Alarcón et al., 2021), online game-based cognitive-behavioral therapy (Tomoiaga & David, 2022), and exercise

programs (Newcombe et al., 2023), enhances the effectiveness of stress management interventions.

The comprehensive array of CBT skills and strategies emphasizes the multifaceted nature of stress management interventions. Psychoeducation, mindfulness, cognitive restructuring, and relaxation techniques equip participants with coping mechanisms. Integrating problem-solving, goal-setting, and communication skills further enhances individuals' ability to navigate stressors effectively. The utilization of various techniques, along with homework assignments and reflective activities, fosters active engagement and skill acquisition.

To conclude, the thematic analysis confirms the positive impact of Cognitive Behavioral Therapy (CBT) on stress management among university students. The adaptability and versatility of CBT interventions, regarding delivery methods, treatment schedules, and supplementary approaches, establish it as a practical and effective strategy for addressing stress-related challenges in university settings. By equipping students with coping strategies, support, and personalized interventions, CBT empowers them to manage stress effectively and enhance overall well-being. The range of approaches showcased in CBT interventions equips universities with valuable tools to support their students' mental health and well-being.

CONCLUSION

The thematic analysis of 27 selected articles provides valuable insights into

the effectiveness of Cognitive Behavioral Therapy (CBT) as a versatile and practical strategy for managing stress among university students. The findings highlight the importance of integrating stress management interventions, particularly CBT, within university settings to effectively enhance students' mental well-being and address stress-related challenges. CBT therapies are sufficiently flexible to accommodate a wide range of requirements and preferences, which enhances patient participation and treatment compliance. By fostering a supportive environment and providing students with effective coping mechanisms, universities can proactively address mental health issues and encourage student success. Supplementary therapies that complement CBT, such as bibliotherapy, mindfulness practices, and exercise regimens, have been shown to maximize stress management interventions and improve mental health outcomes.

However, it is essential to acknowledge certain limitations of this study to refine CBT interventions and strengthen the evidence base to ensure their continued effectiveness in the evolving university landscape. These limitations include potential biases associated with self-reported measures, a lack of population representation (Ashraf et al., 2020; Murad, 2021), and a reliance on self-management rather than discussion forums. Practical challenges, such as a shortage of qualified therapists, have hindered the administration of multiple therapeutic interventions (Agah et al., 2023). Furthermore, the study could

not examine potential variables such as participant education levels, medication use, gender, tribe, age, and religion, which could influence CBT's efficacy (Eneogu et al., 2023).

Future investigations should investigate the mechanisms mediating and moderating intervention effects to address these limitations and advance the field. By identifying the fundamental causes of the intervention's effectiveness, this investigation could enable the design of particularly effective interventions for specific student subgroups. Furthermore, a thorough evaluation of the intervention's long-term impacts and influence on academic outcomes is necessary (El Morr et al., 2020).

Future studies should also evaluate its sustainability over time and determine the need for ongoing support to ensure the lasting impact of the intervention (Hanani et al., 2022). Moreover, understanding the differential effects of these interventions across diverse student cohorts is crucial for maintaining the advantages and how different student cohorts respond to the intervention (Cozzolino et al., 2020). By pursuing these avenues, future research can provide a comprehensive understanding of the intervention's effectiveness, shedding light on its potential to enhance overall student well-being and success (Newcombe et al., 2023).

In summary, this study significantly advances the theoretical foundation and real-world application of stress management among college students by highlighting

the importance of CBT therapies (Table 6). The research highlights CBT's adaptability and usefulness and argues that university mental health frameworks should incorporate it to improve student well-being. Theoretically, CBT's adaptability supports strong theoretical models in a variety of demographic and situational scenarios, indicating that the techniques it employs are generally useful and relevant in a range of contexts. From a practical standpoint, the study emphasizes how crucial it is for universities to implement proactive mental health policies, such as customized CBT programs, to address the particular needs of their student bodies. These treatments continue to be relevant and effective in changing educational environments to ensure that future research should focus on refining these treatments, analyzing their long-term benefits, and investigating their sustainability and adaptation across varied student populations.

Table 6
Summary of major conclusions

Key points	Summary
Thematic analysis of CBT for stress management in university students	• CBT is an effective approach for managing stress among university students. • It should be integrated into university settings to support mental well-being and effective stress coping. • The adaptability of CBT allows personalized approaches, improving engagement and adherence. • Integrating CBT with methods like bibliotherapy and mindfulness can optimize stress management.
Importance of addressing mental health and well-being in university settings	• Universities should create a supportive environment and equip students with coping strategies to promote mental health and success. • Innovative approaches can foster holistic well-being.
The focus of future research: Mediators/moderators, Long-term effects, Academic outcomes	• Future studies should explore the mechanisms of intervention effects and consider tailoring them for specific student subgroups. • Also, examine long-term effects and impact on academic outcomes.
Addressing limitations to enhance the evidence base for CBT interventions	• Addressing biases and representativeness and using diverse approaches can refine CBT interventions to support mental well-being in evolving university landscapes.

Note. The summarized table encapsulates the primary conclusions from this study's outcomes.
Source: Authors' work

ACKNOWLEDGMENT

REFERENCES

Abouzahra, A., Sabraoui, A., & Afdel, K. (2020). Model composition in model driven engineering: A systematic literature review. *Information and Software Technology*, *125*, Article 106316. https://doi.org/10.1016/j.infsof.2020.106316

Agah, J. J., Ede, M. O., Asor, L. J., Ekesionye, E. N., & Ejionueme, L. (2023). Managing examination induced stress among students using FEAR-model of cognitive behavioural intervention: Policy implications for educational evaluators. *Current Psychology*, *42*(4), 3011-3023. https://doi.org/10.1007/s12144-021-01657-z

Ahmad, T. B., & Meriç, M. (2021). The effect of an online psychoeducational stress management program on international students' ability to cope and adapt. *Perspectives in Psychiatric Care*, *57*(4), 1673-1684. https://doi.org/10.1111/ppc.12735

Ashraf, M., Masroor, U., & Khan, M. J. (2020). Effectiveness of cognitive behavior therapy in reduction of academic stress and enhancement of academic performance in university students. *Pakistan Journal of Psychology, 51*(2), 3-13.

Asl, R. G., Taghinejad, R., Parizad, N., & Jasemi, M. (2022). The relationship between professional autonomy and job stress among intensive care unit nurses: A descriptive correlational study. *Iranian Journal of Nursing and Midwifery Research*, *27*(2), 119-124. https://doi.org/10.4103/ijnmr.ijnmr_375_20

Azevedo, M. L.D, & Menezes, C. B. (2020). Effects of the mindfulness-based cognitive therapy program on stress, self-efficacy and mindfulness in university students. *SMAD Revista Eletrônica Saúde Mental Álcool e Drogas (Edição Em Português)*, *16*(3), 44-54. https://doi.org/10.11606/issn.1806-6976.smad.2020.165513

Benoit, S. (2023). Blended online and face-to-face art therapy. *Journal of Applied Arts & Health*, *14*(1), 101-115. https://doi.org/10.1386/jaah_00130_1

Borjalilu, S., Mazaheri, M. A., & Talebpour, A. (2019). Effectiveness of mindfulness-based stress management in the mental health of Iranian university students: A comparison of blended therapy, face-to-face sessions, and mHealth app (*Aramgar*). *Iranian Journal of Psychiatry and Behavioral Sciences*, *13*(2), Article e84726. https://doi.org/10.5812/ijpbs.84726

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Changklang, P., & Ranteh, O. (2023). The effects PF cognitive behavioural therapy on depression, anxiety, stress and self-esteem in public health students, Thailand. *Journal of Education and Health Promotion*, *12*(1), Article 152. https://doi.org/10.4103/jehp.jehp_1274_22

Cozzolino, M., Girelli, L., Vivo, D. R., Limone, P., & Celia, G. (2020). A mind–body intervention for stress reduction as an adjunct to an information session on stress management in university students. *Brain and Behavior*, *10*(6), Article e01651. https://doi.org/10.1002/brb3.1651

El Morr, C., Ritvo, P., Ahmad, F., & Moineddin, R. (2020). Effectiveness of an 8-week web-based mindfulness virtual community intervention for university students on symptoms of stress, anxiety, and depression: Randomized controlled trial. *JMIR Mental Health*, *7*(7), Article e18595. https://doi.org/10.2196/18595

Eneogu, N. D., Ugwuanyi, C. K., & Ugwuanyi, C. S. (2023). Efficacy of cognitive behavioral therapy on academic stress among rural community secondary school economics students: A randomized controlled evaluation. *Journal of Rational - Emotive and Cognitive - Behavior Therapy*, *42*, 245-262. https://doi.org/10.1007/s10942-023-00508-z

Flemming, K., Booth, A., Garside, R., Tunçalp, Ö., & Noyes, J. (2019). Qualitative evidence synthesis for complex interventions and guideline development: Clarification of the purpose, designs and relevant methods. *BMJ Global Health*, *4*(Suppl 1), Article e000882. https://doi.org/10.1136/bmjgh-2018-000882

Hamdan, W., Al Duaijy, L., & Al Sawy, Y. (2021). Effectiveness of bibliotherapy in alleviating examination stress on college students: A quasi-experimental trial. *Health Information and Libraries Journal*, *38*(3), 172-181. https://doi.org/10.1111/hir.12312

Hanani, A., Badrasawi, M., Zidan, S., & Hunjul, M. (2022). Effect of cognitive behavioral therapy program on mental health status among medical student in Palestine during COVID Pandemic. *BMC Psychiatry*, *22*(1), Article 310. https://doi.org/10.1186/s12888-022-03915-1

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343* Article d5928. https://doi.org/10.1136/bmj.d5928

Huberty, J., Green, J., Glissmann, C., Larkey, L., Puzia, M., & Lee, C. (2019). Efficacy of the mindfulness meditation mobile app *"calm"* to reduce stress among college students: Randomized controlled trial. *JMIR MHealth and UHealth*, *7*(6), Article e14273. https://doi.org/10.2196/14273

Hwang, H., Kim, S. M., Netterstrøm, B., & Han, D. H. (2022). The efficacy of a smartphone-based app on stress reduction: Randomized controlled trial. *Journal of Medical Internet Research*, *24*(2), Article e28703. https://doi.org/10.2196/28703

Joompathong, N., Bunrayong, W., & Chupradit, S. (2022). Stress management programme on the stress of Chiang Mai University students: A pilot study. *Journal of Experimental Biology and Agricultural Sciences, 10*(4), 879-885. https://doi.org/10.18006/2022.10(4).879.885

Kaplan, V., Düken, M. E., Kaya, R., & Almazan, J. (2023). Investigating the effects of cognitive-behavioral-therapy-based psychoeducation program on university students' automatic thoughts, perceived stress, and self-efficacy levels. *Journal of Research and Health*, *13*(2), 87-98. https://doi.org/10.32598/JRH.13.2.2125.1

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher*, *42*(8), 846-854. https://doi.org/10.1080/0142159X.2020.1755030

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Technical Report EBSE-2007-01). School of Computer Science and Mathematics, Keele University.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and Software Technology*, *51*(1), 7-15.

Lee, T. Y., Hsieh, T. C., Sung, H. C., & Chen, W. L. (2019). Internet-delivered cognitive behavior therapy for young Taiwanese female nursing students with irritable bowel syndrome-a cluster randomized controlled trial. *International Journal of Environmental Research and Public Health*, *16*(5), Article 708. https://doi.org/10.3390/ijerph16050708

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., John, P. A., Clarke, M., Devereaux, P., Kleijnen, J., & Moher, D.

(2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *Journal of Clinical Epidemiology, 62*(10), e1-e34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Maloney, E., Dowling, C., Deehan, T., & O' Reilly, G. (2022). Comparing people's experience of therapy for complex trauma across three modalities: Face-to-face, mixed, and online. *OBM Integrative and Complementary Medicine*, *7*(4). https://doi.org/10.21926/obm.icm.2204051

Marins, M. da R., Daher, D. V., Pinto, A. A., Cardoso, R. da S. S., & Sá, S. P. C. (2020). Art therapy and education between peers connecting the group: An experience report. *Revista Brasileira de Enfermagem*, *73*(5), Article e20180958. https://doi.org/10.1590/0034-7167-2018-0958

Modrego-Alarcón, M., López-del-Hoyo, Y., García-Campayo, J., Pérez-Aranda, A., Navarro-Gil, M., Beltrán-Ruiz, M., Morillo, H., Delgado-Suarez, I., Oliván-Arévalo, R., & Montero-Marin, J. (2021). Efficacy of a mindfulness-based programme with and without virtual reality support to reduce stress in university students: A randomized controlled trial. *Behaviour Research and Therapy*, *142,* Article 103866. https://doi.org/10.1016/j.brat.2021.103866

Mohamed, M. A. (2022). The effect of applying the cognitive-behavioral therapy program on psychological stress and improving the achievement motivation of first-and second-year university students in Egypt. *Research Journal Specific Education,* (71), 27-48. https://doi.org/10.21608/mbse.2022.156809.1237

Murad, O. S. (2021). Effectiveness of a cognitive-behavioral therapy program on reducing psychological stress and improving achievement motivation among university students. *Universal Journal of Educational Research*, *9*(6), 1316-1322. https://doi.org/10.13189/ujer.2021.090621

Naini, R., Wibowo, M. E., & Mulawarman, M. (2021). Efficacy of online group counseling with mindfulness-based cognitive approach to enhance students' humility. *Islamic Guidance and Counseling Journal*, *4*(1), 78-90. https://doi.org/10.25217/igcj.v4i1.1280

Newcombe, B. C., Olthuis, J. V., MacLean, M., Hamilton, R., & McAulay, T. (2023). CBT + exercise vs treatment as usual in treating anxiety and depression in university students: A pilot study. *Journal of College Student Mental Health*, *38*(2), 311-327. https://doi.org/10.1080/87568225.2023.2181255

Oh, Y., & Cho, E. (2020). Effects and traits of online narrative group counseling using video conferencing modality: Comparing with face-to-face group counseling. *The Journal of the Convergence on Culture Technology*, *6*(4), 401-409. https://doi.org/10.17703/JCCT.2020.6.4.401

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, *37*, 879-910. https://doi.org/10.17705/1CAIS.03743

Page, M. J., Mckenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-wilson, E., Mcdonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372,* Article n71. https://doi.org/10.1136/bmj.n71

Parsons, D., Gardner, P., Parry, S., & Smart, S. (2022). Mindfulness-based approaches for managing stress, anxiety and depression for health students in tertiary education: A scoping review. *Mindfulness*, *13*, 1-16. https://doi.org/10.1007/s12671-021-01740-3

Pasaribu, P. E., & Zarfiel, M. D. (2019). Cognitive behavioral therapy treatment for reducing stress:

A case study of self-acceptance in an early adult college student. *Advances in Social Science, Education and Humanities Research*, *229*, 631-644. https://doi.org/10.2991/iciap-18.2019.54

Purkayastha, S., Addepally, S. A., & Bucher, S. (2020). Engagement and usability of a cognitive behavioral therapy mobile app compared with web-based cognitive behavioral therapy among college students: Randomized heuristic trial. *JMIR Human Factors*, *7*(1), Article e14146. https://doi.org/10.2196/14146

Rackoff, G. N., Fitzsimmons-Craft, E. E., Taylor, C. B., Eisenberg, D., Wilfley, D. E., & Newman, M. G. (2022). A randomized controlled trial of internet-based self-help for stress during the COVID-19 pandemic. *Journal of Adolescent Health*, *71*(2), 157-163. https://doi.org/10.1016/j.jadohealth.2022.01.227

Recabarren, R. E., Gaillard, C., Guillod, M., & Martin-Soelch, C. (2019). Short-term effects of a multidimensional stress prevention program on quality of life, well-being and psychological resources. A randomized controlled trial. *Frontiers in Psychiatry*, *10*, Article 88. https://doi.org/10.3389/fpsyt.2019.00088

Ritvo, P., Ahmad, F., Morr, C. El, Pirbaglou, M., & Moineddin, R. (2021). A mindfulness-based intervention for student depression, anxiety, and stress: Randomized controlled trial. *JMIR Mental Health*, *8*(1), Article e23491. https://doi.org/10.2196/23491

Rodriguez, M., Eisenlohr-Moul, T. A., Weisman, J., & Rosenthal, M. Z. (2021). The use of task shifting to improve treatment engagement in an internet-based mindfulness intervention among Chinese university students: Randomized controlled trial. *JMIR Formative Research*, *5*(10), Article e25772. https://doi.org/10.2196/25772

Shaffril, H. A. M., Ahmad, N., Samsuddin, S. F., Samah, A. A., & Hamdan, M. E. (2020). Systematic literature review on adaptation towards climate change impacts among indigenous people in the Asia Pacific regions. *Journal of Cleaner Production*, *258*, Article 120595. https://doi.org/10.1016/j.jclepro.2020.120595

Shaffril, H. A. M., Samah, A. A., & Kamarudin, S. (2021). Speaking of the devil: A systematic literature review on community preparedness for earthquakes. *Natural Hazards*, *108*(3), 2393-2419. https://doi.org/10.1007/s11069-021-04797-4

Shahrokhian, N., Hassanzadeh, S., Razini, H. H., & Ramshini, M. (2022). The effects of cognitive-behavioral therapy (CBT) in well-being and perceived stress in adolescents with low academic performance during the COVID-19 pandemic. *International Journal of Sport Studies for Health*, *4*(2), 45-52. https://doi.org/10.61838/kman.intjssh.4.2.6

Sharma, M., Kanekar, A., Batra, K., Hayes, T., & Lakhan, R. (2022). Introspective meditation before seeking pleasurable activities as a stress reduction tool among college students: A multi-theory model-based pilot study. *Healthcare*, *10*(4), Article 614. https://doi.org/10.3390/healthcare10040614

Shearin, S. M., & Brewer-Mixon, K. (2020). Effectiveness of a short education series to reduce anxiety for health professions graduate students: A pilot study. *Journal of Physical Therapy Education*, *34*(1), 12-18. https://doi.org/10.1097/JTE.0000000000000124

Terp, U., Bisholt, B., & Hjärthag, F. (2022). A feasibility study of a cognitive behavioral based stress management intervention for nursing students: Results, challenges, and implications for research and practice. *BMC Nursing*, *21,* Article 30. https://doi.org/10.1186/s12912-021-00761-6

Tomoiaga, C., & David, O. (2022). The efficacy of guided and unguided game-based cognitive-behavioral therapy in reducing distress in college

students. *Games for Health Journal: Research, Development, and Clinical Applications*, *11*(6), 403-413. https://doi.org/10.1089/g4h.2021.0195

Weiner, L., Berna, F., Nourry, N., Severac, F., Vidailhet, P., & Mengin, A. C. (2020). Efficacy of an online cognitive behavioral therapy program developed for healthcare workers during the COVID-19 pandemic: The REduction of STress (REST) study protocol for a randomized controlled trial. *Trials*, *21*, Article 870. https://doi.org/10.1186/s13063-020-04772-7

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546-553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Wong, K. P., Bonn, G., Tam, C. L., & Wong, C. P. (2018). Preferences for online and/or face-to-face counseling among university students in Malaysia. *Frontiers in Psychology*, *9*, Article 64. https://doi.org/10.3389/fpsyg.2018.00064

Yusop, Y. M., A. Rahman, N., Zainudin, Z. N., Ismail, A., Othman, W. N. W., & Sumari, M. (2020). The effectiveness of cognitive behavioral therapy in group counselling. *International Journal of Academic Research in Business and Social Sciences*, *10*(9), 359-379. https://doi.org/10.6007/ijarbss/v10-i9/7648

Zhang, Z., Li, Y., Sun, S., & Tang, Z. (2022). Intervention effect of group counseling based on positive psychology on psychological crisis of college student. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2022/3132016

The Impact of Antisocial Media Towards Aggressive Behavior Among Online Gamers in Malaysia

Mohd Ziyad Afiq Mohd Zaharim*, Azlina Mohd Khir, Nik Ahmad Sufian Burhan and Amna Md Noor

Department of Social and Development Sciences, Faculty of Human Ecology, Universiti Putra Malaysia, 43400, Serdang, Malaysia

ABSTRACT

The global craze of digital gaming has attracted millions of individuals worldwide, particularly during the COVID-19 pandemic. While it has allure, it also carries drawbacks. Notably, it may lead to changes in conduct, such as a higher tendency for violence. In light of this, the study aims to identify the predictors of aggressive behaviors among online gamers. We recruited 384 local online gamers from the Klang Valley area using a purposive and snowball sampling technique. Data analysis was performed using the Statistical Equation Model-Partial Least Square (SEM-PLS), incorporating path analysis. The findings revealed a negative relationship between moral identity and self-esteem toward aggressive behavior, indicating that aggressive behavior decreases as moral identity and self-esteem increase. Conversely, a positive relationship was found between antisocial media exposure and aggressive behavior. It means that as antisocial media exposure increases, aggressive behavior also increases. In conclusion, this study demonstrated that high levels of moral identity, self-esteem, and a low level of antisocial media exposure are crucial predictors in decreasing aggressive behavior among local online gamers. Therefore, the study suggests policymakers should develop programs to boost moral identity and self-esteem among Malaysian online gamers to reduce aggression. It also calls for further research on gamers from various regions, other influencing factors, behavioral changes over time, and in-depth qualitative studies.

Keywords: Aggressive, gamers, Malaysia, media exposure, self-esteem, violent

INTRODUCTION

In today's ever-evolving digital landscape brought on by advancing technologies, online communities and social interactions are constantly being transformed. Online communities have become the central hub for fostering social engagement and facilitating digital collaboration. Within this sphere, individuals can communicate and establish connections across geographical boundaries. This transition represents a profound shift in individuals' mode of engagement in the modern digital age. As a result, this phenomenon gives rise to fresh complexities in the intricate relationship between morality and aggressive behavior within the digital sphere, as elucidated by Barlett and Coyne (2019).

The presence of content in the media that may promote antisocial behaviors has always been a topic of interest in scientific studies. Due to the diverse ways violence is depicted in video games, television shows, news, movies, and social media to capture the audience's attention, it has become significant in all media-related production and sales businesses. While scientific research on the effects of media first emerged in the 1920s and 1930s, it only gained wider prominence by the late 1950s. All these previous research developments gave rise to bountiful theories and evidence on media effects. Since the 1990s, several meta-analyses have also been conducted to consolidate the findings of all these empirical studies. It was highlighted that adolescents who are regularly exposed to content on various media platforms

may promote antisocial behaviors such as substance abuse and violence (Strasburger et al., 2010). They are also at an elevated risk of engaging in aggressive acts and displaying problematic alcohol consumption patterns (Khurana et al., 2019). Previous scholars have framed antisocial media as delivering and promoting certain risky behaviors. For example, den Hamer et al. (2017) propose that antisocial and risky behaviors such as theft, excessive drinking, profanity, and physical altercations are more common among young individuals, as observed on various media outlets. Antisocial media content is pervasive on multiple platforms, including television shows (e.g., Hetsroni, 2007), films (e.g., Bushman et al., 2013), video games (e.g., Hartmann et al., 2014), music videos (e.g., Martin & Collins, 2002), and even advertisements (e.g., Jones et al., 2014).

To date, the media has been shown to indirectly stimulate acts of violence, causing an increasing presence of violence in society despite various prevention strategies to combat such negativity. Often, it is challenging to establish a direct relationship between aggression and antisocial media exposure (ASME). Most of the previous research listed ASME only as a causal factor of aggression, as a variety of factors can also cause aggression. So far, research has shown that ASME can predispose to aggressive behavior (AB), aggressive cognition, diminished empathy, and an increase in violence (Coyne, 2016).

In Malaysia, aggressive behavior has been extensively investigated in various

populations, including secondary school students, illegal motorcycle racers (*mat rempit*), road drivers, juveniles, moviegoers, rehabilitation center residents, and prison inmates. However, there is a noticeable gap in research with regard to aggressive behavior in the online gaming community. Nasser et al. (2021) found that most Malaysian youth are involved in online gaming or watching films on YouTube or Netflix. A recent report also highlighted online activities as the preferred media activity in Malaysia (Malaysian Communications and Multimedia Commission [MCMC], 2020). According to Ismail et al. (2021) and T'ng and Pau (2020), it is impossible to completely ban inappropriate media exposure and online games if the goal is to protect children from potential exposure to negative online content. Very often, the negative impact of online media content may differ among youths based on their personality or psychological states (Bushman, 2017). According to the General Aggression Model, psychological traits such as self-esteem (SE) may influence an individual's propensity for negative behavior (Anderson & Bushman, 2002; Balakrishnan & Fernandez, 2018). In other words, when situational factors interact with human personality, the presence and severity of an individual's aggressive behavior may be affected. Past research has highlighted the relationship between low SE and victimization caused by aggressive behavior. For instance, Schultze-Krumbholz et al. (2018) suggested that individuals exhibiting low SE are more prone to becoming victims.

Conversely, other research has explored the association between low SE and higher uptake of aggressive behavior. In view of the growing influence of media exposure and digital activity, it is imperative to investigate whether personal aspects of SE are related to aggressive behavior, especially among the youth of Malaysia who are actively engaged with antisocial media.

While there has been substantial research on the individual concepts of moral identity, self-esteem, and antisocial behavior on social media, there appears to be a significant gap in the literature regarding the interplay among these three variables. For instance, studies have explored the antecedents of moral identity (Smith, 2019), the development of self-esteem (Johnson & Johnson, 2020), and the rise of antisocial behavior on social media (Brown, 2021). However, the potential influence of moral identity and self-esteem on antisocial behavior on social media still needs to be explored. Furthermore, cultural and demographic factors' impact on these relationships needs to be better understood (Williams, 2022; Zhang, 2023). This gap is particularly relevant given the increasing prevalence of antisocial behavior on social media (Brown, 2021) and the recognized importance of self-esteem and moral identity in individuals' behavior (Smith, 2019; Johnson & Johnson, 2020). Therefore, this study addresses this gap by examining the relationships among moral identity, self-esteem, and ASME on social media.

The General Aggression Model (GAM), developed by Anderson and Bushman, provides a comprehensive framework for understanding aggression. GAM predicts the likelihood of aggressive behavior based on a combination of personal and situational factors, and it can be used to understand and potentially reduce violent behavior in various contexts. It incorporates social, cognitive, personality, developmental, and biological factors (Anderson & Bushman, 2002). The model has been criticized for its broadness, lack of specificity, and inability to consider social and cultural factors (Ferguson & Dyck, 2012). Despite facing criticisms, the General Aggression Model (GAM) has found extensive application in diverse contexts. These include the study of the effects of media violence, domestic violence, intergroup violence, temperature effects, pain effects, and the impacts of global climate change (Anderson & Bushman, 2002). Recent research, such as those conducted by Jiang et al. (2022) and Gagnon et al. (2022), has consistently employed the GAM, underscoring its continued relevance. Meanwhile, Simanjuntak et al. (2021) explored the relationship between online game violence and adolescent aggression, establishing a low-to-moderate correlation. Our study builds upon this foundation but with a broader scope.

While valuable in outlining factors influencing aggression, it has limitations (Ferguson et al., 2011). GAM primarily focuses on situational cues and personality traits, neglecting the potential moderating role of moral identity (Carnagey &

Anderson, 2003). This study aims to extend the current understanding of aggression by examining how a strong moral identity might buffer the influence of self-esteem and exposure to antisocial content. Moral identity, a multifaceted construct encompassing cognitive, affective, and self-referent components (Aquino et al., 2023), could protect against the aggression-promoting effects of low self-esteem and exposure to negativity online. Self-esteem, a distinct yet interrelated construct, reflects an individual's overall positive self-evaluation, encompassing their abilities and inherent worth (Gentile, Greer, et al., 2020). Crucially, this study distinguishes between general social media use and exposure to antisocial content (Tedeschi et al., 2021).

The present study delves into the interplay between moral identity, self-esteem, and exposure to antisocial content on social media platforms. Moral identity, a cornerstone of ethical behavior, transcends singular aspects of moral reasoning. It encompasses a multifaceted framework that integrates cognitive (moral reasoning), affective (moral emotions like guilt and empathy), and self-referent components (moral self-image, seeing oneself as a good person) (Aquino et al., 2023). Moral identity is crucial in regulating the impact of other psychological factors on aggression. Individuals with a strong, integrated moral identity are more likely to utilize moral reasoning and experience emotions like guilt or empathy to inhibit aggressive tendencies, even when provoked (Graham et al., 2011). Self-esteem, a distinct yet interrelated

construct, reflects an individual's overall positive self-evaluation, encompassing their abilities and inherent worth (Gentile, Twenge, et al., 2020). Meanwhile, Antisocial media exposure (ASME) refers to exposure to all forms of social media content that depict behaviors considered antisocial, violating established societal norms and legal frameworks (den Hamer et al., 2017). Antisocial media content, characterized by negativity, aggression, and cyberbullying, represents a specific dimension of the online environment with potentially distinct influences (Tedeschi et al., 2021). This study further distinguishes between general social media use and exposure to antisocial content.

This study focused on the potential function of moral identity (MI) in regulating the impacts of ASME on aggression among Malaysian youth. Past studies have shown that MI is correlated to aggression, whereby it acts as a crucial component in regulating the impact of other psychological factors on aggression (Skobkareva, 2020; Teng et al., 2020). Moral identity plays a significant role as a crucial component in regulating the impact of other psychological factors on aggression. Anticipated feelings of guilt mediate the effect of moral identity on antisocial behavior. Individuals with a solid moral identity who contemplate aggressive actions anticipate feelings of guilt, which serves as a deterrent (Gini et al., 2021).

In this context, a solid moral identity could reduce moral disengagement, thereby reducing aggression (Falla et al., 2021). Hence, we set out to determine if the moral

identities of local Malaysian online gamers exhibit a similar influence on SE and deviant peer associations, as well as their role in mitigating the adverse effects of ASME on aggressive behavior. The primary aim of this study was to assess the potential impact of ASME, SE, and MI on aggressive behavior by using the General Aggression Model (GAM) as a theoretical framework.

To summarize, our objective is to examine the significant impact of Antisocial Social Media Exposure (ASME) on aggressive behavior, differentiating it from general social media use. We hypothesize that exposure to content specifically promoting violence, hate speech, and other antisocial behaviors will have a stronger association with aggression. Furthermore, we also aim to explore the interplay between Moral Identity (MI) and Self-Esteem (SE) within the GAM framework. We propose that strong MI, characterized by a well-developed moral compass and high Self-Esteem, can act as protective factors, mitigating the negative influence of ASME on aggressive tendencies among local online gamers in Malaysia (Figure 1).
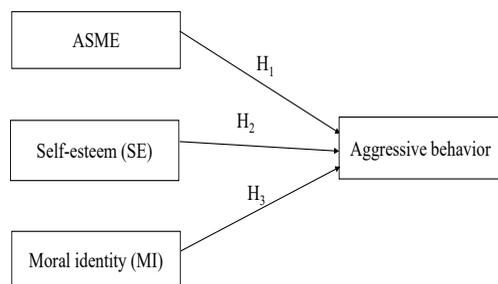

Figure 1. Conceptual framework
Source: Authors' work

METHODS

Sample and Procedure

This quantitative cross-sectional study used a correlational design to examine the relationship between ASME and aggressive behavior. An online survey platform was used to distribute the questionnaire for data collection in the Klang Valley. The area is home to most of the Malaysian population, with 81% of the internet users residing there (MCMC, 2020). In addition, it was also the location for many major e-sports competitions, such as the one held at the Melawati Stadium in Shah Alam in 2020, as Yunus et al. (2021) reported.

Using the formula by Krejcie and Morgan (1970), the minimum sample size required was 384. The inclusion criteria were Malaysian citizens between 18 and 40 who played a massive multiplayer online role-playing game (MMORPG). Individuals who could not understand or communicate in English or Malay were excluded.

Before the data collection, a pilot test was conducted among 30 participants between November and December 2022. Based on Browne's (1995) recommendation, to achieve an 80% upper confidence limit (UCL), a pilot study or pre-test should include 30 subjects. Kieser and Wassmer (1996) also recommended a sample size between 20 and 40 for pilot tests. The Cronbach's alpha value of each construct was greater than 0.70 (Hair et al., 2019), while the UCL was approximately 85%.

Measure

In this study, multiple psychometric instruments were employed to evaluate salient constructs. The Content-based Media Exposure Scale (C-ME), adapted from den Hamer et al. (2017), was utilized to measure ASME. Of the 17 items, eight were under the "antisocial" domain and nine under the "natural" domain. All items were measured on a five-point Likert scale from 1 (never) to 5 (very often), with an α-value of 0.988. The Content-based Media Exposure Scale (C-ME), adapted from den Hamer et al. (2017), is chosen for its unique ability to assess Antisocial Social Media Exposure (ASME), a critical aspect of this study, allowing for a crucial differentiation between exposure to "antisocial" and "natural" content, given the distinct effects of negativity and aggression online on moral development and behavior (Huang et al., 2022; Tedeschi et al., 2021).

Next, the Moral Identity Questionnaire, a 20-item scale adapted from Black and Reynolds (2016), was implemented to assess participants' MI dimensions on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). It demonstrated an alpha (α) of 0.985. The Moral Identity Questionnaire, adapted from Black and Reynolds (2016), is selected for its comprehensive capture of the multifaceted construct of moral identity, aligning with the research focus on understanding how different aspects of moral identity interact with other variables (Aquino et al., 2023; Graham et al., 2011).

Meanwhile, the ten items under the Rosenberg Self-esteem Scale (RSES,1979)

were utilized to measure the level of SE on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree) with an α-value of 0.898. The Rosenberg Self-esteem Scale (RSES; Rosenberg, 1979) is chosen for its wide use and validation in measuring self-esteem, a construct of interest due to its complex relationship with social media use and its potential to make individuals more susceptible to the negative influences of antisocial content, potentially leading to increased aggression (Gentile, Twenge, et al., 2020).

Lastly, the Reactive and Proactive Aggression Questionnaires (RPQ), which encompassed 23 items on a three-point Likert scale from 1 (never) to 3 (often) with an α-value of 0.978, adapted from Raine et al. (2006), was also incorporated in the questionnaire. The RPQ, adapted from Raine et al. (2006), are selected for their ability to differentiate between reactive and proactive aggression, providing a nuanced understanding of how self-esteem and social media exposure might influence distinct types of aggressive behavior (Ferguson et al., 2011). These instruments, all well-established and validated in previous research, ensure the reliability and validity of the findings in this context.

Following data collection, descriptive statistics were used to report the participants' characteristics. Next, the reliability and validity of the constructs were also assessed. Finally, structural equation modeling (SEM) was employed for path analysis.

Reliability and Validity Analysis

Senior academic experts were invited to assess the questionnaire before the pilot study to determine the content validity. Thereafter, amendments and revisions were made to all measurement items in three aspects based on the expert feedback. Firstly, the objective of the statement was edited to provide a clearer purpose for the current research. Secondly, the sociodemographic-related items were also amended. Lastly, some items of the remaining variables were also redesigned.

In the next step, the reliability of each independent variable in the questionnaire was assessed using Partial Least Squares-Confirmatory Composite Analysis (PLS-CCA) to obtain the quality of the reflective measurement model. PLS-CCA aims to assess the reliability and validity of reflective PLS-SEM measurement models via three categories: internal consistency reliability (ICR), convergent validity (CV), and discriminant validity (DV).

Table 1 presents the Composite Reliability (CR) and Cronbach Alpha (CA), representing the internal consistency and reliability of the items in the questionnaire. Each item's standardized factor loadings (estimate) exceeded the critical value 0.5. Based on Table 1, any ICR results above 0.95 indicated problematic data. However, according to Becker et al. (2023), if the researcher has strategically employed appropriate data collection (e.g., avoiding respondents being distracted by the questions and any potential ensuing demand effects) in

optimizing the construct (e.g., never using a redundant and synonymous item in the questionnaire), then the ICR result should be considered as a good match to the concept of reliability. Therefore, the nearer the value is to 1, the more reliable the items are.

Convergent analysis refers to the degree to which a specified construct explains the variance of its indicator reflectively (Hair et al., 2021). In this study, convergent validity was evaluated using the established technique known as "Average Variance Extracted" (AVE). The AVE for each construct was above the recommended value of 0.5 in this study.

The next step was determining the outer loading or indicator to see how each item contributed to that construct (Hair et al., 2017). Table 2 presents the outer loading result as between 0.75 and 0.99. Therefore, several items (D14, E15,17, 23, and 5) were removed to ensure the outer loading met the standard of assessment of the reflective model. Furthermore, the missing data in this study was analyzed by computing the cases for each variable with missing data. Based on the recommendation of Hair et al. (2009), any missing data of less than 10% can be ignored.

Table 1
Internal Consistency Reliability (ICR)

Construct	Items	CA	CR	AVE
Aggressive behavior	AB	0.976	0.981	0.692
Antisocial media exposure	ASME	0.963	0.967	0.765
Moral identity	MI	0.983	0.984	0.745
Self-esteem	SE	0.978	0.983	0.813

Note. CA = Cronbach's alpha, CR = Combined Reality; AVE = Average Variance Extracted
Source: Authors' work

Table 2
Outer loading

Construct	Item of the construct	Outer loading
ASME1	How often do you watch people who fight on the Internet/TV/mobile phone/DVD?	0.995
ASME2	How often do you watch on the Internet/TV/mobile phone/DV people who use drugs?	0.822
ASME3	How often do you watch on the Internet/TV/mobile phone/DVD people who drink a lot of alcohol?	0.979
ASME4	How often do you watch on the Internet/TV/mobile phone/DVD people who are having sex?	0.833

Table 2 *(Continue)*

Construct	Item of the construct	Outer loading
ASME5	How often do you watch people who steal on the Internet/ TV/mobile phone/DVD?	0.999
ASME6	How often do you watch on the Internet/TV/mobile phone/ DVD people who help someone?	0.890
ASME7	How often do you watch on the Internet/TV/mobile phone/ DVD people who openly talk about sex?	0.920
ASME8	How often do you watch on the Internet/TV/mobile phone/ DVD people who destroy someone else's belongings?	0.877
SE1	On the whole, I am satisfied with myself.	0.999
SE2	At times, I think I am no good at all.	0.836
SE3	I feel that I have several good qualities.	0.894
SE4	I can do things as well as most other people.	0.845
SE5	I feel I do not have much to be proud of.	0.896
SE6	I certainly feel useless at times.	0.849
SE7	I feel that I am a person of worth.	0.886
SE8	I wish I could have more respect for myself.	0.855
SE9	All in all, I am inclined to think that I am a failure.	0.881
SE10	I take a positive attitude toward myself	0.857
MI1	I try hard to act honestly in most things I do.	0.969
MI0	If no one is watching or will know, it does not matter if I do the right thing.	0.886
MI1	It is more important that people think you are honest than being honest.	0.851
MI2	If no one could find out, stealing a small amount of money or other things that no one will miss is okay.	0.820
MI3	If a cashier accidentally gives me RM 50 extra change, I usually act as if I did not notice it.	0.947
MI5	Lying and cheating are just things you have to do in this world.	0.843
MI6	Lying and cheating are just things you have to do in this world.	0.848
MI7	Doing things that some people might view as dishonest does not bother me.	0.903
MI8	If people treat me badly, I will treat them in the same manner.	0.757
MI19	I will go along with a group decision, even if I know it is morally wrong.	0.898
MI2	It is more important that people think you are honest than being honest.	0.824
MI20	Having moral values is worthless in today's society.	0.933

Table 2 *(Continue)*

Construct	Item of the construct	Outer loading
MI3	If no one could find out, stealing a small amount of money or other things that no one will miss is okay.	0.947
MI5	If a cashier accidentally gives me RM 50 extra change, I usually act as if I did not notice it.	0.843
MI6	Lying and cheating are just things you have to do in this world.	0.848
MI7	Doing things that some people might view as dishonest does not bother me.	0.903
MI8	If people treat me badly, I will treat them in the same manner.	0.757
MI9	As long as I make a decision to do something that helps me, it does not matter much if other people are harmed.	0.898
AB1	Yelled at others when they have annoyed you	0.997
AB10	Hurt others to win a game	0.887
AB 11	Become angry or mad when you do not get your way	0.789
AB 12	Use physical force to get others to do what you want	0.912
AB 13	Get angry or mad when you lose a game	0.753
AB 14	Got angry when others threatened you	0.703
AB 16	Used force to obtain money or things from others	0.900
AB 8	Made obscene phone calls for fun	0.910
AB 9	Hit others to defend yourself	0.782
AB 2	Had fights with others to show who was on top	0.802
AB 20	Gotten others to gang up on someone else	0.923
AB 21	Carried a weapon to use in a fight	0.925
AB 22	Yelled at others so they would do things for you	0.868
AB 3	Reacted angrily when provoked by others	0.756
AB 4	Taken things from other students	0.722
AB 6	Vandalized something for fun	0.914
AB 7	Had temper tantrums	0.775
AB 8	Damaged things because you felt mad	0.834
AB 9	Had a gang fight to be cool	0.461

Source: Authors' work

Discriminant Validity Analysis Using the Fornell and Lacker Criterion and Heterotrait-Monotrait Ratio (HTMT)

Two methods were utilized to assess the discriminant validity of the constructs, i.e., the Fornell-Larcker criterion (Table 3), as well as item loading and cross-loading in Table 4. The results showed that all the measures exhibited appropriate discriminant validity. Furthermore, the Fornell and

Lacker ratio should be less than 0.9 for the measurement model to demonstrate discriminant validity. As evident in Table 3, the HTMT ratios for all the comparisons of constructs ranged between 0.324 and 0.823 (below 0.9). Thus, the measurement model achieved satisfactory discriminant validity according to the HTMT criterion.

Table 3
Discriminant Validity using the Fornell and Lacker Criterion and HTMT

Fornell and Lacker	AB	ASME	MI	SE
AB	0.832	-	-	-
ASME	0.626	0.501	-	-
MI	-0.426	-0.328	0.863	-
SE	-0.716	-0.683	0.291	0.902
Heterotrait-Monotrait Ratio (HTMT)				
AB	-	-	-	-
ASME	0.620	-	-	-
MI	0.424	0.266	-	-
SE	0.708	0.438	0.285	-

Note. AB = Aggressive Behavior; MI = Moral Identity; ASME = Antisocial Media Exposure
Source: Authors' work

In addition, the cross-loading matrix was also employed as the third approach to assess discriminant validity. The process involved examining the item loadings of a construct on other constructs. For discriminant validity, the items of a particular construct should have higher loadings on their intended construct than other constructs (Straub et al., 2004). This study used the cross-loading method to validate the discriminant validity. Table 4 indicates that all items demonstrated the highest loadings on their respective intended constructs, thus confirming the discriminant validity of the measurement model. In short, the measurement model was verified and established, and the structural model assessment could be adopted in the next step.

Table 4
Cross loading result

Item	Aggressive behavior	Antisocial media exposure	Moral identity	Self-esteem
Item_ASME1	0.583	0.931	-0.240	-0.423
Item_ASME2	0.594	0.949	-0.284	-0.395
Item_ASME 3	0.473	0.756	-0.209	-0.337
Item_ASME 4	0.559	0.893	-0.240	-0.374

Table 4 *(Continue)*

Item	Aggressive behavior	Antisocial media exposure	Moral identity	Self-esteem
Item_ASME 5	0.476	0.760	-0.191	-0.374
Item_ASME 6	0.553	0.883	-0.239	-0.375
Item_ASME 7	0.512	0.818	-0.210	-0.408
Item_ASME 8	0.614	0.981	-0.263	-0.416
Item_SE1	-0.839	-0.543	0.338	1.171
Item_SE10	-0.599	-0.361	0.237	0.836
Item_SE2	-0.640	-0.366	0.234	0.894
Item_SE 3	-0.605	-0.413	0.299	0.845
Item_SE 4	-0.642	-0.362	0.246	0.896
Item_SE 5	-0.608	-0.391	0.256	0.849
Item_SE 6	-0.634	-0.400	0.243	0.886
Item_SE 7	-0.612	-0.385	0.255	0.855
Item_SE 8	-0.631	-0.371	0.260	0.881
Item_SE 9	-0.614	-0.380	0.245	0.857
Item_MI1	-0.413	-0.230	0.969	0.253
Item_MI10	-0.377	-0.244	0.886	0.288
Item_MI11	-0.362	-0.210	0.851	0.247
Item_MI12	-0.349	-0.208	0.820	0.242
Item_MI13	-0.403	-0.254	0.947	0.271
Item_MI15	-0.359	-0.232	0.843	0.231
Item_MI16	-0.361	-0.209	0.848	0.230
Item_MI17	-0.384	-0.249	0.903	0.272
Item_MI18	-0.322	-0.200	0.757	0.245
Item_MI19	-0.382	-0.237	0.898	0.260
Item_MI2	-0.351	-0.222	0.824	0.244
Item_MI20	-0.464	-0.281	1.089	0.348
Item_MI3	-0.355	-0.232	0.834	0.248
Item_MI4	-0.345	-0.237	0.810	0.220
Item_MI5	-0.352	-0.228	0.826	0.241
Item_MI6	-0.347	-0.222	0.816	0.238
Item_MI7	-0.357	-0.235	0.838	0.224
Item_MI8	-0.342	-0.257	0.803	0.220
Item_MI9	-0.329	-0.231	0.772	0.226
Item_AB1	1.025	0.583	-0.342	-0.791
Item_AB10	0.887	0.586	-0.391	-0.616
Item_AB11	0.789	0.486	-0.299	-0.579
Item_AB12	0.912	0.582	-0.419	-0.641

Table 4 *(Continue)*

Item	Aggressive behavior	Antisocial media exposure	Moral identity	Self-esteem
Item_AB13	0.753	0.535	-0.275	-0.514
Item_AB14	0.703	0.482	-0.231	-0.471
Item_AB16	0.900	0.567	-0.402	-0.638
Item_AB18	0.910	0.579	-0.470	-0.605
Item_AB19	0.782	0.497	-0.358	-0.547
Item_AB2	0.802	0.445	-0.316	-0.612
Item_AB20	0.923	0.596	-0.430	-0.633
Item_AB21	0.925	0.612	-0.399	-0.628
Item_AB22	0.868	0.539	-0.341	-0.636
Item_AB3	0.756	0.444	-0.233	-0.576
Item_AB4	0.722	0.406	-0.418	-0.543
Item_AB6	0.914	0.569	-0.430	-0.642
Item_AB7	0.775	0.524	-0.232	-0.567
Item_AB8	0.834	0.509	-0.357	-0.627
Item_AB9	0.461	0.241	-0.386	-0.345

Source: Authors' work

Data Collection

The data collection involved a multi-stage approach. Firstly, ethical clearance was obtained from the Institutional Review Board (IRB) of the University's Ethics Committee for Research Involving Human Subjects (JKEUPM). The survey started in December 2022. A purposive and snowball sampling was used to recruit online gamers on gaming platforms such as Quora, Facebook forums, MMORPG forums, Instagram, WhatsApp or Telegram gaming forums, and various gaming events. The questionnaire encompassed variables on demographic characteristics, ASME, MI, SE, and aggressive behavior that could be completed in 20 minutes. A cover letter was attached at the beginning of the questionnaire to explain the study objective. The researcher's email address and mobile number were also included if the participants needed to obtain further information about the study.

RESULTS

Participant Demographics

Table 5 shows the baseline demographic characteristics of the study participants. Among the 384 participants, two-thirds (67.2%) were males. Most were young adults aged between 18 and 24 (71.4%), with only a small number of older adults aged 32 and above (6.8%). All participants reported that they had experienced ASME.

Table 5
Demographics of study participants (n = 384)

Demographic Profile	n	%
Gender		
Female	126	32.8
Male	258	67.2
Age		
18–24	274	71.4
25–31	84	21.9
32 and above	26	6.8
Antisocial Media Exposure (e.g., online games and film)		
Yes	384	100.0

Source: Authors' work

Path Analysis

Direct Effect Analysis

Figure 2 shows the results of a direct effect analysis examining the relationships between ASME, MI, SE, and aggressive behavior.

Table 6 shows a path coefficient result; a negative relationship was observed between MI, SE, and aggressive behavior (β = –0.14, p < 0.00; β = –0.27, p < 0.00), indicating that aggressive behavior would decrease when MI and SE increased. A positive relationship was also found between ASME and aggressive behavior (β = 0.25, p < 0.000). In other words, when ASME increases, aggressive behavior also increases. These results align with previous findings reported by Shafti et al. (2021) and Przybylski and Weinstein (2019), who all reported MI and SE as important factors in reducing aggressive behavior. Meanwhile, ASME was an important predictor of aggressive behavior.
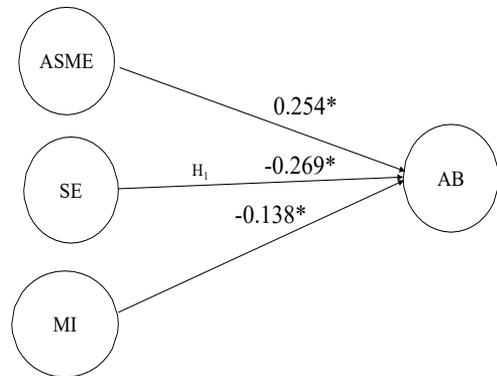

Figure 2. Path Coefficients Analysis. The main entries are standardized coefficients, β (Sig. level *p<0.05, ** *p<0.05).
Source: Authors' work

Furthermore, coefficients of determination (R^2), effect size (f^2), and predictive relevance (Q^2) were tested in the analysis. The f^2 value for the relationship between MI and AB (0.051) was lower than 0.15 (Hair et al., 2017), indicating a small effect. However, the value of f^2 for the relationship between ASME and AB (0.185) was higher than 0.35, revealing a significant effect. Meanwhile, including R^2

Table 6
Path coefficients

Hypotheses	Coefficients	Std Error	*t*-values	*p*-values	R^2	f^2	Q^2	Decision
ASME -> AB	0.254	0.041	6.139	0.000		0.185		Supported
SE -> AB	-0.269	0.037	7.300	0.000		0.149		Supported
MI -> AB	-0.138	0.033	4.223	0.000		0.351		Supported
IV predict DV							0.681	Highly predictive
DV predict IV					0.694			Highly predictive

Note. R^2 = coefficients of determination; f^2 = effect size; Q^2 = predictive relevance; Significance Level = $p > 0.05$

Source: Authors' work

in the analysis contributed to 70% of the variance in aggressive behavior. In general, the model's overall predictive strength was considered highly acceptable ($R^2 = 0.694$; Henseler & Chin, 2010). In addition, the Q^2 value for AB was 0.681. Any value higher than zero indicates a good predictive relevance of the model (Byrne, 2016). In other words, independent variables in this model could explain a significant portion of the variance in aggressive behavior.

DISCUSSION

This study examined the factors influencing aggressive behavior in online gamers, focusing on Antisocial Social Media Exposure (ASME), Self-Esteem (SE), and Moral Identity (MI). All three hypotheses received significant support.

The study results revealed a significant and positive association between ASME and aggressive behavior, concordance with prior research (Przybylski & Weinstein, 2019; Teng et al., 2020). Several mechanisms might explain this association. Online environments saturated with negativity can desensitize individuals to aggressive behavior (Ferguson & Kilpatrick, 2009). Witnessing such behavior on social media might be seen as a model for responding to conflict (Bandura, 1977). Additionally, exposure to violent content can trigger feelings of anger and hostility (Anderson & Bushman, 2002), increasing the risk of impulsive and aggressive responses.

However, Self-Esteem (SE) acts as a protective factor against aggressive behavior. Individuals with high Self-Esteem are less susceptible to internalizing negative messages from online content (Rosenberg, 1965). This resilience allows them to favor assertive communication styles over resorting to aggression when faced with conflict or frustration. Similarly, a robust Moral Identity (MI) also buffers against aggressive behavior. Gamers with a well-developed moral compass are more likely to critically evaluate aggressive

messages embedded in social media content (Shaffer et al., 2016). This critical evaluation process helps them distance themselves from negativity and avoid internalizing aggressive messaging. Additionally, their strong moral compass guides them to behave in ways that align with their internal ethical standards, reducing the likelihood of engaging in aggressive behavior that contradicts their core values.

Nevertheless, our findings also show a different light on aggressive behavior in the Malaysian context, whereby two personal factors in the GAM, i.e., SE and MI, can decrease the level of aggressive behavior among local online gamers. This finding is very important to stakeholders in developing suitable strategies to mitigate various delinquencies among local online gamers. Apart from that, multinational companies (MNCs) involved in producing and marketing online games in Malaysia must revise their policies to avoid the increasing rates of aggressiveness among local online gamers, given the evidence linking exposure of antisocial elements to aggressive behavior. The Malaysian Communications and Multimedia Commission (MCMC) plays a significant in ensuring the implementation and enforcement of relevant policies that can reduce aggressive tendencies among gamers.

CONCLUSION

The findings of this study underscore the pivotal role that Antisocial Media Exposure (ASME) plays in fostering aggressive behavior. This groundbreaking evidence not only enriches the General Aggression Model (GAM) by integrating a new social factor but also illuminates the intricate cognitive processes within the GAM. The study's findings reveal that Moral Identity (MI) and Self-Esteem (SE) serve as cognitive buffers, mitigating the likelihood of aggressive outcomes in response to ASME. From a practical perspective, these findings carry profound implications for Malaysia's rapidly evolving online landscape. They substantiate concerns about the capacity of social media platforms to escalate negativity and aggression, especially among Malaysia's young demographic with high social media usage rates (Ho et al., 2022). It calls for more rigorous content moderation practices by platforms like Facebook and YouTube, notably popular in Malaysia (Kemp, 2023). Moreover, implementing robust age verification systems can limit minors' access to harmful content, providing an extra layer of protection for this vulnerable group (UNICEF Malaysia, 2021). The study also underscores the protective role of MI and SE. This insight can guide the creation of targeted educational interventions within Malaysia's national curriculum.

By equipping students with digital media literacy skills, we can enable them to critically evaluate online content and responsibly navigate the digital world (Ministry of Education Malaysia, 2023). Furthermore, mental well-being programs that nurture self-esteem and foster positive online communities can enhance resilience against negative influences. Public awareness campaigns can enlighten parents

and the broader population about the risks of ASME and strategies for responsible social media use, potentially capitalizing on existing initiatives by the MCMC (2021). Lastly, forging partnerships with social media platforms is essential. Through collaboration, policymakers and platforms can devise evidence-based interventions to stem the tide of harmful content and foster positive online interactions. This comprehensive approach can assist Malaysia in establishing a safer and more responsible online environment, promoting positive social interactions and curbing aggression incited by social media. This study, therefore, serves as a beacon, guiding us toward a more harmonious digital future.

Implication for Theory and Practice

This study has significant implications for both the theoretical understanding and practical management of aggression, particularly within the context of the General Aggression Model. This research offers valuable insights that refine the General Aggression Model (GAM) when applied to online gaming environments. Traditionally, GAM emphasizes internal cues (frustration, anger) and external cues (witnessing aggression) as triggers leading to aggressive behavior (Berkowitz, 1962). This study broadens the understanding of external cues by highlighting the role of Antisocial Social Media Exposure (ASME). Frequent exposure to negativity, violent content, and hate speech online can be seen as a social learning context within the GAM framework. Through repeated exposure,

gamers might observe and learn aggressive behaviors displayed on social media, potentially increasing their likelihood of adopting them. It suggests that the online environment can be a significant source of cues that trigger aggressive tendencies.

Furthermore, the study elevates the role of individual differences within GAM. Traditionally, the model focuses on immediate situational factors leading to aggression. This research emphasizes the role of moderators like Self-Esteem and Moral Identity. These personal characteristics influence how individuals process and respond to social learning experiences like ASME exposure. Individuals with high Self-Esteem have a more positive self-perception, making them less susceptible to internalizing negativity online. This resilience allows them to favor assertive communication styles over resorting to aggression in online conflicts. Gamers with a well-developed Moral Identity possess a strong sense of right and wrong. This strong moral compass guides their online interactions and reduces the likelihood of engaging in aggressive behavior that contradicts their core values. They might critically evaluate aggressive messages, allowing them to distance themselves from the negativity and avoid internalizing aggressive messaging.

This study suggests a more nuanced understanding of GAM in online gaming contexts. It expands the model by identifying antisocial social media exposure as a potential social learning context that triggers aggressive behavior. It also highlights the

crucial role of individual differences like self-esteem and moral identity in moderating the influence of social learning experiences on aggressive tendencies. These findings encourage further exploration of how the GAM can be adapted to better capture the complex interplay between online environments, individual characteristics, and aggressive behavior in online gamers.

This study enhances the available evidence of the association between ASME and aggressive behavior. ASME was established to impact the portrayal of aggressive behavior significantly. In contrast, MI and SE could be protective against individuals' reactions to violent online content. In conclusion, these findings can be instrumental in the development of digital detox programs aimed at preventing online media-related violence. Digital detox programs have emerged as a potential solution to curb online media-related violence. However, these programs can offer more than just reduced screen time. By incorporating evidence-based strategies, digital detox programs can equip participants with the skills to navigate the online world more critically and foster a more positive online environment. One approach involves enhancing media literacy. Digital detox programs can integrate training that empowers participants to evaluate online content critically. This training could focus on identifying red flags associated with Antisocial Social Media Exposure (ASME), such as inflammatory language, the normalization of violence, and the promotion of hate speech. By developing these media literacy skills, participants become more responsible online consumers, actively seeking prosocial content that contributes to a richer online experience.

Digital detox programs should not be seen solely as periods of isolation. These programs can be designed to foster a sense of community and belonging even with reduced online engagement. Support groups or online forums designed for digital detox participants can provide a safe space for individuals to share their experiences, challenges, and successes. It fosters a sense of accountability and motivates participants to stay committed to their digital detox goals. A crucial aspect of digital detox programs involves encouraging participants to explore alternative activities and hobbies. It could involve reconnecting with nature, engaging in physical activities, or pursuing creative endeavors. By developing a more prosperous offline life, participants are less likely to feel the pull of potentially harmful online environments.

Finally, digital detox programs should be realistic and acknowledge the potential for relapse. Equipping participants with relapse prevention strategies is crucial for long-term success. It might involve developing personalized trigger lists identifying situations where they are more likely to return to unhealthy online habits. Participants can then create coping plans to address those triggers and maintain their digital detox goals. By incorporating these research-informed strategies, digital detox programs can move beyond simply reducing screen time and offer a more holistic

approach to curbing online media-related violence. These programs can empower individuals to navigate the online world more critically, fostering a positive and responsible online environment.

Several proposals are proposed for expanding future research. Firstly, broadening the demography to include players from diverse geographical places or countries is recommended. This methodology has the potential to provide a more thorough comprehension of the topic since it would incorporate a wider range of cultural and socioeconomic circumstances. Additionally, it is advisable to investigate other variables that influence aggression conduct. It can provide more insights into such behavior's fundamental origins and triggers, enhancing the current knowledge base. Furthermore, it is suggested that longitudinal investigations be carried out. Longitudinal studies, which examine people over long durations, have the potential to illuminate the progression of aggressive behaviors, providing a dynamic viewpoint that cross-sectional research may need to improve. Finally, it is advisable to use qualitative research methodologies. Qualitative approaches differ from quantitative methods in exploring humans' subjective experiences and views rather than focusing on numerical data and statistical analysis.

It has the potential to provide a more profound comprehension of the events being examined, providing subtle insights that quantitative data could fail to encompass. If these ideas are put into practice, they can greatly improve the quality and breadth of future study in this area. Their approach to studying aggressive behavior among gamers is broad, aiming to foster a comprehensive knowledge of the subject.

Limitations and Recommendations for Future Research

There are several limitations to this study. Firstly, the questionnaire survey in this study only focused on gamers in Malaysia, thus limiting the generalizability of the results. Furthermore, it will be valuable to examine the long-term effects of violent media exposure and moral identity on aggressive behavior in future studies. Longitudinal research can help to establish causal relationships and provide insights into the developmental trajectories of aggression, as well as the potential roles of moral identity and self-esteem in mitigating or exacerbating these effects over time. As we employed a purposive and snowball sampling approach of individuals living in the Klang Valley, our study results would not be generalizable to the entire community of online gamers in Malaysia.

ACKNOWLEDGMENTS

REFERENCES

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, *53,* 27-51. https://doi.org/10.1146/annurev. psych.53.100901.135231

Aquino, K., Freeman, T. E., & Lindquist, C. A. (2023). The antecedents of moral identity: A meta-analytic review. *Journal of Personality and Social Psychology*, *124*(1), 18-48.

Balakrishnan, V., & Fernandez, T. (2018). Self-esteem, empathy, and their impacts on cyberbullying among young adults. *Telematics and Informatics*, *35*(7), 2028-2037. https://doi.org/10.1016/j.tele.2018.07.006

Bandura, A. (1977). *Social learning theory.* Prentice-Hall.

Barlett, C. P., & Coyne, S. M. (2019). Electronic bullying: An overview of the literature. *Journal of Educational Psychology*, *111*(3), 414-428.

Becker, J.-M., Cheah, J.-H., Gholamzade, R., Ringle, C.M., & Sarstedt, M. (2023). PLS-SEM's most wanted guidance. *International Journal of Contemporary Hospitality Management, 35*(1), 321-346. https://doi.org/10.1108/IJCHM-04-2022-0474

Berkowitz, L. (1962). *Aggression: A social psychological analysis.* McGraw-Hill.

Black, J. E., & Reynolds, W. M. (2016). Development, reliability, and validity of the Moral Identity Questionnaire. *Personality and Individual Differences, 97,* 120-129. https://doi.org/10.1016/j.paid.2016.03.041

Brown, A. (2021). The rise of antisocial behavior on social media. *Journal of Social Media Studies, 12*(3), 45-60.

Browne, R. H. (1995). On the use of a pilot sample for sample size determination. *Statistics in Medicine*, *14*(17), 1933-1940. https://doi.org/10.1002/sim.4780141709

Bushman, B. J. (2017). Aggressive cues: Weapons and violent media. In B. J. Bushman (Ed.), *Aggression and violence: A social psychological perspective* (pp. 184-198). Routledge/Taylor & Francis Group.

Bushman, B. J., Jamieson, P. E., Weitz, I., & Romer, D. (2013). Gun violence trends in movies. *Pediatrics*, *132*(6), 1014-1018. https://doi.org/10.1542/peds.2013-1600

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). Routledge. https://doi.org/10.4324/9781315757421

Carnagey, N. L., & Anderson, C. A. (2003). Theory in the study of media violence: The General Aggression Model. *Communication Research, 30*(1), 6-32.

Coyne, S. M. (2016). Effects of viewing relational aggression on television on aggressive behavior in adolescents: A three-year longitudinal study. *Developmental Psychology*, *52*(2), 284-295. https://doi.org/10.1037/dev0000068

den Hamer, A. H., Konijn, E. A., Plaisier, X. S., Keijer, M. G., Krabbendam, L. C., & Bushman, B. J. (2017). The content-based media exposure scale (C-ME): Development and validation. *Computers in Human Behavior*, *72*, 549-557. https://doi.org/10.1016/j.chb.2017.02.050

Falla, D., Romera, E. M., & Ortega-Ruiz, R. (2021). Aggression, moral disengagement and empathy. A longitudinal study within the interpersonal dynamics of bullying. *Frontiers in Psychology, 12*, Article 703468. https://doi.org/10.3389/fpsyg.2021.703468

Ferguson, C. J., & Dyck, D. (2012). Paradigm change in aggression research: The time has come to retire the General Aggression Model. *Aggression and Violent Behavior, 17*(3), 220-228. https://doi.org/10.1016/j.avb.2012.02.007

Ferguson, C. J., & Kilpatrick, D. A. (2009). The public health risks of media violence: A meta-analytic review. *Journal of Adolescent Health, 46*(4), 274-289.

Ferguson, C. J., Rueda, M. R., & Bushman, B. J. (2011). The antisocial effects of violent video

games on youth: A progress report. *Perspectives on Psychological Science, 6*(1), 65-72.

Gagnon, J., Quansah, J. E., & McNicoll, P. (2022). Cognitive control processes and defense mechanisms that influence aggressive reactions: Toward an integration of socio-cognitive and psychodynamic models of aggression. *Frontiers in Human Neuroscience, 15,* Article 751336. https://doi.org/10.3389/fnhum.2021.751336

Gentile, D. A., Greer, A., Wachs, B. D., Moffitt, A. J., & Walsh, B. T. (2020). Exploring the mechanisms relating media use to self-esteem: A review and directions for future research. *Psychological Bulletin, 146*(2), 180-215.

Gentile, D. A., Twenge, J. M., Campbell, W. K., & Rubinstein, E. (2020). The effects of social media use on self-esteem: A meta-analysis. *Psychological Science, 31*(1), 60-71.

Gini, G., Thornberg, R., Bussey, K., Angelini, F., & Pozzoli, T. (2021). Longitudinal links of individual and collective morality with adolescents' peer aggression. *Journal of Youth and Adolescence, 51*, 524-539. https://doi.org/10.1007/s10964-021-01518-9

Graham, J., Noack, L., Harmon-Jones, E., & Sinnott-Armstrong, W. (2011). Moral hypocrisy and dispositional moral identity. *Journal of Personality and Social Psychology, 101*(4), 743-756.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.

Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). Advanced diagnostics for multiple regression. In *Multivariate Data Analysis* (7th ed., pp. 761-775). Prentice Hall.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM*; 3rd ed.). Sage Publications.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

Hartmann, T., Krakowiak, K. M., & Tsay-Vogel, M. (2014). How violent video games communicate violence: A literature review and content analysis of moral disengagement factors. *Communication Monographs*, *81*(3), 310-332. https://doi.org/10.1080/03637751.2014.922206

Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(1), 82-109. https://doi.org/10.1080/10705510903439003

Hetsroni, A. (2007). Four decades of violent content on prime-time network programming: A longitudinal meta-analytic review. *Journal of Communication*, *57*(4), 759-784. https://doi.org/10.1111/j.1460-2466.2007.00367.x

Ho, S. S., Mohamed, H., & Ibrahim, Y. (2022). Social media usage patterns and mental health among Malaysian adolescents: A systematic review. *Journal of Educational and Developmental Psychology, 12*(2), 189-204.

Huang, X., & Fan, P. (2022). The dark side of social media in the workplace: A social comparison perspective. *Computers in Human Behavior*, *136*, Article 107377. https://doi.org/10.1016/j.chb.2022.107377

Ismail, N., Tajjudin, A. I., Jaafar, H., Nik Jaafar, N. R., Baharudin, A., & Ibrahim, N. (2021). The relationship between internet addiction, internet gaming and anxiety among medical students in a Malaysian public university during Covid-19 pandemic. *International Journal of Environmental Research and Public Health*, *18*(22), Article 11870. https://doi.org/10.3390/ijerph182211870

Jiang, X., Li, X., Dong, X., & Wang, L. (2022). How the Big Five personality traits related to aggression from perspectives of the benign and malicious envy. *BMC Psychology, 10,* Article 203. https://doi.org/10.1186/s40359-022-00906-5

Johnson, L., & Johnson, M. (2020). The development of self-esteem: A longitudinal study. *Journal of Personality and Social Psychology, 118*(2), 421-438.

Jones, T., Cunningham, P. H., & Gallagher, K. (2014). Violence in advertising: A multilayered content analysis. In N. J. Rifon, M. B. Royne, & L. Carlson (Eds.), *Advertising and violence: Concepts and perspectives* (pp. 255-289). M E Sharpe.

Kemp, S. (2023). Digital 2023: Malaysia. *DataReportal.* https://datareportal.com/reports/digital-2023-malaysia

Khurana, A., Bleakley, A., Ellithorpe, M. E., Hennessy, M., Jamieson, P. E., & Weitz, I. (2019). Sensation seeking and impulsivity can increase exposure to risky media and moderate its effects on adolescent risk behaviors. *Prevention Science, 20*(5), 776-787. https://doi.org/10.1007/s11121-019-0984-z

Kieser, M., & Wassmer, G. (1996). On the use of the upper confidence limit for the variance from a pilot sample for sample size determination. *Biometrical Journal, 38*(8), 941-949. https://doi.org/10.1002/bimj.4710380806

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*(3), 607-610. https://doi.org/10.1177/001316447003000308

Malaysian Communications and Multimedia Commission. (2020). *Industry Performance Report 2020*. https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/MCMC-IPR_2020.pdf

Malaysian Communications and Multimedia Commission. (2021). *Media Matters* (Vol. 4; Advance Copy; Digital Society Research Report) https://repo.uum.edu.my/id/eprint/30381/1/MCMC_DSRG_04_2021%2044-52.pdf

Martin, B. A., & Collins, B. A. (2002). Violence and consumption imagery in music videos. *European Journal of Marketing, 36*(7/8), 855-873. https://doi.org/10.1108/03090560210430845

Ministry of Education Malaysia. (2023). *Teks Perutusan Ketua Pengarah Pendidikan Malaysia Sempena Sambutan Hari Guru Peringkat Kebangsaan Kali ke-53 Tahun 2024* [Address from the Director of Education Malaysia in Commemoration of the 53rd National Teacher's Day Celebration 2024]. https://www.moe.gov.my/storage/files/shares/Teks%20Ucapan%20Dan%20Slide/2024/Perutusan%20Ketua%20Pengarah%20Pendidikan%20Malaysia%20Hari%20Guru%202024.pdf

Nasser, N. M. A., Mohamad, M., Chan, L. F., Suppiah, S., Yusoff, A. N., Abd Hamid, A. I., Ismail, W. S. W., & Sim, S. T. (2021). The task-based fMRI study on neural activity of internet gaming addiction among Malaysian adolescents: A stroop color task. *Neurology Asia, 26*(3), 561-574. https://doi.org/10.54029/2021msd

Przybylski, A. K., & Weinstein, N. (2019). Violent video game engagement is not associated with adolescents' aggressive behaviour: Evidence from a registered report. *Royal Society Open Science, 6*(2), Article 171474. https://doi.org/10.1098/rsos.171474

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*(2), 159-171. https://doi.org/10.1002/ab.20115

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton University Press. https://doi.org/10.1515/9781400876136

Rosenberg, M. (1979). *Conceiving the self.* Basic Books.

Schultze-Krumbholz, A., Hess, M., Pfetsch, J., & Scheithauer, H. (2018). Who is involved in cyberbullying? Latent class analysis of cyberbullying roles and their associations with aggression, self-esteem, and empathy. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 12*(4), Article 2. https://doi.org/10.5817/CP2018-4-2

Shaffer, D. R., Kipp, K., Wood, E., & Willoughby, T. (2016). *Developmental psychology: Childhood and adolescence* (1st ed.). Cengage Learning.

Shafti, M., Taylor, P. J., Forrester, A., & Pratt, D. (2021). The co-occurrence of self-harm and aggression: A cognitive-emotional model of dual-harm. *Frontiers in Psychology, 12,* Article 586135. https://doi.org/10.3389/fpsyg.2021.586135

Simanjuntak, A., Rahmawati, I. S., & Purnamawati, H. (2021). The effects of online game violence exposure on teenagers' aggressive behaviour. *SOSHUM Jurnal Sosial dan Humaniora, 11*(1), 69-78.

Skobkareva, A. (2020). *Aggression in dark personalities: The role of self-esteem* [Doctoral dissertation, University of Windsor (Canada)]. Electronic Theses and Dissertations. https://scholar.uwindsor.ca/etd/8483

Smith, J. (2019). Exploring the antecedents of moral identity: A socio-cultural perspective. *Journal of Moral Education, 48*(4), 490-505.

Strasburger, V. C., Jordan, A. B., & Donnerstein, E. (2010). Health effects of media on children and adolescents. *Pediatrics*, *125*(4), 756-767. https://doi.org/10.1542/peds.2009-2563

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13*, 380-427. https://doi.org/10.17705/1CAIS.01324

T'ng, S. T., & Pau, K. (2020). Validation of a translated Internet Gaming Disorder Scale (short form) and measurement invariance across sex groups in Malaysian samples. *Current Psychology*, *39*(6), 1977-1989. https://doi.org/10.1007/s12144-020-00668-6

Tedeschi, R. G., Feldstein Russell, D. W., & Park, C. L. (2021). Cyberbullying and adolescent moral reasoning: The mediating role of self-efficacy and moral emotions. *Journal of Youth and Adolescence, 50*(11), 2062-2076.

Teng, Z., Nie, Q., Zhu, Z., & Guo, C. (2020). Violent video game exposure and (cyber) bullying perpetration among Chinese youth: The moderating role of trait aggression and moral identity. *Computers in Human Behavior*, *104*, Article 106193. https://doi.org/10.1016/j.chb.2019.106193

UNICEF Malaysia. (2021). *UNICEF Malaysia Annual Report 2021: #ForEveryChild, a future beyond COVID-19.* United Nations Children's Fund. https://www.unicef.org/malaysia/media/3756/file/Annual%20Report%202021.pdf

Williams, R. (2022). Cultural and demographic factors in social behavior: An integrative review. *Journal of Cross-Cultural Psychology, 53*(1), 1-20.

Yunus, M. Y. H., Yusoff, N. H., & Ng, C. Y. (2021). Factors influencing the involvement of Malaysian youths in Massively Multiplayer Online Role-Playing Games (MMORPGS). *Journal of Techno-Social*, *13*(1), 59-67. https://doi.org/10.30880/jts.2021.13.01.008

Zhang, Y. (2023). Understanding the impact of demographic factors on social behavior: A meta-analysis. *Journal of Social and Personal Relationships*, *40*(1), 7-25.

The Digital Dynamics of Political Engagement Among Filipino Youth: Examining Participation in Social Media Platforms

Joseph Ching Velasco*, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan, Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

Department of Political Science and Development Studies, De La Salle University, Manila, Philippines

ABSTRACT

Youth participation has historically played a critical role in shaping the political landscape of the Philippines. Being immersed in digital technology, Generation Z has mostly shifted their activism from the streets to the virtual realm of diverse social media platforms. Through digital interactions, they champion causes and establish political dialogue by initiating socio-political movements online. This study utilized the Social Media Political Participation Scale to put into perspective the online political behaviors of Filipinos who are part of Generation Z. Specifically, the study centers on Metro Manila, comparing their engagements across various online platforms, including Facebook, Twitter, Instagram, YouTube, and TikTok. Social media engagements were divided into four aspects: latent engagement, follower engagement, expressive engagement, and system engagement, to identify the digital activities that encourage respondents to participate. The study highlights the nuanced dynamics of digital political engagement and its role in shaping political participation. Results reveal that only Instagram and TikTok encourage respondents to participate politically on different social media platforms.

Keywords: Expressive engagement, follower engagement, generation Z, latent engagement, online engagements, political participation, social media, system engagement

INTRODUCTION

The ubiquitous impact of social media on society has resulted in significant social transformations. It enabled global communication and seamless access to

information. As digital capabilities expand, social media has become prominent in daily life. It subsequently gave rise to Generation Z, a tech-savvy cohort deeply immersed in these platforms (Kim et al., 2015; Radut, 2018). This generation stands as the most engaged age group on social media, showcasing technological competence and a commitment to leveraging platforms. Despite the increase in the said generation's political participation, several studies discovered that this generation may be politically apathetic, politically performative, or suffering from political fatigue. It stems from the rising competition for online attention, fame, and social capital in social media networks (Baym & Evans, 2022), as observed in the social media landscape of the Philippines, where innumerable individuals have become prominent for creating videos with falsified information regarding COVID-19 and the presidential campaign of Ferdinand Marcos Jr. (Manavis, 2020; Santos, 2022).

Nevertheless, the surge in performative political acts among Generation Z has transformed social media into a new avenue for engagement, fostering social and political conversations. In the last few years, the hashtag function has been utilized to express dissents related to the state's aggression towards different sectors of society, particularly #NoToExtraJudicialKillings for the condemnation of *Oplan Tokhang*, #JunkAntiTerrorBill for the dangerous provisions of the Anti-Terrorism Bill, and #NasaanAngPangulo for former President Rodrigo Duterte's absence during a typhoon

(Alawi et al., 2020; Madarang, 2020; Macaraig & Hameleers, 2022).

As such, it is evident that existing scholarship continues to advance the notion that social media is a legitimate channel for Generation Z to engage in political participation, which empowers active involvement in the political sphere. However, political involvement comes in many forms. Considering the diverse spectrum of political participation, there is still a dearth of scholarship exploring the actual behavior this generation exhibits online. Likewise, several studies only associate social media political participation with platforms limited to Facebook and Twitter, leaving other platforms unexplored, resulting in limited information and understanding of the capability of networks like TikTok, Instagram, and YouTube as a conduit for political expression, mobility, and participation. Thus, the study aims to uncover the different forms of Generation Z's political behavior and its impact on the Philippine political landscape.

LITERATURE REVIEW

Looking at the data over the years, the Philippines topped the world six times with the most time spent on social media and was only outranked in 2021 by South Africa. In the Digital Global Overview Report for January 2023, there were already 84.45 million social media users, equating to approximately 72.5% of the country's total population. Filipinos spend at least three hours and 43 minutes on social media daily. Kemp (2023) reported that

the top social media platforms in the country are Facebook, with an average of 75% active users, 50.5% on YouTube, 49.4% on TikTok, 16.7% on Instagram, and 12.6% on Twitter. From this estimate, more than 50% of these active users belong to Generation Z. Ardèvol-Abreu et al. (2020) discovered that Generation Z's utilization of social media is driven by the sense of belongingness it provides to the said cohort. The engagement of different social groups and online communities on social media platforms influenced Generation Z to express themselves and support certain causes or political advocacy. Lim (2009) further highlights that social media does increase political participation, as it mobilizes individuals and facilitates their political agendas.

In 2015, students from the University of the Philippines initiated an online movement in support of a vulnerable indigenous group known as Lumads (Marcaida, 2020). Their primary hashtag, #StopLumadKillings, was used by thousands of Twitter users, resulting in various information-sharing initiatives and other necessary actions. At the same time, due to restricted healthcare access and incessant lockdowns, some individuals have turned to social media to convey their frustrations and urge the government to administer more comprehensive and humane policies. Similarly, during the 2022 national elections campaign, candidates and voters also relied on producing and consuming visual content on applications such as TikTok and YouTube (Arugay, 2022).

Hite (2010) argues that socioeconomic factors such as household income are to be considered influential toward the political engagement of citizens. Ultimately, she inferred that more financially secure constituents have lesser political efficacy and do not gravitate greatly toward political affairs compared to those who are not as financially stable. Moreover, the sex of individuals as an influential factor has long been incorporated by scholars who have tackled political participation. Bimber et al. (2021) established that similarly, both males and females exhibit their engagement on social media through information dissemination and responding to content online.

Adedokun (2022) further highlighted how social media empowers citizens to be vocal about certain government actions. The study demonstrates the dynamics between social media and political participation. Given that the interface between political participation and social media is becoming more extensive, Mendenhall and Sodani (2021) assert that new platforms should be examined. As a case in point, their study, which focused solely on TikTok, asserts that individuals choose to produce political content on the said platform since its features provide an efficient and convenient experience for younger audiences. Meanwhile, Mariano et al. (2021) found that, for the 2022 national elections, Facebook was generally the primary source of information compared to other social media platforms. Given these initial forays into examining political participation and

social media, our study examined the relationship between digital platforms such as social media and political engagement. Our study builds upon initial investigations of social media and political participation to examine the potential influence of digital platforms on political engagement and participation.

Numerous academic disciplines have explored the notion of political participation. However, digital innovations such as social media have necessitated a continuous examination of its manifestations. As inferred from Hosseinmardi et al. (2021), involvement in political affairs on social media platforms such as YouTube is not limited to nor bounded by simply viewing content. This study, therefore, highlights that social media is not merely regarded as an ordinary tool for participation but rather as an instrument that facilitates deep engagement in political affairs. The insights derived from this study would facilitate further exploration of potential future trends in relation to political participation to effectively and constructively leverage different platforms. In examining online political participation, different forms of engagement were considered, as shown in Table 1, primarily sourced from Waeterloos et al. (2021).

Generation Z's Political Participation in Social Media

Generation Z's profound engagement on social media platforms has transformed them into influential players in political discourse. This study employs the Social Media Participation Scale to delve into the nuances of political behavior within the digital landscape, focusing on the distinctive ways Generation Z actively shapes the political narrative. The study delimits Generation Z as individuals born between 1993 and 2005 and considers that the influence of five major social media platforms significantly drives them to participate politically. The researchers employed a multi-faceted approach, seeking to answer three pivotal questions that shape the narrative of Generation Z's political involvement on social media. More specifically, the study seeks to respond to these questions:

1. How does respondents' interaction with political content on social media translate to active and passive political engagement?

2. What factors play a role in shaping respondents' decision to participate politically through social media platforms?

3. How do the different types of online engagement contribute to the respondents' political participation and political behavior?

The study's research questions aim to explicate the intricacies inherent in Generation Z's interaction with political content across social media platforms. Through a quantitative approach, the article endeavors to discern the patterns, preferences, and frequency that characterize

Generation Z's engagement with political content. Moreover, the research gauges the implications of these interactions on real-world political engagement, which involves electoral participation and community organizing. In addition, examining factors influencing Generation Z's decisions to participate actively in political discourse through social media reveals the complex interplay among technology, sociopolitical conditions, and even personal motivations. This analytical endeavor aims to enhance the nuanced comprehension of the motivating forces underlying Generation Z's involvement with political content in the online domain. Therefore, it ascertains whether social media operates as a catalyst for constructive political involvement or if it fosters polarization and disengagement.

Conceptual Framework

The widespread use of social media platforms such as Facebook, Twitter, Instagram,

YouTube, and TikTok increased the political engagement of the youth in Metro Manila through the ability of these platforms to facilitate social influence, mobilize political action, increase accountability, and foster participation. Such participation may lead to four main forms of engagement: latent, follower, expressive, and system. It is primarily derived from Waeterloos et. al. (2021) (Table 1). Latent engagement is the passive consumption of political material, whereas expressive engagement is the active sharing and voicing of political beliefs. Follower engagement, on the other hand, pertains to the integration of the public figure and audiences in any action initiated politically. Meanwhile, expressive engagement is directed toward a wider audience to convey one's viewpoint. Finally, system engagement is a more structured and institutionalized form of political participation (Figure 1).
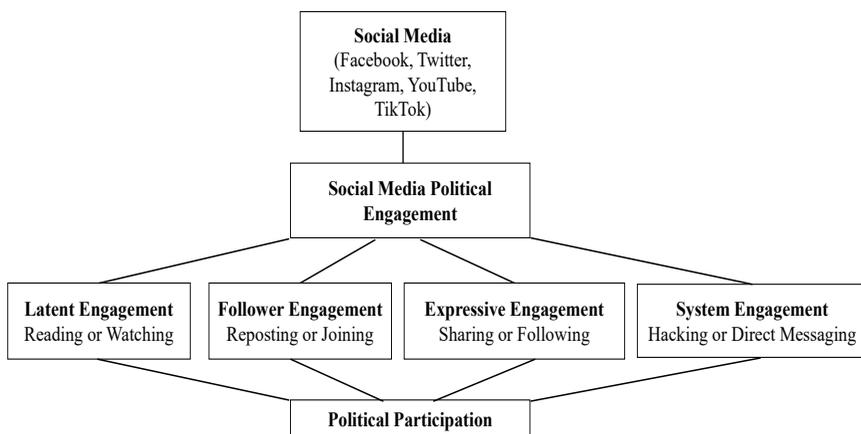

Figure 1. Social media engagement in relation to political participation

Source: Authors' work

Joseph Ching Velasco, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan,
Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

Table 1

Forms of engagement and sample items

Form of Engagement	Example of Manifestation
Latent engagement	Read the whole posts related to political issues and news on Facebook
	Watched videos and lives linked to political issues and news on Instagram, YouTube, or TikTok
Follower engagement	Signed petitions that were seen on Facebook, Twitter, Instagram, YouTube, or TikTok
	Became a member of a politically related group or community on Facebook, Twitter, or YouTube
Expressive engagement	Posted a text, photo, video, or other forms of creative media to express my opinions on Twitter politically
	Publicly retweeted or reposted politically related information and news
	Watched online protests and engaged in political discourse in the comment section
System engagement	Participated in the mass emailing activity of politicians or any government agency to express political dissent
	Privately messaged a politician, government agency, or any political organization and figure to express political suggestions
	Hacked a website or social media account of a political figure to achieve a particular political objective

Source: Waeterloos et al. (2021)

MATERIALS AND METHODS

Purposive sampling was utilized to recruit 300 eligible survey respondents. Specifically, the study looked into Generation Z respondents in Metro Manila who are social media users, most of whom are university students. The study considered the homogeneity of respondents to pool their shared characteristics and interests that serve the purpose and object of this study, hence allowing the researchers to select efficiently online. Quotas were considered to select respondents based on particular criteria such as age and socioeconomic level to guarantee that the sample accurately reflects the range of the examined population. The online survey was designed as a five-point Likert scale to easily measure the frequency of Generation Z's online activity and social media political participation. It was divided into seven parts to maintain the questionnaire's cohesiveness and prevent survey exhaustion. These parts involved data privacy and consent, the personal information of participants, and political engagement on Facebook, Twitter, Instagram, YouTube, and TikTok, respectively.

Since social media remains a broad digital sphere, the study limited its measurement of political participation to the five platforms above. These selections were based on constituents' utilization of

such platforms and the youth's popularity levels on social media networks. Social media engagements were divided into four categories: latent engagement, follower engagement, expressive engagement, and system engagement, to identify the digital activities that encourage respondents to participate. The data was analyzed through SPSS, particularly through multiple regression, linear regression, and analysis of variance (ANOVA).

RESULTS AND DISCUSSION

Type of Social Media Usage

In total, 300 respondents agreed to participate, with 68% females (n=204) and 32% males (n=96). Most participants were 21 years old (mean=21.03, min=18, max=25, SD=1.23). Most Generation Z respondents who fall between 1993 and 2005 belong to the younger end of the spectrum. Their social media usage results indicated that respondents are only active online users on Instagram, constituting 56% (n=167) of the population. Meanwhile, most respondents display passive online usage habits on the remaining platforms, such as Facebook, Twitter, YouTube, and TikTok.

Table 3 shows the descriptive results of the profile variables in the study, where most respondents (N=300) are female (n=204, 68%), while males constituted less than half of the population (n=96, 32%).

Table 2

Demographic information of the respondents and social media platform usage (n= 300)

	n	% of sample
Gender		
Male	96	32
Female	204	68
Age		
18	9	3
19	24	8
20	46	15.33
21	130	43.33
22	61	20.33
23	23	7.67
24	5	1.67
25	2	0.67
Household Income		
Less than Php25,000	49	16.33
Php25,000 to Php50,000	50	16.67
Php50,000 to Php75,000	47	15.67
Php75,000 to Php100,000	46	15.33
above Php100,000	108	36

Joseph Ching Velasco, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan,
Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

Table 2 *(Continue)*

	n	% of sample
Type of Social Media Usage		
Facebook		
Active	127	42.33
Passive	133	57.67
Twitter		
Active	112	37.33
Passive	188	62.67
Instagram		
Active	167	55.67
Passive	133	44.33
YouTube		
Active	66	22
Passive	234	78
TikTok		
Active	97	32.33
Passive	203	67.67

Source: Authors' work

Based on social media usage, the female respondents are only active online users on Instagram but remain passive online users on Facebook, Twitter, YouTube, and TikTok. Contrary to the male respondents, most of their population are passive online users on the mentioned social media platforms. The data reported that both sexes have a low influence on college respondents' political participation in social media. For female respondents, Facebook is the closest to a moderate degree of influence (M=2.26, SD=0.04), followed by Twitter (M=2.01, SD=0.05), Instagram (M=1.82, SD=0.05), TikTok (M=1.61, SD=0.04), and YouTube (M=1.54, SD=0.04). Similarly, with male respondents, sex has a limited impact on their political participation on Facebook (M=2.06, SD=0.06), Twitter (M=1.77, SD=0.07), Instagram (M=1.65, SD=0.07), YouTube (M=1.52, SD=0.06), and TikTok (M=1.49, SD=0.04).

Likewise, in terms of monthly household income, the findings show that living under Php25,000 to Php50,000 income moderately influences respondents' political participation on Facebook. Although all the income ranges on the platform remain low, its degree levels are less than Php25,000 (M=2.17, SD=0.09), Php50,000 to Php75,000 (M=2.24, SD=0.09). Php75,000 to Php100,000 (M=2.15, SD=0.09) and above Php100,000 (M=2.14, SD=0.06) are the closest to moderate influence. Regardless, all the income categories have a low influence on respondents' online political participation for the remaining popular social media platforms.

Table 3
Profile of the respondents and their online participation

Online Participation

Variable	Facebook				Twitter				Instagram			
	Mean	SD	Social Media Use		Mean	SD	Social Media Use		Mean	SD	Social Media Use	
			Active	Passive			Active	Passive			Active	Passive
Sex												
Male	2.06	0.06	44	52	1.77	0.07	25	71	1.65	0.07	49	47
Female	2.26	0.04	83	121	2.01	0.05	87	117	1.82	0.05	118	86
Monthly Household Income												
Less than Php 25,000	2.17	0.09	-	-	1.92	0.10	-	-	1.72	0.10	-	-
Php 25,000 to Php 50,000	2.34	0.08	-	-	2.13	0.10	-	-	1.81	0.10	-	-
Php 50,000 to Php 75,000	2.24	0.09	-	-	1.94	0.11	-	-	1.77	0.10	-	-
Php 75,000 to Php 100,000	2.15	0.09	-	-	1.84	0.11	-	-	1.74	0.10	-	-
Above Php 100,000	2.14	0.06	-	-	1.88	0.07	-	-	1.78	0.07	-	-

Online Participation

Variable	YouTube				TikTok			
	Mean	SD	Social Media Use		Mean	SD	Social Media Use	
			Active	Passive			Active	Passive
Sex								
Male	1.52	0.06	33	63	1.49	0.06	26	70
Female	1.54	0.04	33	171	1.62	0.04	71	133
Monthly Household Income								
Less than Php 25,000	1.60	0.08	-	-	1.55	0.09	-	-
Php 25,000 to Php 50,000	1.58	0.08	-	-	1.68	0.09	-	-
Php 50,000 to Php 75,000	1.58	0.08	-	-	1.64	0.09	-	-
Php 75,000 to Php 100,000	1.40	0.08	-	-	1.54	0.09	-	-
Above Php 100,000	1.51	0.05	-	-	1.54	0.06	-	-

Note: n=300. Low - 1.00 - 2.33; Moderate - 2.34 - 3.66; High - 3.67 - 5.00
Source: Authors' work

Joseph Ching Velasco, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan,
Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

Table 4

Social media participation and social media usage

Platform	Online Engagement	Mean	SD	Interpretation
Facebook	Latent	3.11	0.71	Moderate
	System	1.21	0.58	Low
	Follower	2.34	0.88	Moderate
	Expressive	2.12	0.77	Low
	Overall Mean	2.19	0.78	Low
Twitter	Latent	2.76	1.11	Moderate
	System	1.14	0.55	Low
	Follower	2.07	0.97	Low
	Expressive	1.77	0.85	Low
	Overall Mean	1.93	0.67	Low
Instagram	Latent	2.34	1.07	Moderate
	System	1.12	0.49	Low
	Follower	1.83	0.87	Low
	Expressive	1.78	0.81	Low
	Overall Mean	1.77	0.50	Low
Youtube	Latent	2.25	0.99	Low
	System	1.11	0.48	Low
	Follower	1.43	0.65	Low
	Expressive	1.34	0.58	Low
	Overall Mean	1.53	0.50	Low
TikTok	Latent	2.25	1.06	Low
	System	1.12	0.53	Low
	Follower	1.53	0.75	Low
	Expressive	1.41	0.61	Low
	Overall Mean	1.58	0.48	Low

Note: n=300. Low: 1.00–2.33, Moderate: 2.34–3.66, High: 3.67–5.00.

Source: Authors' work

The findings showed respondents' low political participation on the five popular social media platforms. It can be seen in the low overall marks of Facebook (M=2.19, SD=0.78), Twitter (M=1.93, SD=0.67), Instagram (M=1.77, SD=0.50), YouTube (M=1.53, SD=0.50), and TikTok (M=1.58, SD=0.48) indicating that the political engagements of the respondents on popular social media platforms are weak and only limited to selective political actions. Although there is a predominantly low level of participation amongst the respondents, the respondents exhibit moderate political participation in Follower (M=2.34, SD=0.88) engagement on Facebook, which

shows that the respondents are not limited to passive political online actions. They participate actively through digital petitions, post-sharing, and the like. Additionally, this moderate level can also be observed in the latent engagements for Facebook (M=3.11, SD=0.71), Twitter (M=2.76, SD=1.11), and Instagram (M=2.34, SD=1.07). This data shows that respondents tend to engage more politically on these platforms passively.

Similarly, latent engagement remains the most practiced form of political participation by the respondents on YouTube and TikTok. However, the respondents exhibited only low-level scores (M=2.25, SD=0.99) and (M=2.25, 1.06) of political participation in the mentioned platforms, respectively. Hence, passive political participation is the most practiced form of online engagement across all platforms, limited to reading and watching social media content. Concurrently, system engagement is the lowest-scored form of online engagement practiced by the respondents across all five platforms. Specifically, it only exhibited a weak degree of political participation on Facebook (M=1.21, SD=0.58), Twitter (M=1.14, SD=0.55), Instagram (M=1.12, SD=0.49), YouTube (M=1.11, SD=0.48), and TikTok (M=1.12, SD=0.53) entailing that respondents have a very low tendency to participate in online political actions that involve hacking, spreading information, and the like.

The table presented above indicates the results of a regression analysis examining how respondents' online political engagement on Facebook is affected by their platform usage. The r values of 0.147–0.176 generally indicate a very weak correlation strength between the two, conveying that their use of Facebook has a limited impact on their political behavior. Therefore, the platform does not notably push the respondents to participate politically.

Based on the findings, latent engagement was significantly predicted by respondents' Facebook usage (F=3.1541, p=0.0252) and was the most influenced, with a 3.10% (r2=0.0310) variance. It indicates that they primarily interact with political content by viewing posts. Regarding follower and expressive engagement, they were also significantly predicted by time spent on Facebook (F=2.5954, p=0.0527) and (F=0.2861, p=0.0368), respectively. The said usage influenced the respondents' follower engagement by 2.56% (r2=0.0256) and 2.83% (r2=0.283) for the respondents' expressive engagement variance, illustrating that the passive viewing of posts is followed by more active forms, such as sharing posts, joining groups, and signing petitions. However, for the remaining engagement, the results indicated that respondents' Facebook usage could not predict their system engagement (F=2.1924, p=0.0890). It is the least influenced by online consumption on Facebook as it can only explain 2.17% (r2=0.0217) of its variance.

Based on the results above and the r values of 0.054-0.270, there is a very weak to weak correlation strength between Generation Z's online usage on Twitter and political participation. Among the four types of online engagements, expressive

Joseph Ching Velasco, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan,
Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

Table 5
Generation Z's political participation on Facebook

Social Media Usage	Latent Engagement $r^2 = 0.0310, F = 3.1541***$, Pr > F = 0.0252			System Engagement $r^2 = 0.0217, F = 2.1924***$, Pr > F = 0.0890			Follower Engagement $r^2 = 0.0127, F = 1.2692***$, Pr > F = 0.2851			Expressive Engagement $r^2 = 0.0067, F = 0.6705***$, Pr > F = 0.5710		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
0–3 hrs	-0.2264	0.1584	0.1540	-0.2209	0.1592	0.1662	-0.3005	0.1589	0.0595	-0.3062	0.1587	0.0545
4–6 hrs	-0.1068	0.1509	0.4797	-0.1561	0.1517	0.3042	-0.1998	0.1514	0.1879	-0.0416	0.1512	0.3499
7–9 hrs	-0.0263	0.1012	0.7949	0.0102	0.01017	0.9204	-0.0453	0.1015	0.6558	-0.1171	0.1013	0.2488
More than 10 hrs	-			-			-			-		

Note: Values are regression coefficients with standard errors in parentheses. *$p<0.05$, **$p<0.01$, ***$p<0.001$.

Source: Authors' work

Table 6
Generation Z's political participation on Twitter

Social Media Usage	Latent Engagement $r^2 = 0.0614, F = 6.4589***$, Pr > F = 0.0003			System Engagement $r^2 = 0.0029, F = 0.2881***$, Pr > F = 0.834			Follower Engagement $r^2 = 0.0498, F = 5.1695***$, Pr > F = 0.002			Expressive Engagement $r^2 = 0.0731, F = 7.7855***$, Pr > F = <0.0001		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
0–3 hrs	-0.1563	0.1693	0.3564	0.1053	0.1745	0.5466	0.0748	0.1703	0.6610	-0.0781	0.1682	0.6428
4–6 hrs	0.1001	0.1558	0.5211	0.0664	0.1606	0.6798	0.1114	0.1568	0.4780	0.1650	0.1548	0.2875
7–9 hrs	0.0097	0.0959	0.9194	0.0132	0.0988	0.8942	0.1144	0.0965	0.2365	0.1184	0.0953	0.2149
More than 10 hrs	-			-			-			-		

Note: Values are regression coefficients with standard errors in parentheses. *$p<0.05$, **$p<0.01$, ***$p<0.001$.

Source: Authors' work

engagement was predicted significantly by the generation's Twitter usage (F=7.7855, p=<.0.0001) and considered the most influenced with a 7.31% (r^2=0.0731) variance. Subsequently, follower engagement was statistically predicted by the same usage (F=,5.1695, p=<.0.002), and 4.98% of its variance is influenced, indicating a higher yet moderate association level between the two variables. However, for latent and system engagement, social media usage has a weak impact, which only explains 6.14% and 0.29% of their variance, respectively. The multiple linear regression analysis showed no statistically significant relationship between the four types of social media engagements and time spent on Twitter. Thus, Twitter usage does not notably influence the respondents to participate politically. In latent, the results showed a positive relationship wherein the engagement could increase to 0.010 per unit change in the 7–9 hours of usage. However, this relationship remains statistically insignificant as the p-value, 0.919, is greater than the set alpha, 0.05, indicating that it only occurred by chance.

Based on the data presented above, there is a very weak to weak level of correlation between Instagram usage and respondents' online political participation, as supported by the *r* values of 0.126-0.233. Of the four types, follower engagement was the most influenced by respondents' Instagram usage, with a 5.42% (r^2=0.0542) variance and significant predictability of (F=5.6490, p=<0.001). Following this, latent engagement was significantly predicted by respondents' time spent on Instagram (F=5.3346, p=<0.001) with a variance of 5.13% (r^2=0.0513). Expressive engagement was significantly predicted by the same variables (F=4.8437, p=<0.003) and displayed a variance of 4.68% (r^2=0.0468). However, the results indicated that respondents' Instagram usage could not predict system engagement (F=1.5853, p=0.193). It is even least influenced by online consumption on Instagram as it can only explain 1.58% (r^2=0.0158) of its variance.

The results of the multiple linear regression showed a statistically significant negative relationship in latent engagement within 0–3 hours, 4–6 hours, and 7–9 hours of Instagram consumption. At a 99.9% confidence level, the latent engagement of respondents decreases by -0.6600 per unit change in their Instagram exposure within the 0–3 hours range. Meanwhile, there is a decrease of -0.4982 within the 4–6 hours and -0.2484 within the 7–9 hours usage at a 99% and 95% confidence level, respectively. Hence, in this aspect, the platform's usage remarkably propels the respondents to engage in political matters.

On the contrary, there is no statistically significant relationship between respondents' follower engagement, expressive engagement, and system engagement based on the mentioned data. Although their political behavior is manifested through their frequent viewing of political posts, stories, and reels, such passive action does not necessarily translate to active engagements. The *p*-values for the

Joseph Ching Velasco, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan,
Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

Table 7
Generation Z's political participation on Instagram

Social Media Usage	Latent Engagement			System Engagement			Follower Engagement			Expressive Engagement		
	$r^2 = 0.0513$, $F = 5.3346$***, Pr > F = <0.001			$r^2 = 0.0158$, $F = 1.5853$***, Pr > F = 0.193			$r^2 = 0.0542$, $F = 5.6490$***, Pr > F = <0.001			$r^2 = 0.0468$, $F = 4.8437$***, Pr > F = <0.003		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
0–3 hrs	-0.6600	0.1955	0.0008	0.0742	0.1992	0.7096	-0.3089	0.1952	0.1147	-0.3060	0.1960	0.1195
4–6 hrs	-0.4892	0.1861	0.0090	0.1926	0.1895	0.3103	-0.0751	0.1858	0.6865	-0.1049	0.1865	0.5742
7–9 hrs	-0.2484	0.1118	0.0271	0.0630	0.1139	0.5807	-0.0448	0.1117	0.6889	-0.0218	0.1121	0.8462
More than 10 hrs	-			-			-			-		

Note: Values are regression coefficients with standard errors in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Source: Authors' work

Table 8
Generation Z's political participation on YouTube

Social Media Usage	Latent Engagement			System Engagement			Follower Engagement			Expressive Engagement		
	$r^2 = 0.0092$, $F = 0.9133$***, Pr > F = 0.4338			$r^2 = 0.0073$, $F = 0.7304$***, Pr > F = 0.5346			$r^2 = 0.0127$, $F = 1.2692$***, Pr > F = 0.2851			$r^2 = 0.0067$, $F = 0.6705$***, Pr > F = 0.5710		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
0–3 hrs	0.0206	0.1633	0.8998	0.1168	0.1635	0.4755	-0.1093	0.1630	0.5031	-0.0343	0.1635	0.8341
4–6 hrs	0.1091	0.1563	0.4857	0.0298	0.1565	0.8492	-0.1291	0.1560	0.4086	0.0057	0.1565	0.9712
7–9 hrs	-0.0098	0.0938	0.9167	0.0248	0.0939	0.7921	-0.1547	0.0937	0.0996	-0.0832	0.0939	0.3766
More than 10 hrs	-			-			-			-		

Note: Values are regression coefficients with standard errors in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Source: Authors' work

remaining three engagements are below the significance level of 0.05, and no domain in Instagram usage significantly predicts respondents' political activity in terms of the engagements.

Based on the results and as indicated by the *r* values of 0.082–0.113, there is a very weak association between Generation Z's online political participation and YouTube usage. With 1.27% (r^2=0.0127) of its variance, follower engagement (F=0.0127, *p*=0.2851) was the most influenced among the four types of online engagement. It is followed by the data emanating from latent engagement (F=0.9133, *p*=0.4338) and system engagement (F=0.0073, *p*=0.5346), with a variance of 0.92% (r^2=0.0092) and 0.73% (r^2=0.0073) that did not fully produce a statistically significant association with usage. Finally, expressive engagement (F=0.0067, *p*=0.5710) holds the least influenced value attributed to its variance of 0.67% (r^2=0.0067). Meanwhile, the multiple regression analysis indicates no statistically significant relationship between the four types of online engagement and time spent on YouTube. That being the case, time spent on YouTube does not propel political participation among the respondents. Although latent engagement could decrease to -0.098 in 7–9 hours of YouTube usage, this relationship does not equate to a statistically significant figure as the *p*-value 0.917 is greater than the set alpha, 0.05, indicating that it may have occurred coincidentally.

For TikTok, the *r* values of 0.146–0.372 illustrate a very weak to weak link

Table 9

Generation Z's political participation in TikTok

Social Media Usage	Latent Engagement $r^2 = 0.1384, F= 15.8529^{***}$, Pr > F = <0.0001			System Engagement $r^2 = 0.0212, F= 2.1457^{***}$, Pr > F = 0.095			Follower Engagement $r^2 = 0.1120, F= 12.4459^{***}$, Pr > F = <0.0001			Expressive Engagement $r^2 = 0.1161, F= 12.9579^{***}$, Pr > F = <0.0001		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
0–3 hrs	-0.5198	0.1265	<0.0001	-0.1898	0.1348	0.1603	-0.4666	0.1284	0.0003	-0.4992	0.1281	0.0001
4–6 hrs	-0.1527	0.1236	0.2177	-0.0610	0.1317	0.6437	-0.1380	0.1255	0.2722	-0.1640	0.1252	0.1911
7–9 hrs	-0.1382	0.0852	0.1060	-0.1389	0.0909	0.1274	-0.1159	0.0865	0.1817	-0.1651	0.0863	0.0568
More than 10 hrs	-			-			-			-		

Note: Values are regression coefficients with standard errors in parentheses. *$p<0.05$, **$p<0.01$, ***$p<0.001$.

Source: Authors' work

between Generation Z's usage and political participation. The latent engagement was found to be significantly predicted (F=15.8529, p=<0.0001) and the most influenced by Generation Z's time spent on TikTok with a 13.84% (r^2=0.1384) variance. It is followed by expressive engagement (F=12.9579, p=<0.0001) and follower engagement (F=12.4459, p=<0.0001) with a variance of 11.61% (r^2=0.1161) and 11.20% (r^2=0.1120) correspondingly. Similar to latent engagement, expressive and follower engagement were significantly predicted by the generation's TikTok usage. The figures are observably much lower for System engagement, given that only 2.13% (r^2=0.0212) of its variance is explained by Generation Z's TikTok usage (F=2.1457, p=0.095). It indicates that system engagement is the least influenced and is not significantly predicted using the given platform.

In the multiple regression analysis, three of the four forms of engagement all have a statistically significant relationship when Generation Z's TikTok usage is within the 0–3-hour range. At a 99.9% confidence level, latent engagement decreases by -0.5198 per unit change in online exposure on TikTok within the range above. At the same confidence level and hours spent on TikTok, follower and expressive engagement decreased by -0.4666 and -0.4992 per unit change, respectively. With this, the respondents' usage of the given platform notably influences their participation.

In contrast, the 0.095 p-value in system engagement is greater than the set alpha.

Thus, there is no significant relationship between system engagement and TikTok usage. Although a -0.1389 per unit change is present within the 7–9-hour range, it is still not statistically significant as its p-value of 0.1274 is greater than the set significance level.

DISCUSSION

The study incorporated Waeterloos et al. (2021) four categories of online engagement (i.e., latent, follower, expressive, and system) and their classifications as either active or passive to examine the dynamics of participation in the social media realm. Ultimately, the findings reject the argument that usage of all five social media platforms influences participation amongst Generation Z. Findings from the multiple linear regression analysis highlight that exposure only to TikTok and Instagram holds the potential to influence the respondents' online political engagement.

The results revealed that political content could spur participation in three distinct manners: passive latent political engagement, more active modes of follower engagement, and expressive engagement. Such findings are validated by Waeterloos et al. (2021) assertion that online political participation is determined to be manifested through the said forms of engagement. As discussed by Mendenhall and Sodani (2021), TikTok emerges as a catalyst for diverse forms of political involvement due to its attributes that enable content creators to present comprehensible political information tailored to a younger

audience. It was particularly evident in the controversial 2022 Philippine presidential and vice-presidential elections, where content about President Bongbong Marcos Jr. and his family was widely circulated on the platform. Mendoza (2022) explained that aside from disinformation machinery and campaign teams, the popular videos were also directly sourced from ordinary social media users. It substantiates the theory that latent, follower, and expressive engagement are utilized to participate in political affairs. However, the findings also indicate that prolonged exposure to TikTok correlates with diminished political participation. With its predominant content focus on humor and pop culture-related material (Vijay & Gekker, 2021; Mendoza, 2022), individuals inclined toward entertainment may encounter a scarcity of political content on their personalized homepages. It implies the limits of the relationship dynamics between social media usage and political participation.

Instagram is also extensively used by political media outlets, institutions, parties, and officials to connect more intimately with constituents, which Generation Z predominantly employs for political education and understanding (Beriansyah & Qibtiyah, 2023). Despite the observable consumption of political information, the study's findings indicate that the youth's inclination to learn about politics weakens as they continue to access the platform. Compared to TikTok, participants' political engagement on Instagram falls within the less active spectrum, a paradox when contrasted with their 'active' response concerning online utilization on the platform. Instead, their activities revolve around cognitive engagement, encompassing activities like reading posts, visiting profiles of political figures, and consuming reels and stories. Such findings align with the study's results, highlighting how individuals manifest their involvement in socio-political matters, specifically through latent engagement. Additionally, the results indicate a heightened inclination among participants to seek information on Instagram compared to other social media platforms. It contrasts with Mariano et al.'s (2021) study, which determined that Facebook was the most used platform to retrieve political content during the 2022 elections.

On the other hand, a very weak correlation exists between respondents' engagement on Facebook, Twitter, and YouTube. Despite the Philippines being widely recognized as the 'social media capital' of the world due to its extensive Facebook usage, the findings surprisingly indicate a moderate level of participation, with latent engagement emerging as the dominant mode. Generation Z often concentrates on information-seeking activities, prioritizing self-education about pertinent matters over concrete actions on specific issues. The findings convey that Generation Z primarily engages on Twitter through expressive participation. The influential nature of Twitter, which contributes to creating a 'temporal unfolding' of incidents, enables individuals to construct larger narratives, often shaped by the different narratives aggregated by hashtags

(Yang, 2016). Meanwhile, with YouTube's built-in recommendation system as a pivotal indicator of maximal content consumption, Hosseinmardi et al. (2021) highlight how political engagement extends beyond merely seeking further insight into political matters. The findings validate this notion, indicating that Generation Z primarily expresses political participation on YouTube through follower engagement. However, as indicated by the findings, such platform usage does not significantly affect online participation, thus leading to the rejection of the assertion.

Apart from the vast array of features possessed by each platform, household income and sex were also delved into as potential factors influencing online political participation. As revealed by the findings, one's monthly household income holds a moderate degree of influence, underscoring that socioeconomic status does not dictate the extent of online political engagement among the respondents. It puts a different perspective from Western democracies, where socioeconomic status is often used to explain political participation, positing that individuals of higher status are more inclined to engage in political activities (Verba & Nie, 1974; Hite, 2010). Accounting for the influence of clientelism in the Philippines, this model does not necessarily align well with the political psychology in the country. This finding supports Hite's (2010) proposition that individuals with greater financial stability tend to exhibit lower political efficacy. Subsequently, this leads to reduced engagement in political actions and behaviors. Likewise, the findings suggest

that gender divisions and societal norms only partially influence engagement. It further supports the conclusion of Bimber et al. (2021) that both sexes express themselves politically through sharing and commenting on social media platforms. Nevertheless, these scholars argue that gender-specific political voices exist, with women often encountering reduced political expression in open public discourse spaces.

CONCLUSION

The study examined four types of political engagement, i.e., latent, system, follower, and expressive online political engagement. These were examined across five prominent social media platforms: Facebook, Twitter, Instagram, YouTube, and TikTok. Among these platforms, the study's findings suggest that only Instagram and TikTok have effectively advanced political participation among the respondents. As noted in this study, Instagram seems to be the principal source of political information for the younger generation, which predominantly elicits passive participation. On the other hand, TikTok possibly facilitated further political involvement through its algorithm, enabling users to access political themes and issues readily.

TikTok's tools allowed creators to create appealing video content, subsequently creating user engagement. Respondents employed TikTok to express political viewpoints by uploading content, engaging in peer messaging, and sharing politically oriented information. Nonetheless, given TikTok's primary entertainment

and socializing orientation, the impact of political engagement waned as users spent more time on the platform. Individuals from Generation Z are extensively immersed in various online social media platforms. Consequently, emerging new and evolving networks like Threads and Snapchat may deepen the exploration of how Generation Z leverages social media to navigate socio-political matters.

Implication to Theory and Practice

Political participation is gradually being redefined with the introduction of emerging technologies and platforms. Traditional understanding of political participation includes voting, campaigning, protesting, and engaging in civic education. However, the changing communication landscape calls for a broader consideration of the practice of political participation. Specifically, social media has become an inherent part of society, changing how individuals communicate, consume information, and create communities. It subsequently opened new opportunities for mobilization and even activism in the digital realm. Since the youth are primarily adept at taking advantage of these new communication platforms, studies must explore how new and emerging technologies might hinder or enable political participation. The idea of amplification is linked with political participation in social media. Traditionally marginalized voices, which might have gone unnoticed previously, are now amplified.

Aside from amplifying voices from marginalized communities, this also points toward another outcome: democratizing political discourse. New social media platforms have created new opportunities for political participation beyond traditional boundaries. It allows for creating communities of individuals with shared ideologies and interests. While the potential for political participation on social media is immense, such comes with pitfalls. Issues such as disinformation and the creation of echo chambers are counterproductive outcomes against informed political participation. Likewise, the opaque algorithms of platforms influence what users see on their feeds, limiting an individual's exposure to diverse political perspectives.

Limitations and Recommendations for Future Research

This study seeks to expound on the political participation of individuals from Generation Z in the Philippines. There are some limitations to consider in terms of the possible temporal validity. Platforms are continuously evolving, requiring continuous examination of the phenomenon. Future studies could delve into Generation Z's information-seeking behavior and background research, particularly considering their inclination toward passive engagement as the predominant form of participation. This avenue of investigation might further elucidate how online interactions translate into active involvement in social justice rallies and protests. It leads to implications both in theory and practice. Specifically, the extensive use of these social media platforms enables

Joseph Ching Velasco, Jhazmin Joi C. Manguera, Eunice Julia V. Navalan,
Julianna Nicole C. Limchiko, Trixie Anne Marie D. Bayot and Marjorie Grace V. Mercado

new forms of socio-political engagements. Considering the respondent's predilection for passive engagement on social media, there is a need to investigate how this shift in consuming information influences their ability to discern and assess credible sources of information.

Geographic coverage could also be broadened to encompass regions beyond Metro Manila. It enriches the dataset for a comprehensive understanding of political engagement. Moreover, statistical analyses beyond ANOVA and multiple linear regression could provide deeper insights into the relationship between social media usage and user perspectives. Future research might pivot to qualitative methods for a more holistic grasp of social reality and user behaviors. Alternatively, drawing on studies like Waeterloos et al. (2021) and other scholarly articles could serve as a foundation or framework for refining the Social Media Participation Scale, tailoring it to reflect Filipino behaviors and integrating the cultural and societal context of the country more accurately.

ACKNOWLEDGEMENT

REFERENCES

Adedokun, T. A. (2022). Social media as a strategy for protest movements: A study of #EndSARS in Nigeria. *International Journal of Research in Business and Social Science*, *11*(6), 438–450. https://doi.org/10.20525/ijrbs.v11i6.1927

Alawi, A., Wei, Y., & Suh, J. (2020). *ECAASU condemns Philippines anti-terror law. East Coast Asian American student union.* https://www.ecaasu.org/ecaasu-condemns-philippines-anti-terror-law/

Ardèvol-Abreu, A., Delponti, P., & Rodríguez-Wangüemert, C. (2020). Intentional or inadvertent fake news sharing? Fact-checking warnings and users' interaction with social media content. *Profesional de la Información, 29*(5), 1-13. https://doi.org/10.3145/epi.2020.sep.07

Arugay, A. (2022). *Stronger social media influence in the 2022 Philippine elections.* ISEAS Yusof Ishak Institute. https://www.iseas.edu.sg/articles-commentaries/iseas-perspective/2022-33-stronger-social-media-influence-in-the-2022-philippine-elections-by-aries-a-arugay/

Baym, N. K., & Evans, J. M. (2022). The audacity of clout (chasing): Digital strategies of black youth in Chicago DIY Hip-Hop. *International Journal of Communication*, *16*, 2669-2687.

Beriansyah, A., & Qibtiyah, M. (2023). Instagram and political literacy generation Z. *Jurnal Kajian Tunggal Ika: Kajian Teori Dan Praktik Pendidikan PKN, 10*(1), 134–149. https://doi.org/10.36706/jbti.v10i1.20463

Bimber, B., Koc-Michalska, K., & Lilleker, D. G. (2021). Women learn while men talk?: Revisiting gender differences in political engagement in online environments. *Information, Communication & Society, 24*(14), 2037–2053. https://doi.org/10.1080/1369118x.2021.1961005

Hite, N. (2010). *Political engagement, social change and the political economy of financial modernization: Experimental evidence from the Philippines.* Social Science Research Network. https://doi.org/10.2139/ssrn.1674102

Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D. M., & Watts, D. J. (2021). Examining the consumption of radical

content on YouTube. *Proceedings of the National Academy of Sciences*, *118*(32). https://doi.org/10.1073/pnas.2101967118

Kemp, S. (2023, February 9). Digital 2023: The Philippines. *Data Reportal*. https://datareportal.com/reports/digital-2023-philippines

Kim, J., Lee, C., & Elias, T. (2015). Factors affecting information sharing in social networking sites amongst university students. *Online Information Review, 39*(3), 290–309. https://doi.org/10.1108/OIR-01-2015-0022

Lim, N. N. (2009). Novel or novice: Exploring the contextual realities of youth political participation in the age of social media. *Philippine Sociological Review*, *57*, 61–78.

Macaraig, A., & Hameleers, M. (2022). #DefendPressFreedom: Paradigm repair, role perceptions and Filipino journalists' counterstrategies to anti-media populism and delegitimizing threats. https://doi.org/10.1080/1461670X.2022.2138949

Madarang, C. R. S. (2020, March 12). Metro Manila residents are panic buying amid COVID-19. Why they do it and why they shouldn't. *Interaksyon.* https://interaksyon.philstar.com/rumor-cop/2020/03/12/164036/panic-buying-metro-manila-covid-19/

Manavis, S. (2020, March 10). On TikTok, coronavirus is just another way to gain clout. *The New Statesman.* https://www.newstatesman.com/science-tech/2020/03/tiktok-coronavirus-just-another-way-gain-clout

Marcaida, M. Y. (2020). Student activism offline and online: A mixed-methods study on college students' protest participation in the Philippines. *Philippine Social Sciences Review, 72*(1), 1-22.

Mariano, J. B., De Torres, M., & Vargas, D. (2021). *Social media and political decision making.* Social Science Research Network. https://doi.org/10.2139/ssrn.3791368

Mendenhall, S., & Sodani, T. (2021). Binge-swiping through politics: TikTok's emerging role in American government. *Journal of Student Research, 10*(2). https://doi.org/10.47611/jsrhs.v10i2.1777

Mendoza, M. E. H. (2022). Philippine elections 2022: TikTok in Bongbong Marcos' presidential campaign. Contemporary Southeast Asia: *A Journal of International and Strategic Affairs 44*(3), 389-395.

Radut, F. (2021). *Generation "Z" and social networks*. Social Science Research Network. https://ssrn.com/abstract=3918595

Santos, A. P. (2022, August 22). A masterclass in propaganda - political vloggers in the Philippines. *Al Jazeera Media Institute.* https://institute.aljazeera.net/en/ajr/article/1966?fbclid=IwAR0izCZpPDleg2G-G27iP-x-rswL7AP9mydsJOnjkGWgzvIKdfvdio8HOkY

Verba, S., & Nie, N. H. (1974). Participation in America: Political democracy and social equality. *Social Forces*, *53*(1), 133. https://doi.org/10.2307/2576853

Vijay, D., & Gekker, A. (2021). Playing politics: How Sabarimala played out on TikTok. *American Behavioral Scientist, 65*(5), 712–734. https://doi.org/10.1177/0002764221989769

Waeterloos, C., Walrave, M., & Ponnet, K. (2021). Designing and validating the social media political participation scale: An instrument to measure political participation on social media. *Technology in Society, 64,* Article 101493. https://doi.org/10.1016/j.techsoc.2020.101493

Yang, G. (2016). Narrative agency in hashtag activism: The case of #BlackLivesMatter. *Media and Communication, 4*(4), 13-17. https://doi.org/10.17645/mac.v4i4.692

Assessing Lecturers' Belief on Assessment in Nigeria Using Brown's Factor Model of Assessment Belief: A Study of Content-based Validity

Yaro Umar and Lim Hooi Lian*

School of Educational Studies, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia

ABSTRACT

The objectives of this study are threefold: First, to highlight the importance of assessment beliefs and their contribution to the practice of assessment. Second, the current literature on assessment beliefs in different contexts should be assessed, including the relationship between lecturers' beliefs on assessment and their evaluations of student performance. Third, the content-based validity of an adapted instrument (Brown conception of Assessment III Abridged Survey) will be measured using the Lawshe content validation ratio for use among lecturers within the context of Nigeria. This psychometric study utilized a 27-item survey from Brown (2006), covering four dimensions: improvement of teaching and learning, student accountability, institutional accountability, and assessment as irrelevant. Seven experts were randomly selected to judge the relevancy of each item to the domain construct based on the defined four-point scale. The Content Validity Ratio (CVR), based on the Lawshe Validity Model (1975), is used to analyze the data collected. Results indicated that items under the "assessment as irrelevant" were deemed non-essential. As a major contribution to this study, the existing literature on content-based validity studies is expanded by applying the Lawshe content validation ratio. Based on the study result, a three-dimensional conceptual framework for analyzing lecturers' beliefs on assessment is also suggested. It calls for additional psychometric properties, particularly construct validity testing, to ensure the convergent and discriminate validity of the instrument within the context of TVET tertiary institutions in Nigeria.

Keywords: Assessment beliefs, Brown four-factor model of assessment, content-based validity, lecturers

INTRODUCTION

Assessment in any instructional context, including higher education, is to provide evidence of student learning, student progress, teaching quality, and institutional accountability (Fletcher et al., 2012; Paris et al., 2012; Suleman et al., 2020). Thus, given the importance of assessment, issues related to assessment beliefs deserve investigation, particularly those connected to lecturers (Elshawa et al., 2017). Beliefs in assessment represent a critical issue in the field of assessment research. Even though this research area affects policy and practice in tertiary institutions, little is known about lecturers' beliefs about assessment (Opre, 2015). There may be numerous factors that influence the intended assessment practice. For instance, individual assessment beliefs and their varied purposes of assessment affect their judgment on what assessment methods to adopt (Osman et al., 2021).

Additionally, evidence from empirical research indicates that beliefs regarding assessment significantly influence the choice of assessment methods (Thomas, 2012; Vandeyar et al., 2007). For example, researchers have illustrated that different assessment beliefs lead to different assessment practices (Thomas, 2012; Vandeyar et al., 2007). As a result, lecturers who believe in assessing student learning use traditional assessment practices. In contrast, lecturers who believe in assessment for student learning will be motivated to use alternative assessment approaches. Thus, assessment practices employed in assessing students' learning differ depending on the

beliefs of assessment, teaching, and learning theories (Moiinvaziri, 2015). Given this, Brown et al. (2011) suggested that it is essential to consider educators' assessment beliefs to appreciate them and, if necessary, find ways to improve their assessment practices. Therefore, understanding the assessment beliefs of a lecturer, especially at the tertiary level, is essential since it can help improve the quality of assessment practices for the effective implementation of any assessment policy reform (Elshawa et al., 2017; Mohd et al., 2013).

Therefore, the connection between assessment beliefs and practice is not quite direct and simple. It is influenced by numerous distinct factors, particularly context, where the teaching activity plays a significant role. For example, in Malaysia, English language instructors believe that the assessment aims to improve teaching and learning, which involves providing information about students' progress and giving feedback to students (Elshawa et al., 2017). However, in Colombia, a study of 62 lecturers found a contradiction between their reported beliefs and practices, suggesting that lecturers need opportunities for reflection, self-assessment, and more guidance on formative assessment practices (Muñoz et al., 2012). Moreover, in higher education in New Zealand, faculty and undergraduate students across four tertiary institutions (two universities, one indigenous tertiary institution, and a polytechnic) held differing views. The results indicate that faculty viewed assessment as a trustworthy process aiding teaching and learning. In contrast,

students viewed assessment as focusing primarily on accountability and perceived assessment as irrelevant or ignored in the teaching and learning process (Fetcher et al., 2012). Hence, the outcomes of these studies brought forward three perspectives on how belief-practice relationships function. Beliefs and practices influence each other, resulting in a contradiction between what is believed and what is practiced and a different opinion regarding the purpose of assessment. It clearly shows that the relationship between belief and practice is complex and varies across contexts and individuals. Thus, assessment of lecturers' beliefs have great consequences as they influence their assessment practices (Harris & Brown, 2009).

Therefore, in technical and vocational education and training (TVET), assessment is essential for determining students' learning outcomes. It provides accurate feedback on whether students have successfully achieved the intended learning outcomes (Yusop et al., 2022). Assessing students' competence in TVET, whether through written or practical methods, is crucial to ensuring they have demonstrated mastery of practical skills and essential abilities required to carry out tasks based on specific curriculum standards (Mazin et al., 2020). However, research suggests that assessment practices in TVET can sometimes be inconsistent with their intended purposes (Yusop et al., 2023). More importantly, assessment serves multiple purposes, such as providing information about student learning and progress, teaching quality, and

program and institutional accountability (Opre, 2015). In his various studies, Brown (2004, 2008) identified four main purposes teachers' beliefs about assessment can serve. These beliefs on assessment vary widely. Some educators see it as a powerful tool to enhance teaching and learning, emphasizing helpful feedback (Brown et al., 2019). Others view assessment as holding students accountable through scores, grades, and certifications (Brown et al., 2019). Another perspective sees assessment as a way to gauge the effectiveness of schools and teachers, providing valuable data for improvement (Brown et al., 2019). Finally, some educators believe the assessment is irrelevant or harmful, questioning its accuracy and preferring to focus solely on teaching (Brown et al., 2019). This diverse range of beliefs highlights the complex role assessment plays in education.

Consequently, the role of assessment in TVET is to ensure that students' knowledge, skills, and attitudes are assessed (Yusop et al., 2023). These assessments manifest in two crucial phases: formative with a future goal, which is to gain feedback, and summative, which assesses concrete achievement and acquires evidence (Black et al., 2018). Thus, constant improvements from these assessments are required to ensure the development of knowledgeable and skilled students (Shepard et al., 2018). Hence, lecturers' understanding of the purpose and function of assessment is closely related to how they implement it in their classroom practice. While using assessment to improve teaching and learning

may be an essential factor in being a lecturer in TVET, their beliefs depend on the sociocultural context and policy framework within which they operate.

Therefore, the conception of the Assessment III Abridged Survey from Brown (2006) as a measurement tool was developed to measure teachers' beliefs about the four major purposes of assessment. The term "conception" of assessment denotes educators' belief systems about the nature and purpose of assessment and encompasses their cognitive and affective responses (Xu & Brown, 2016). Literature confirmed that researchers of assessment beliefs used diverse subsuming terms such as conception and values to define variables of interest. For example, scholars such as Remesal (2011) distinguish between the term's conceptions and beliefs, while others such as Calveric (2010), Vardar (2010), and Yidana et al. (2018) decided to use the two terms interchangeably. The concept of conception incorporates knowledge and beliefs into a singular construct to provide a framework for describing this study's overall perception and awareness of assessment.

Hence, the application of this instrument requires that validity be reported in content and construct. For example, Brown conducted studies in the context of basic education in New Zealand, focusing only on construct validity. Such studies include Teachers' Conceptions of Assessment: Implications for Policy and Professional Development (Brown, 2004) and Teachers' Conceptions of Assessment: Validation of an Abridged Version (Brown, 2006). Other

studies that applied and adapted the same instrument in the context of higher education show no evidence of content or construct validity. In Nigeria, for example, Raji et al. (2020), in their study titled "Teachers' Conceptions and Choices of Assessment Tasks in a Nigerian Postgraduate Teacher Training," adopted the same instrument but reported no content or construct validity evidence. Hence, using the instrument in the TVET context in Nigeria requires the establishment of psychometric properties in terms of content and construct validity. This evidence suggests that while the Conceptions of Assessment Abridged survey might be suitable for the New Zealand context, its validity needs to be established through content and construct validation before application in other contexts, especially the TVET context or those significantly different, like high education or different countries. Thus, this study intends to address this critical research gap identified in the TVET context: the absence of established content validity for the "Conceptions of Assessment Abridged" survey.

Accordingly, it is observed from the literature that the evaluation of content validity has been much neglected in the literature on assessment beliefs. While Brown (2019) lamented that context, culture, and local factors shape teachers' conceptions of assessment, a literature review creates a gap that the present study intends to fill. Nevertheless, a lack of precise specifications and standards challenges understanding assessment beliefs and their underlying principles in the Nigerian

higher education context. Subsequently, wide variations in its conceptualization and the resulting practical implications are expected. As a result, more data is needed on the Nigerian TVET higher education context to contextualize assessment beliefs, experiences, practices, and professional development needs, particularly during this era of accountability. Therefore, research on the validation process and the construction and adaptation of measurement instruments are still insufficient in Nigeria's education field. The current study is a step forward in this direction by using the Content Validity Ratio (CVR) based on the Lawshe Validity Model (1975) to determine the content validity of an adapted Brown Conception of Assessment Abridge Version for use among lecturers in the context of Nigerian TVET tertiary institutions.

LITERATURE REVIEW

Content-based Validity

Content validation is one of the numerous steps in the instrument development process that requires careful attention. Content validity refers to the extent to which a measurement instrument accurately represents the content being measured, thereby establishing its validity (Akmal et al., 2022). It is the first and inevitable step in assessing other sources of validity. Yet, it is ranked among instrument development's less frequently reported psychometric properties (Zapata et al., 2022). As such, it is regarded as a challenge that needs to be overcome by acknowledging its worth. Therefore, since content validity is essential to ensuring

overall validity, it is important to conduct content validation systematically using evidence and established best practices.

Moreover, researchers in psychometrics, namely Galicia et al. (2017) and Taherdoos (2016), identified three validation methods: content, construct, and criterion, with construct validity being the most widely used. Contrary to the literature reports on the commonly employed construct validity in this study, content-based validation criteria are considered. Therefore, the usual procedure for assessing the content and validity of an instrument is to consult experts. Experts are selected to evaluate and criticize an assessment method and are usually chosen based on their knowledge of the subject matter under consideration (Yusoff, 2019). The way these experts are chosen is, thus, fundamental since they must be experts in the field, either because of expertise or work experience. Once selected, the experts assess the different dimensions of the instrument using a numerical scale as a part of the procedure (Galacia et al., 2017). The process that the experts undertake becomes a real effort to eliminate unimportant aspects, modify aspects that need modification, and incorporate relevant aspects (Rubio et al., 2003). Thus, in this study, the researcher considers this process thoroughly and decides what to modify, enhance, or eliminate from the Conception of Assessment III Abridged Survey adapted for use among lecturers in TVET tertiary institutions using content validation procedures.

Assessment Beliefs

Various studies in assessment literature have investigated assessment beliefs and how they relate to assessment practice (Ha et al., 2021; Ha & Murray, 2021). Studies investigating beliefs in relation to assessment used terms like conceptions and values to describe variables of interest (Opre, 2015; Osman et al., 2021). Therefore, beliefs about assessment denote the convictions of lecturers held about the nature and purpose of assessment that incorporate their cognitive and affective responses (Xu & Brown, 2016). Empirical research on beliefs about assessment is directed at studying beliefs in various contexts and discovering the primary factors that constrain or facilitate these beliefs being translated into assessment practice (Elshawa et al., 2017). As such, cross-cultural research recommends that beliefs on assessment differ across contexts, reflecting the internalization of society's cultural priorities and practices (Barnes et al., 2015; Brown & Harris, 2009). For example, a study in the Malaysian context revealed that English language instructors believed that the purpose of assessment was to improve teaching and learning (Elshawa et al., 2017). Equally concerning the assessment beliefs linked to the assessment purposes, data analyses revealed that the items that received the highest agreement were identifying student strengths and weaknesses, giving information about students' progress, and giving feedback as they learn.

Additionally, a study by Narathakoon et al. (2020) investigated teachers' beliefs about assessment and how they are congruent with their actual assessment practices. The research demonstrated that the teachers' approaches to assessment changed from employing a range of assessment techniques to employing a more restricted set of methods because of the O-NET tutoring policy. Incongruence between teachers' beliefs and practice could be a result of contextual factors such as educational policy and a lack of assessment knowledge.

Moreover, the literature also confirmed that studies about teachers' beliefs on assessment were conducted in New Zealand, Australia, Spain, China, Ghana, Nigeria, and Iran using Browns' (2008) COA-III (full and abridged version). The results from administering this instrument in various countries indicate different factor structures and variations in the pattern and strength of agreement on each factor (Barnes et al., 2015). For example, a study by Osman et al. (2021) on Basic School Teachers' Conceptions of Assessment in the Sissala East Municipality revealed that the participants demonstrated positive conceptions of assessment as a means for ensuring student and school accountability as well as improving teaching and learning, with assessment for student accountability yielding the highest mean value.

Similarly, a global phenomenon or a global localism, research originating in New Zealand with the Teacher Conceptions of Assessment self-report inventory has been replicated in multiple localities and languages (Brown et al., 2019), for example, (English in New Zealand, Hong Kong,

India, and Queensland; Greek in Cyprus; Arabic in Egypt; and Spanish in Spain and Ecuador) and at different ranks of instructional contexts (primary, secondary, senior secondary, and teacher education). Findings indicate that while the inventory can be applied cross-culturally after localized adaptations, no single model is universally applicable—culture, context, and the nature of teacher beliefs about assessment. Hence, judging from the literature, understanding assessment beliefs in the context of Nigerian tertiary institutions would provide further insights into cross-cultural differences in the assessment beliefs reported in the literature.

Nevertheless, empirical evidence about lecturers' views concerning classroom assessment in Nigerian TVET tertiary institutions remains scarce. A recent literature search identified a few studies on lecturers' views on classroom assessment (Raji et al., 2020). Therefore, it is crucial to understand lecturers' beliefs concerning assessment to foster sustainable collaboration among stakeholders within Nigerian TVET tertiary institutions. It is because studies concerning lecturers' and other stakeholders' perspectives have started to note the influence of lecturers' beliefs on assessment (Elshawa et al., 2017; Raji et al., 2020). For example, Raji et al. (2020) examined teachers' conceptions and choices of assessment tasks in Nigerian postgraduate teacher training. Results show that teachers practice assessments of learning and assessment for learning tasks but with preferences for assessment of learning tasks. A substantial difference

in teachers' conception of assessment was recorded.

Therefore, key findings from the reviewed studies brought forward four perspectives: First, assessment beliefs and purposes, as highlighted by Elshawa et al. (2017) and Narathakoon et al. (2020), indicate that the primary purpose of assessment is to improve teaching and learning. Second, the influence of context shows that assessment practice can be influenced by contextual factors such as educational policies and a lack of assessment knowledge. Third is cross-cultural variation, which highlights that while there are commonalities in assessment beliefs, there are variations in the factor structures and patterns of agreement among different cultural contexts. Fourth, the impacts of beliefs on assessment practice indicate that assessment beliefs can influence the choice of assessment tasks, as Raji et al. (2020) researched.

However, some gaps and limitations have been observed in the reviewed studies. For example, there is a limited focus on contextual factors, which suggests a need for in-depth exploration of these factors, including policy changes and professional development opportunities. There is a lack of comprehensive studies in specific regions, pointing out that contexts like Nigeria have been relatively understudied. As a result, more research focusing on this context would provide insights into Nigeria's unique challenges and opportunities for assessment beliefs. Equally, all the reviewed studies also show no evidence of content

validation for the adapted instruments. Consequently, these gaps and limitations suggest more comprehensive, cross-cultural, and context-specific research to advance our understanding of assessment beliefs and their implications in Nigeria. Therefore, this study intends to address this gap by providing evidence of the content-based validity of the adapted instrument, upon which a proposed framework can be suggested for use among lecturers in Nigeria. As such, the conception of the Assessment III Abridged Survey from Brown (2006) is adapted to reports on content-based validity within the context of TVET tertiary institutions in Nigeria.

Four-factor Model of Assessment Conception

Brown (2003) acknowledged four factors that explain how teachers conceptualize assessment: (1) improvement of teaching and learning, (2) accountability of teachers and schools, (3) accountability of students, and (4) irrelevance conceptions. The first dimension of how assessment improves teaching and learning is optimizing students' learning processes by giving them constructive feedback, which encourages

their commitment to assessment through self-assessment or peer assessment. Simultaneously, the assessment provides lecturers with the necessary information to enhance their teaching activities. The second assessment dimension considers that students are accountable for their learning and must acquire the essential qualifications to access different educational levels. The third assumption concerning assessment beliefs addresses assessment from the viewpoint of its suitability or utility in measuring lecturers' work and educational institutions in contrast to predetermined standards (Brown, 2003). It further shows that any deficiency in the level of student performance is assigned to those two actors: lecturers and institutions. Finally, the fourth type of belief is those conceptions that describe the assessment as irrelevant. Specifically, this perspective upholds that the usefulness of the assessment process is rejected in education. That assessment is considered to have negative consequences for the educational process, students, and lecturers (Opre, 2015). The four-factor model is presented in Table 1.

Table 1

Brown's four-factor model of assessment (Conception)

Instrument	Dimensions
COA III Abridged version	Institutional Accountability
	Students Accountability
	Improvement of teaching and learning
	Irrelevance

Source: Authors' work

METHODOLOGY

This methodological study forms an integral component of a comprehensive research endeavor employing a quantitative approach to assess the content validity of an adapted instrument, the Conceptions of Assessment of the III Abridged Survey, utilizing the content validation ratio (CVR). A panel of seven experts, each representing a distinct field of expertise, meticulously evaluates the relevance of each item within the adapted instrument to the underlying domain construct. The calculated CVR values are subsequently compared against established acceptance criteria to determine the fate of each item: inclusion in the final instrument, acceptance with modifications, or exclusion. This crucial phase of the study enhances the content validity of the measurement tool, ensuring its suitability for the specific context of Nigerian TVET tertiary institutions.

Sample (Expert Validation)

The validation panel consists of experts who were chosen to judge whether the item is necessary for operating a construct in a set of items or not for inclusion or exclusion in the instrument. Therefore, expert panel comprises individuals with knowledge and proficiency in a specified field. Hence, the participants of this study were selected using random sampling from different institutions and consisted of seven experts from different fields of knowledge with practiced experience (in research and conducting research) ranging from 15 to 30 years of experience (Table 2).

Table 2
Judges' areas of knowledge and years of working experience

Judges	Area of expertise/Academic training	Years of work experience
1	Technical Education	29
2	Technical Education	30
3	Agricultural Education	19
4	Business Education	20
5	Office Technology and Management	26
6	Education (Test & Measurement)	20
7	Education (Test & Measurement)	26

Source: Authors' work

Instrument

The content validation process using the Lawshe procedure is applied to the 27-item adapted conceptions of the Assessment III Abridged Survey from Brown (2006). The questionnaire consists of four dimensions: (1) assessment improves teaching and learning (12 items); (2) assessment is about certification of student learning (three items); (3) assessment demonstrates the quality of school (three items); and (4)

assessment is irrelevant to the work of lecturers and students learning (nine items). The inventory was judged by assessing each item using expert judgment regarding the relevancy of each item to the domain construct based on the defined four-point scale, which includes (i) the item not relevant to the measured domain, (ii) the item somewhat relevant to the measured domain, (iii) the item quite relevant to the measured domain, and (iv) the item highly relevant to the measured domain (Table 3). Equally, the expert was also

asked to evaluate the degree of language appropriateness regarding clarity, suitability, and ease of understanding. The instrument was developed by Brown (2004) within the theoretical framework delineated by the other literature on classroom assessment (Rural, 2021). The COA III Abridge Version is very relevant and applicable in gathering information for professional development and policy planning in such a way that assessment increases the student's achievement standard and enhances the quality of lecturers' teaching.

Table 3

Items on the conceptions of assessment subscales relate to the four main purposes of assessment

Assessment Conceptions Sub-scales	Items Number
Institutional accountability	1, 19, 10
Students' accountability	2, 20, 11
Improvement of teaching and learning	4, 3, 22, 12, 5, 14, 13, 21, 23, 24, 6, 15,
Assessment is Irrelevance	9, 18, 8, 27, 7, 25, 17, 16, 26

Source: Brown (2006)

Data Collection Procedure

Approval to partake in the study was sought from the expert judges in two ways: some in person and others by email to obtain their agreement to take part in the study to collect the required data. The online procedure for the assessment and validation process was selected because some experts were from different geographical locations. The panel judges were informed about the purpose of the study and their roles in being selected. They were provided with the content validation form and detailed instrument validation instructions. The expert judges were asked to provide

objective and constructive recommendations on improving the sentence structure and clarity of the items. The expert panels were given three weeks to assess and validate the 27 items of the adapted instrument.

Data Analysis

The content validity ratio (CVR) method is employed to assess the content validity of the adapted instrument: Conceptions of Assessment III Abridged Survey by Brown (2006). CVR is an item measurement suitable for accepting or rejecting individual items and is universally recognized as a way of reporting content validity (Zamanzadeh

et al., 2015). The content validation ratio engages a group of expert panels to examine the degree to which each item reflects the domain construct of an instrument. Therefore, the expert judges for this study were asked to provide their judgments regarding the relevancy of each item to the domain construct based on the defined four-point scale (Table 3). The CVR technique was chosen in this study as it is practical in terms of time and cost, straightforward, simple, and easy to apply.

Adjustment of Lawshe CVR Model

As stated earlier, the content validation ratio is used to determine the viewpoints of the panel of experts. In the earliest Lawshe Model, CVR engages a group of expert judges to assess the suitability of an instrument's items that reflect the domain construct on a three-point scale: (1) essential, (2) useful but not essential, and (3) not essential. However, there have been some criticisms of Lawshe's CVR model in assessing the agreement and response of the panels (Chalavi et al., 2015). Thus, to prevent different misinterpretations related to Lawshe's codes and to provide more significant differentiation in panels' ratings, Lawshe's three-point rating scales were expanded to a four-point scale (Chalavi et al., 2015).

Quantification of Content Validity

The content validity of the adapted instrument is computed using the viewpoint of the panel experts. Thus, the agreement of judgments among experts' panels on the relevancy of including an item in the measure can be calculated by determining the content validity ratio (CVR). As such, the decisions of an expert who makes relevant (3) and highly relevant (4) selections will be calculated using the content validity ratio formula in Equation 1.

$$CVR = \frac{n_e - (\frac{N}{2})}{N/2} \qquad (1)$$

The n_e from the formula represents the number of panels that make an appropriate and highly relevant choice, while **N** refers to the total panels. The result from this calculation can be explained as the CVR value being closer to 1 when all the judges agree that the item is relevant or highly relevant. The CVR values range from 0 to 1 when over half of the judges make relevant (3) and highly relevant (4) choices. The CVR is negative when less than half of the judges make the selection relevant (3) and highly relevant (4). The CVR value acceptance criteria on items are based on the revised version of the reference table (Wilson et al., 2012), initially developed by Lawshe (1975). The revised version of the reference table is shown in Table 4, as adopted by Chong et al. (2021).

With seven experts, the CVR critical value for each item must be equal to or greater than 0.741 at α = .05 level of significance for two-tailed tests (Table 4). Any item that fails to reach a CVR value of 0.741 will be excluded from the instrument.

Table 4
CVR acceptance values based on Lawshe (1975) as revised by Wilson et al. (2012)

N	**Level of Significance for One-Tailed Test**					
	.1	.05	.025	.01	.005	.00
	Level of Significance for Two-Tailed Test					
	.2	.1	.05	.02	.01	.00
5	.573	.736	.877	.99	.99.	.99
6	.532	.672	.800	.950	.99	.99
7	.485	.622	.741	.879	.974	.99
8	.453	.582	.693	.822	.911	.99
9	.427	.548	.653	.775	.859	.99
10	.405	.520	.620	.736	.815	.97
11	.387	.496	.591	.701	.777	.93
12	.370	.475	.566	.671	.744	.89
13	.356	.456	.544	.645	.714	.85
14	.343	.440	.524	.622	.688	.82
15	.331	.425	.506	.601	.665	.79
16	.321	.411	.490	.582	.644	.77
17	.311	.399	.475	.564	.625	.75
18	.302	.388	.462	.548	.607	.72
19	.294	.377	.450	.534	.591	.70
20	.287	.368	.438	.520	.576	.69
21	.280	359	.428	.508	.562	.67
22	.273	.351	.418	.496	.549	.65
23	.267	.343	.409	.485	.537	.64
24	.262	.336	.400	.475	.526	.63
25	.256	.329	.392	.465	.515	.61
26	.251	.323	.384	.456	.505.	.60
27	.247	.317	.377	.488	.496	.59
28	.242	.311	.370	.440	.487	.58
29	.238	.305	.364	.432	.478	.57
30	.234	.300	.358	.425	.470	.56
31	.230	295	.352	.418	.463	.55
32	.227	.291	.346	.411	.455	.54
33	.223	.286	.341	.405	.448	.53
34	.220	.282	.336	.399	.442	.53
35	.217	.278	.331	.393	.435	.52
40	.203	.260	.310	.368	.407	.48

Source: Chong et al. (2021)

Computation of Mean Value of Judgments Made by Experts' Values

According to Lawshe's suggestion, the following conversion rule is adapted for the validation process to determine the mean judgment values concerning each instrument item (Chalavi et al., 2015).

1. Highly relevant or quite relevant is substituted with 2,
2. Somewhat relevant is substituted with 1,
3. Not relevant is substituted with zero.

The total value of highly relevant or quite relevant (2), somewhat relevant (1), and relevant (0) for each item is added up, and we divided by the total number of experts. The items that fail to meet the value of the minimum requirements are excluded from the final instrument.

Acceptance or Rejection of An Item (Decision Rule)

The following benchmarks are established to include or exclude items in the instrument:

- An item is unconditionally included if its CVR equals or exceeds 0.741. This value is set given the number of judges (7).
- An item is accepted if its CVR ranges from 0–0.741 and the numerical mean of judgment is equal to or greater than 1.5. Such a CVR value indicates that more than half of the panel experts made "highly relevant" or "quite relevant"

choices. The mean value equal to or greater than 1.5 indicates that the mean judgment is closer to "highly relevant" and "quite relevant" choices. On the other hand, the mean value equal to or greater than 1.5 suggests that the mean judgment is equal to or greater than 75% of the maximum mean (2), which is greater than the minimum acceptable (Chalavi et al., 2015).

- An item is either revised or omitted from the instrument if CVR is less than or equal to 0 and the mean is less than 1.5. These indices indicate that at least half of the panel did not judge it to be "highly relevant or quite relevant" (an essential item on Lawshe's scale) and that it possesses a mean of judgments that is closer to "somewhat relevant or not relevant" (an unessential item on Lawshe's scale).

RESULTS

Based on the analysis of Table 5, 12 of the initial 27 items failed to meet the inclusion criteria. Three items from the "improving teaching and learning" sub-construct, and nine items from the "assessment as irrelevant" sub-construct were deemed not essential to the overall framework of the study. Table 6 presents a detailed overview of the rejected items and their respective constructs.

Consequently, only fifteen items remained for further analysis. These items were distributed across three dimensions:

Table 5

CVR, means of judgment, and acceptance/rejection results of assessment belief construct

Item Number	Total (N=7) CVR ≤0.741	Mean of Judgments <1.50	Item Status	Item Number	Total (N=7) CVR ≤0.741	Mean of Judgments <1.50	Item Status
1	1.000	2.00	Accepted	15	**0.428**	**1.42**	**Rejected**
2	1.000	2.00	Accepted	16	1.000	2.00	Accepted
3	1.000	2.00	Accepted	17	**0.428**	**1.42**	**Rejected**
4	1.000	2.00	Accepted	18	**0.428**	**1.42**	**Rejected**
5	1.000	2.00	Accepted	19	**0.142**	**1.42**	**Rejected**
6	1.000	2.00	Accepted	20	**-0.142**	**0.85**	**Rejected**
7	1.000	2.00	Accepted	21	**-0.714**	**0.42**	**Rejected**
8	1.000	2.00	Accepted	22	**0.428**	**1.42**	**Rejected**
9	1.000	2.00	Accepted	23	**-0.428**	**0.71**	**Rejected**
10	1.000	2.00	Accepted	24	**-0.428**	**0.57**	**Rejected**
11	1.000	2.00	Accepted	25	**-0.428**	**0.57**	**Rejected**
12	1.000	2.00	Accepted	26	**-0.428**	**0.57**	**Rejected**
13	1.000	2.00	Accepted	27	**-0.428**	**0.28**	**Rejected**
14	1.000	2.00	Accepted				

Source: Authors' work

institutional accountability (3 items), students' accountability (3 items), and improvements in teaching and learning (9 items). Subsequently, these items would be subjected to confirmatory factor analysis to assess their construct validity.

Notably, two factors—institutional and student accountability—had only three indicators. While some researchers, such as Marsh et al. (1998), advocate for many indicators per factor, others, such as Kenny (1979) and Kenny et al. (1998), argue that a minimum of three indicators per factor is sufficient to avoid model identification problems. This perspective aligns with the work of Koran (2020), who recommended a minimum of three indicators per factor

to ensure the over-identification of all models and prevent complications arising from empirical under-identification. This approach ensures that the subsequent confirmatory factor analysis will yield reliable and valid results.

The study aims at achieving three objectives, and the results are presented and discussed based on these objectives. First, the results provide insights into the significance of assessment beliefs in the context of educational assessment practices. The assessment belief construct was evaluated based on the mean of judgments and CVR values. The findings demonstrate that some items were accepted (with a mean of judgments ≥ 1.50 and a CVR >

Table 6
Assessment belief construct and rejected items

Construct	Dimensions	Item number	CVR ≤ 0.741	$\bar{x}$ <1.50	Total item rejected
Assessment belief	Institutional Accountability	3	1.000	2.00	
			1.000	2.00	
			1.000	2.00	
	Students Accountability	3	1.000	2.00	
			1.000	2.00	
			1.000	2.00	
	Improvement of teaching and learning	12	1.000	2.00	3
			1.000	2.00	
			1.000	2.00	
			1.000	2.00	
			1.000	2.00	
			1.000	2.00	
			1.000	2.00	
			1.000	2.00	
			0.428	**1.42**	
			1.000	2.00	
			0.428	**1.42**	
			0.428	**1.42**	
	Irrelevant	9	**0.142**	**1.42**	9
			-0.142	**0.85**	
			-0.714	**0.42**	
			0.428	**1.42**	
			-0.428	**0.71**	
			-0.428	**0.57**	
			-0.428	**0.57**	
			-0.428	**0.57**	
			-0.428	**0.28**	
			Total items rejected		12

Source: Authors' work

0.741), while others were rejected due to lower mean judgments and CVR values. Secondly, the study aimed to assess the existing literature on assessment beliefs across different contexts. The literature review reveals that assessment beliefs vary across different contexts and countries. Cultural, societal, and educational factors play a significant role in shaping educators' beliefs about assessment. Thirdly, another objective was to determine the content-based validity of the adapted instrument,

specifically the "Brown Conception of Assessment III Abridged Survey," within the Nigerian context. The Lawshe Content Validation Ratio (CVR) results are presented in the table. Items with CVR values greater than 0.741 and mean judgments higher than 1.50 were accepted, while those failing to meet these criteria were rejected. The study thus quantitatively evaluated the validity of the adapted instrument.

In summary, the study objectives revolved around emphasizing the importance of assessment beliefs, evaluating the existing literature, and assessing the content validity of an adapted instrument. The provided results contribute to achieving these objectives by shedding light on the validity of the instrument and offering insights into how well it aligns with the targeted constructs, particularly within the Nigerian context.

Discussion on Findings

The study adapts Brown's four-factor model of assessment beliefs, which encompasses improvement of teaching and learning, institutional and student accountability, and irrelevance conceptions. This model serves as the basis for understanding the assessment beliefs of lecturers in the context of TVET Nigeria. The findings reveal a detailed content validity assessment process analysis using the Content Validity Ratio (CVR) method, which involves expert judgments. The CVR values and mean judgments for each item are presented, leading to acceptance or rejection decisions. It further revealed that contextual differences are an important factor that defines assessment beliefs at different levels of instructional contexts. This finding is supported by Brown et al. (2019), who opined that while the inventory can be used cross-culturally after localized adaptations, there is indeed no single universally applicable model. Context, culture, and local factors shape the belief in assessment. Equally, a study by Astuti (2015) confirmed the findings of this study, which found that most participants tend to disagree with the view that assessment is irrelevant.

Thus, based on the findings of this study, 15 items are retained. These items were distributed across three dimensions: institutional accountability (3 items), students' accountability (3 items), and improvements in teaching and learning (9 items). The findings exclude the dimension of "assessment as irrelevant," suggesting its limited relevance in the context of TVET Nigeria. Furthermore, the outcome addresses a gap in understanding lecturers' viewpoints on assessment in the Nigerian TVET context, where research is scarce. Equally, the findings align with existing research suggesting lecturers prioritize assessment to improve teaching and learning, similar to findings in other contexts like Malaysia.

CONCLUSION

This study aimed to adapt an assessment belief instrument for use among lecturers in the Nigerian TVET context. The initial 27-item instrument, based on Brown's model, underwent a rigorous content validation process involving expert judgments.

Following analysis, twelve items were excluded. Remarkably, nine items belonged to the "assessment as irrelevant" sub-construct, suggesting that this dimension might not be relevant in the Nigerian TVET context. Additionally, three items were removed from the "improving teaching and learning" sub-construct. Table 6 provides a detailed breakdown of the rejected items.

Therefore, only fifteen items remained for further analysis. These retained items are distributed across three dimensions: institutional accountability (3 items), students' accountability (3 items), and improvement in teaching and learning (9 items). In the next stage of the research process, these items will be subjected to confirmatory factor analysis. This analysis will assess their construct validity and determine if they accurately measure the intended dimensions of assessment beliefs in the Nigerian TVET context.

In conclusion, this study has successfully reduced the initial instrument to a more concise and contextually relevant set of 15 items. The next phase of the research will involve confirmatory factor analysis to evaluate the construct validity of these items and ensure they accurately measure the intended dimensions of assessment beliefs among lecturers in Nigerian TVET institutions.

Implications for Theory and Practice

This study regarding assessment beliefs in the Nigerian TVET context has offered valuable insight into theory and practice.

On the theory side, the study reinforces the idea that assessment belief models like Brown's require contextual adaptation. The finding that the "assessment as irrelevant" dimension was excluded suggests that Nigerian TVET lecturers generally value assessment and its role in education. It contributes to the ongoing development of contextually sensitive models of assessment beliefs. Furthermore, the three identified dimensions—institutional accountability, student accountability, and improvement of teaching and learning—provide a focused framework for understanding lecturers' assessment beliefs in Nigerian TVET. This framework can guide future research to explore these dimensions in greater depth.

The findings can be used to develop assessment practices that align with lecturers' beliefs in TVET, leading to more effective learning and improved student outcomes. The emphasis on institutional accountability, student accountability, and improvements in teaching and learning provides a practical framework for TVET lecturers to design and implement meaningful assessments. Equally, policymakers can use the findings to develop assessment policies that harmonize with lecturers' beliefs and promote effective teaching and learning.

Limitations and Recommendations for Future Research

While offering valuable insights, this study has limitations that suggest paths for future research. The study involved a limited number of expert judges for content validation. Expanding the pool of

judges in future studies can enhance the generalizability of the findings. Moreover, the study focused solely on lecturers' assessment beliefs. Including students' perspectives on assessment in future research could provide a more holistic understanding of the assessment landscape in Nigerian TVET.

On the other hand, as mentioned in the conclusion, the next phase of this research should involve confirmatory factor analysis to assess the construct validity of the remaining 15 items. It will ensure that the three identified dimensions—institutional accountability, student accountability, and improvement of teaching and learning— are accurately measured. Additionally, once the instrument's construct validity is established, future research can delve into the predictive validity of the instrument. It would investigate the relationship between lecturers' assessment beliefs and their actual assessment practices to provide professional development interventions.

REFERENCES

Akmal, N., Nasir, M., Singh, P., Narayanan, G., Hani, A., Habali, M., & Rasid, S. (2022). Development of mathematical thinking test: Content validity process. *ESTEEM Journal of Social Sciences and Humanities, 6*(2), 18-29.

Azis, A. (2015). Conceptions and practices of assessment: A case of teachers representing improvement conception. *TEFLIN Journal, 26*(2), 129-154. http://doi.org/10.15639/teflinjournal.v26i2/129-154

Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 284-300). Routledge. https://www.researchgate.net/publication/283326946_Teachers'_beliefs_about_assessment

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551-575. https://doi.org/10.1080/0969594X.2018.1441807

Brown, G. T. L. (2003). *Teachers' conceptions of assessment* (Doctoral dissertation, University of Auckland). DuraSpace. https://researchspace.auckland.ac.nz/handle/2292/63

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice, 11*(3), 301-318. https://doi.org/10.1080/0969594042000304609

Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged version. *Psychological Reports, 99*(1), 166-170.

Brown, G. T. L. (2011). New Zealand prospective teacher conceptions of assessment and academic

performance: Neither student nor practicing teacher. In R. Kahn, J. C. McDermott & A. Akimjak (Eds.), *Democratic access to education* (pp.119-132). Antioch University. file:///C:/Users/OPTIPLEX%205250%20AIO/Downloads/2011BrownprospectiveteachersTCoAandperformance.pdf

Brown, G. T. L., & Hirschfeld, G. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice, 15*(1), 3-17. https://doi.org/10.1080/09695940701876003

Brown, G. T. L., & Harris, L. R. (2009). Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of Multidisciplinary Evaluation, 6*(12), 68-91. https://doi.org/10.56645/jmde.v6i12.236

Brown, G. T. L., Gebril, A., & Michaelides, M. P. (2019) Teachers' conceptions of assessment: A global phenomenon or global localism. *Frontiers in Education, 4,* Article 16. https://doi.org/10.3389/feduc.2019.00016

Calveric, S. B. (2010). *Elementary teachers' assessment beliefs and practices* [Doctoral Dissertation, Virginia Commonwealth University]. VCU Scholars Compass. https://doi.org/10.25772/0XTP-8W56

Chalavi, L. H., Alamdari, S., & Asadi, K. (2015). Improving indices in evaluation of medical research centers in Iran. *International Journal of Advanced Research in Biological Sciences*, *2*(2), 41-52.

Chong, J., Mokshein, S. E., & Mustapha, R. (2021). A content validity study for vocational teachers' assessment literacy instrument (VoTAL). *International Journal of Academic Research in Business and Social Sciences*, *11*(4), 868-883. https://doi.org/10.6007/IJARBSS/v11-i4/9547

Elshawa, N. R. M., Abdullah, N. A., & Md, Rashid. S. (2017). Malaysian instructors' assessment beliefs in tertiary ESL classrooms. *International Journal of Education and Literacy Studies*, *5*(2), 29.

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students' conceptions of assessment in higher education. *Higher Education*, *64*(1), 119–133. https://doi.org/10.1007/s10734-011-9484-1

Galicia Alarcón, L. A., Balderrama Trápaga, J. A., & Edel Navarro, R. (2017). Content validity by experts' judgment: Proposal for a virtual tool. *Apertura*, *9*(2), 42–53. https://doi.org/10.32870/Ap.v9n2.993

Ha, X. V., Murray, J. C., & Riazi, A. M. (2021). High school EFL students' beliefs about oral corrective feedback: The role of gender, motivation and extraversion. *Studies in Second Language Learning and Teaching, 11*(2), 235- 264. https://doi.org/10.14746/ssllt.2021.11.2.4

Ha, X. V., Tran, N. G., & Tran, N. H. (2021). Teachers' beliefs and practices regarding assessment in English as a foreign language classroom in Vietnam. *The Qualitative Report, 26*(11), 3457-3475. https://doi.org/10.46743/2160-3715/2021.5063

Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, *16*(3), 365–381. https://doi.org/10.1080/09695940903319745

Koran, J. (2020). Indicators per factor in Confirmatory Factor Analysis: More is not always better. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(5), 765–772. https://doi.org/10.1080/10705511.2019.1706527

Kenny, D. A. (1979). *Correlation and causality*. Wiley. https://www.abebooks.com/9780471024392/Correlation-Causality-Kenny-David-A-0471024392/plp

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S., Fiske & G. Lindzey (Eds.), *The handbook of social Psychology* (pp. 233-265). McGraw-Hill. https://www.columbia.edu/~nb2229/docs/KennyKashyBolger1998-Data_analysis.pdf

Lawshe, C. H. (1975). A Quantitative approach to content validity. *PersonnelPsychology*, *28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975. tb01393.x

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in Confirmatory Factor Analysis. *Multivariate Behavioral Research*, *33*(2), 181–220. https://doi.org/10.1207/ s15327906mbr3302_1

Mazin, K. A., Norman, H., Nordin, N., & Ibrahim, R. (2020). Student self-recording videos for TVET competency in MOOCs. *Journal of Physics: Conference Series*, *1529*(4). https://doi. org/10.1088/1742-6596/1529/4/042061

Mohd, R. M., Sardareh, S. A., & Ambarwati, E. K. (2013). Iranian secondary school EFL teachers' assessment beliefs and roles. *Life Science Journal, 10*(3), 1638-1647.

Moiinvaziri, M. (2015). *Teachers' conception of assessment: A structural equation modeling approach*. https://api.semanticscholar.org/ CorpusID:113404239

Muñoz, A. P., Palacio, M., & Escobar, L. (2012). Teachers' beliefs about assessment in an EFL context in Colombia. *Profile, 14*(1), 143-158.

Narathakoon, A., Sapsirin, S., & Subphadoongchone, P. (2020). Beliefs and classroom assessment practices of English teachers in primary schools in Thailand. *International Journal of Instruction, 13*(3), 137-156. https://doi.org/10.29333/ iji.2020.13310a

Opre, D. (2015). Teachers' conceptions of assessment. *Procedia - Social and Behavioral Sciences*, *209*, 229–233. https://doi.org/10.1016/j. sbspro.2015.11.222

Osman, S., Bordoh, A., & Eshun, I. (2021). Basic school teachers" conceptions of assessment in the Sissala East Municipality. *International Journal of Research and Innovation in Social Science*, *5*(3), 311–324. https://doi.org/10.47772/ IJRISS.2021.5320

Paris, D., Eaton, J., & Geary Schneider, C. (2012). *Committing to quality guidelines for assessment and accountability in higher education.* https:// www.learningoutcomesassessment.org/wp-content/uploads/2019/08/Committing-to-Quality.pdf

Raji, M. O., Daramola, D. S., & Oladele, J. I. (2020). Teachers' conceptions and choices of assessment tasks in a Nigerian postgraduate teacher training. *Asian Journal of Assessment in Teaching and Learning*, *10*(2), 82–93. https://doi. org/10.37134/ajatel.vol10.2.9.2020

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education, 27*(2), 472–482. https://doi.org/10.1016/j.tate.2010.09.017

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94–104. https://doi.org/10.1093/ swr/27.2.94

Rural, J. D. (2021). Teachers' conceptions of assessment using Brown's Four-Factor Model and DepEd assessment policy: Implications for policy review. *Turkish Journal of Computer and Mathematics Education* (TURCOMAT), *12*(3), 5320–5326. https://doi.org/10.17762/turcomat. v12i3.2170

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice, 37*(1), 21–34. https://doi. org/10.1111/emip.12189

Sulaiman, T., Kotamjani, S. S., Rahim, S. S. A., & Hakim, M. N. (2020). Malaysian public university lecturers' perceptions and practices of formative and alternative assessments. *International Journal of Learning, Teaching and Educational Research*, *19*(5), 379–394. https:// doi.org/10.26803/IJLTER.19.5.23

Taherdoost, H., (2016). Validity and reliability of the research instrument: How to test the validation of a questionnaire/Survey in research. *International Journal of Academic Research in Management (IJARM), 5*(3), 28-36. https://doi.org/10.2139/ssrn.3205040

Thomas, M. (2012). Teachers' beliefs about classroom assessment and their selection of classroom assessment strategies. *Journal of Research and Reflections, 6*(2), 103-112.

Vandeyar, S., & Killen, R., (2007). 'Educators' conceptions of classroom assessment in post-apartheid South Africa', *South African Journal of Education, 27*(1), 101-115.

Vardar, E. (2010). *Sixth, seventh and eighth grade teachers' conception of assessment* [Master's thesis, Middle East Technical University]. OpenMETU. https://etd.lib.metu.edu.tr/upload/12612141/index.pdf

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development, 45*(3), 197–210. https://doi.org/10.1177/0748175612440286

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58,* 149–162. http://doi.org/10.1016/j.tate.2016.05.010

Yidana, M. B., & P., A. P. (2018). Economics teachers' conceptions of classroom assessment. *International Journal for Innovation Education and Research*, *6*(10), 153–174. https://doi.org/10.31686/ijier.vol6.iss10.1176

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Yusop, S. R. M., Rasul, M. S., Yasin, R. M., & Hashim, H. U. (2023). Identifying and validating vocational skills domains and indicators in classroom assessment practices in TVET. *Sustainability*, *15*(6), Article 5195. https://doi.org/10.3390/su15065195

Yusop, S. R. M., Rasul, M. S., Yasin, R. M., Hashim, H. U., & Jalaludin, N. A. (2022). An assessment approaches and learning outcomes in technical and vocational education: A systematic review using PRISMA. *Sustainability*, *14*(9), Article 5225. https://doi.org/10.3390/su14095225

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences*, *4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017

Zapata-Ospina, J. P., & García-Valencia, J. (2022). Validity based on content: A challenge in health measurement scales. *Journal of Health Psychology*, *27*(2), 481–493. https://doi.org/10.1177/1359105320953477

Factors Affecting Livelihood Diversification of Women of Farm Families Working in Textile Industries in Faisalabad District, Pakistan

Munaza Zafar¹, Rana Muhammad Amir¹*, Hafiz Ali Raza¹, Sidiqat Aderinoye-Abdulwahab², Inam Irshad³ and Tawalkalitu Abiola Dolapo²

¹*Institute of Agricultural Extension, Education and Rural Development, University of Agriculture Faisalabad, Punjab, Pakistan*
²*Department of Agricultural Extension and Rural Development, University of Ilorin, Ilorin, Nigeria*
³*Institute of Soil and Environmental Sciences, University of Agriculture, Faisalabad, Punjab, Pakistan*

ABSTRACT

The economy of Pakistan is agro-based, and rural women's livelihoods are linked to men's. In the meantime, livelihood diversification and labor mobility have impacted working men and women. This study highlights the factors influencing women's decision-making abilities to diversify livelihood strategies from on-farm to off-farm activities in Faisalabad district, Pakistan. Two textile industries were selected in Faisalabad, and 115 females participated in the study. The interview schedule was used to elicit information from women from farm families working in the textile industries. Data collected were analyzed using descriptive and inferential statistics. Most women reported that daily-based wages (85.2%), poultry (63.5%), livestock (55.7%), farming (40.9%) and casual labor (40.0%) were among the available livelihood diversification strategies in use. The foremost factors influencing the decision to engage in more sources of income were inter-role conflict (3.78±.88), family pressure (3.71±.81), and stereotypic influences (3.68±.96). It was concluded that the women who have diversified their livelihood were faced with gender-based discrimination (3.84±.67), lack of education (3.78±.73), non-cooperation of colleagues (3.77±.82), and low wages (3.69±.76) in the textile industry. The study suggested gender-friendly initiatives that will minimize conflicts and family pressures, particularly for women, as this will foster a successful livelihood diversification approach.

Keywords: Decision-making, diversification, farming households, income, livelihood, non-farm

INTRODUCTION

The agricultural sector, particularly at the rural level, is indispensable to Pakistan's economic growth, food security, employment generation, and poverty alleviation. It contributes 19.2% to the gross domestic product (GDP) and employs around 38.5% of the labor force. More than 65–70 percent of the population depends on agriculture for its livelihood (Pakistan Economic Survey, 2015), and agriculture remains a high-risk enterprise while women's role has continued to generate resource base and power issues; yet they must find ways to cope with a variety of socio-economic, environmental, and decision-making capabilities daily (Chimgonda-Nkhoma et al., 2019). Given the unpredictability of farmers' revenue due to the high-risk nature of agriculture, off-farm diversification is a key approach for agricultural households to safeguard their incomes (Iqbal et al., 2016). Climate shocks such as floods, droughts, insufficient agricultural yields, market failures and price changes have necessitated that farmers have had to transfer their labor to off-farm businesses to sustain household food demand (Aderinoye-Abdulwahab et al., 2015). Additionally, poor rural families are more predisposed to seasonal risk variables than urban families; hence, the capacity to diversify into multiple income sources is important for their survival. Women account for about half of the population in Pakistan. They contribute substantially to the labor force and perform as many tasks as the males in agricultural activities. In addition, they also take care of their homes and families.

Female workers in various parts of the world, including Asia, put in long hours from dawn to night-time, yet society does not recognize their efforts. They are also not adequately compensated, particularly in rural and most urban regions. Hence, women face several hurdles compared to their male counterparts. Family life balance, arduous work and its conditions, family and society-related assistance, culture and norms-related challenges, and workplace structure are a few of the difficulties women encounter (Javaid et al., 2020). Working women of all ages, married or single, face multiple stereotyped obstacles at work, and these cause stress, sadness, and lack of confidence. These make them unable to strike a balance between work and family life owing to job insecurity issues (Javaid et al., 2020).

It has been observed that rural livelihood diversification is usually influenced by a variety of push and pull forces. Despite the recognized importance of improving rural household income, Pakistan remains one of the least studied countries while trying to understand how rural household inclination and off-farm rural activities impact incomes at the micro level. The lack of capital, poor infrastructure, lack of access to finance, lack of access to markets, and shortage of farmlands are among the problems rural households encounter in diversifying their livelihoods (Kassa, 2019). Certain attitudes, such as desire, past behavior and perceived behavioral control, equally influence decision-making ability (Sayginer & Kurtsan, 2022). For these reasons, this study aims to analyze the

decision-making ability of women in farm families in Faisalabad District, Pakistan, with special emphasis on their livelihood diversification status.

Studies have equally shown the economic activities and livelihood diversification of women in rural Faisalabad, Pakistan (Haq et al., 2017; Haq et al., 2021), but no study has presented factors responsible for the choice of livelihood activities of the women. Thus, it is evident that there is a dearth of knowledge of the factors that affect women's decision to diversify their livelihoods in textile industries in Faisalabad District, Pakistan. Consequently, the main objective of this study is to provide empirical data for a clear understanding of the factors affecting their livelihood diversification.

The specific objectives were to:
1. Describe the socio-economic characteristics of the women in the textile industry in Faisalabad, Pakistan.
2. Determine the livelihood diversification strategies of rural women in the textile industry in Faisalabad, Pakistan.
3. Identify the factors influencing the decision-making ability of the women in the textile industry to diversify their livelihoods.
4. Identify the constraints faced by the women in the textile industry in Faisalabad, Pakistan.

Literature Review

Irrespective of the region or the type of civilization around the world, working women have faced a variety of difficulties in human culture. In Asian organizations, particularly in the Indo-Pak region, the problem is getting worse. Furthermore, the challenges become more intense when it comes to working women. Numerous women from the Indo-Pak sub-continent continue to face serious challenges, especially in terms of financial rewards and social perspectives. As a result of Pakistan's socio-cultural structure, the country's labor force is not treated equally in terms of capacity and productivity. This tendency imposes numerous limits, such as gender-based discrimination on female labor market participation.

Small-scale farmers' livelihood diversification is critical for reducing hunger and mitigating the negative effects of climate change, given that agriculture continues to be the most important source of income. Despite various agricultural promotion measures, Pakistan's agricultural productivity is still rated low, even as livelihood diversification from agriculture tends to be gaining more prominence. It, therefore, becomes more pertinent for small-scale farmers' livelihoods to be improved to promote livelihood diversification. It can be achieved if the government takes appropriate policy actions, including creating national jobs (Shah et al., 2021). Such strategies and policies had previously placed heavy emphasis on farming-related activities while ignoring the potential for non-farm diversification, even though non-farm sources of income have often proven productive (Naudiyal et al., 2019). The

textile industry is equally one sector that offers great potential in terms of off-farm livelihood opportunities.

Farm households must diversify their income streams to meet their food and subsistence needs. However, it has not always been easy to grow income sources for rural residents and small-scale farmers due to labor market constraints, land market conditions, unavailability of natural resources, and lack of institutional support (that aids farmers' yield improvement and the choice of more diversification options for livelihoods) (Aderinoye-Abdulwahab et al., 2015; Shah et al., 2021). Diversification is the process of creating income-generating portfolios with low covariate risk (Asravor, 2018). Hence, the ultimate purpose of livelihood diversification is to provide long-term benefits such as greater income, enhanced food security, reduced vulnerability, and increased welfare (Dinku, 2018). Human, financial, social, and natural- and physical assets are some determinants that influence rural communities to broaden their livelihood potentials from on-farm to off-farm activities. Diversification of livelihoods is a common strategy for dealing with economic and environmental shocks, as well as a key component of poverty reduction (Gautam & Andersen, 2016).

Theorizing the Phenomenon

This study adopts three theories that guide livelihood diversification. They are household economic theory (Singh, 1986), livelihood approach theory (Chambers & Conway, 1992) and theory of choice (Etuk et al., 2018). The household theory regards farm households as production units that maximize satisfaction. It is achieved by combining time and other inputs to generate output, but subject to price and resource constraints. According to the theory, diversification is a function of returns to labor from farm activities. The theory helps to examine household production and off-farm labor allocation decisions.

The livelihood approach theory has been used in studies relating to livelihood strategies, poverty, and livelihood diversification. The theory states that "livelihood diversification is beneficial to the rural poor because it reduces risks and stabilizes income flow and consumption. It leads to improvement in quality of life, wealth accumulation and food security". The theory aids in the assessment of people's livelihood assets. According to the theory of choice, diversification represents the initiative of selecting diversity over consistency (De Giorgi & Mahmoud, 2016). Several studies have shown that the socio-economic characteristics of rural women influenced their decision to adopt livelihood diversification. According to Kongla and Akoh (2019), being unmarried and divorced negatively influenced the diversification of livelihood. On the other hand, age is a positive factor that influenced the choice and nature of diversification of rural women's activities in Cameroon. Oludipe (2019) reported that the household size of rural women had no significant

influence on livelihood diversification in Nigeria, while Islam et al. (2022) also reported that educational status, family size, work experiences, personal income, savings, and training significantly affected rural women's involvement in off-farm activities in Bangladesh. Hafeez and Ahmed (2002) found females' educational level to be a strong and positive factor that influenced the decision of educated married women to participate in the labor market in the Punjab district of Pakistan, while Habib et al. (2022) similarly found that educational level was a significant factor for adopting livelihood diversification among rural women in Pakistan.

Thus, the study tested a hypothesis using Spearman's correlation to further establish the relationship between rural women's socio-economic factors and their decision-making ability in the textile industry in Faisalabad, Pakistan.

H1: Socio-economic factors do not significantly affect the decision to diversify livelihood.

METHODS

Study Area

The current study, which focuses on the analysis of livelihood diversification of farm families with special reference to decision-making by the women working in textile industries, was conducted in Faisalabad district, Pakistan (Figure 1). The textile sector in Pakistan employs over 15 million people and accounts for 9.5% of the country's GDP. Pakistan is Asia's fourth-largest cotton producer, with the third-largest spinning capacity after China and India. The country accounts for 5% of the global capacity of cotton production. In many developing countries, like Pakistan, women are the primary labor providers in the textile sector, which is notorious for its poor working conditions and lack of worker rights (Mahboob & Anita, 2016). Asia exports 58.4% of all clothes and textiles in the globe (Khan et al., 2020), while textiles involving regional diversity in embroideries may be found in many regions

Figure 1. Map of district Faisalabad, Pakistan
Source: Authors' work

of Pakistan. Punjab (primarily Faisalabad) is the center of the textile industry, while women's contributions to the development of textiles and textile-related industries have received little attention (Baghal et al., 2019). District Faisalabad was chosen for this study because, compared to other cities, it is Pakistan's Manchester because of its abundant textile and other industries.

Research Design

The research is explanatory in nature and aims to provide a clear understanding of the factors affecting the livelihood diversification of women in farm families working in textile industries in Faisalabad District, Pakistan. The study relied on descriptive and inferential statistics to analyze data collected from the participants.

Sample Size

A sample is the smallest single unit that possesses all the characteristics of a population. Simple random sampling was used to select two textile industries for the study. These industries specialize

Figure 2. Determination of sample size
Source: Sample Size Calculator (n.d)

in spinning, yarn dyeing, stitching and power plants. Pre-survey visits to these industries showed that 163 farm women were working there. Therefore, 163 female laborers in these industries constituted the study population from which the sample was drawn. A sampling calculator with a 5% confidence level and 95% confidence interval was used to determine the study's sample size (Figure 2). It allows for determining a sample that can produce results that precisely capture the target population. The sampling calculator gave a sample percentage of 70.5%, which translates into 115 women from the entire population of 163. Thereafter, the names of the 163 women were listed on paper, a number was assigned to each name from 1 to 163, and every alternate number was picked until a total of 115 sample size was attained.

Research Instrument and Measurement

Female workers were interviewed in the textile industry, and the interview schedule (See Appendix) was prepared in English to facilitate the process of getting precise information. However, the questions were delivered in the respondents' local languages (Urdu and Punjabi). The questions were interpreted and validated by experts with a full grasp of both languages to ensure accurate information about the research. A reliable and valid interview schedule was developed to collect data that covered the objectives of the research. Before finalizing the interview schedule, relevant experts ensured face and content validity. Reliability

was measured by carrying out a pre-test on ten respondents. After taking responses from the respondents, the Cronbach Alpha test was used to check its reliability. The value of the Cronbach Alpha test was 0.811, which means it was fit and reliable for data collection.

Respondents were asked to indicate their level of education on a four-point Likert-type (1932) ordinal scale of illiterate, primary, middle, and high school while 1, 2, 3 and 4 numerals were assigned, respectively. Respondents were also provided with a list of possible sources of earning additional income for livelihood diversification. Such income sources included farming, agricultural labor, casual labor, and government employment, and they were requested to indicate the diversification strategies being used on a 2-point nominal scale of yes (diversified) or no (no diversification). At the same time, 1 and 0 numerals were assigned, respectively. Regarding constraints faced in the textile industry, respondents were given a list of 10 possible challenges (lack of skills, lack of communication skills, lack of education) to indicate the severity on a five-point Likert-type ordinal scale of very low, low, medium, high, and very high. At the same time, 1, 2, 3, 4 and 5 numerals were assigned respectively. Respondents were asked to indicate the level of effect of decision-making ability from the list of 7 factors (illiteracy, male dominance, family pressure) to determine the factors affecting respondents' decision to adopt livelihood diversification. A five-point Likert-type ordinal scale of very low, low, medium, high, and very high was used with 1, 2, 3, 4 and 5 numerals assigned, respectively.

Data Analyses

The collected data were prepared and arranged on the spreadsheet for analysis. Statistical Package of Social Sciences (SPSS) software was used to analyze the data using different techniques. These techniques were descriptive analysis (frequency, percentage, mean and standard deviation). At the same time, chi-square and Spearman's correlation were used for inferential analysis between socio-economic characteristics and the sample's decision-making ability.

RESULTS AND DISCUSSION

Objective 1: Socio-economic Characteristics of Rural Women in Textile Industry in Faisalabad, Pakistan

It is important to state that economic activities and livelihood diversification of women in rural Pakistan have already been recorded (Haq et al., 2017; Haq et al., 2021), but the factors responsible for the rationale behind women's choice of livelihood activities are yet to be documented. The current study, therefore, bridges the gap between the decision-making abilities of women working in textile industries and their livelihood diversification strategies in the Faisalabad district of Pakistan. The study found that a considerable part of the sample was relatively young, as 44% were in their thirties (Table 1). Age is depicted by the respondent's chronological number of

years counted from the first day of birth of the individual to the moment of information collection. Table 1 showed that 44.4% of the participants belonged to the third age group (29–33 years). It shows that female workers in textile industries were young and would be expected to be agile enough to contribute productively to the textile industry and family farming. The study found that a good proportion of the respondents were not lettered. Regarding respondents' education, around one-fourth (27.8%) of the participants were illiterate, indicating that the remaining majority were literate. High-level literacy among female textile workers is expected to help them be analytical and make logical decisions, consequently improving their decision-making power in various farm families. Islam et al. (2022) similarly opined that socio-economic factors such as educational status, family size, and work experiences significantly affected rural women's involvement in off-farm activities in Bangladesh, while Hafeez and Ahmed (2002) found that females' educational level is a strong and positive factor that could influence the decision of educated married women. In the same vein, Habib et al. (2022) also found that educational level was a significant factor in adopting livelihood diversification among rural women in Pakistan.

On marital status, a significant proportion (64.3%) of the participants were married, divorced, or widowed in the same manner that the majority (67.0%) were living in joint or extended family systems. Kongla and Akoh (2019) also affirmed that

being unmarried and divorced negatively influenced the diversification of livelihoods of rural women in Cameroon. These findings implied that the decision-making ability of the married females working in the textile industry might have been compromised and could have been transferred or designated

Table 1
Socio-economic characteristics of rural women in the textile industry in Faisalabad, Pakistan (n=115)

	f	**%**
Age groups		
19–23 years	26	22.6
24–28 years	26	22.6
29–33 years	51	44.4
34–38 years	12	10.4
Education		
Illiterate	32	27.8
Primary (1–5)	33	28.7
Middle (6–8)	45	39.1
High school (9–12)	5	4.4
Marital status		
Married, Divorce, Widow	74	64.3
Unmarried	41	35.7
Family structure		
Nuclear	38	33.0
Joint	77	67.0
Predominant occupation of the respondent		
Daily wages	70	60.9
Farming	28	24.3
Small Business	6	5.2
Others	11	9.6
Land holdings		
Yes	7	6.1
No	108	93.9
Income (PKR)		
5000–8000	3	2.6
8001–10000	28	24.4
10001–14000	69	60.0
Above 14000	15	13.0

Source: Authors' work

to their husbands or even the heads of the families in situations where such women were widowed. Such decisions might include matters related to work done in the textile industry or even activities on family farms.

Results in Table 1 further showed that the women engaged in hired labor, farming, and petty businesses. It was found that a substantial proportion (60.9%) were predominantly laborers with daily wage earnings; 24.3% relied heavily on farming as their main occupation, while only 5.2% engaged in small/petty businesses. Again, it was reported that personal income and savings could influence livelihood diversification (Islam et al., 2022). It shows that economic engagements that will yield daily wages were the predominant occupation and main sources of income for female textile workers in the study. Food is not only necessary to keep body and soul together, but it also serves to an end in the same manner that money is required to eat, live a good life, and obtain the other necessities of life. As a result, people take up various occupations; therefore, it was not surprising that the respondents in this study diversified into different occupations (on-farm and off-farm). It aligns with the household theory, which opined that farming households, which serve as production units, are needed to maximize satisfaction, hence the need for diversification.

On land holdings, only 6.1% of the sample had agricultural land, and a majority (93.9%) were landless. This finding may be attributed to the issue of marriage as husbands (males) in most rural settings owned the land while their wives remained landless. Therefore, the decision-making ability to diversify into farming and the size of land cultivated by the female textile industry workers may be weak, as this greatly depended on their husbands. It is also noteworthy that women are usually not allowed to participate in family decision-making meetings, let alone contribute to matters relating to access and ownership of productive resources (Bayeh, 2017), and women have grown to accept such subsidiary roles in traditional African societies. Also, in Table 1, a large portion (60.0%) of the participants earned Rs. 10001–14000, the equivalent of less than $100 (≡$53.27–$74.57). The livelihoods approach theory established that livelihoods can only be sustainable when physical, social, and economic assets are maintained at a level that translates to improved quality of life. It posits that livelihood diversification benefits the rural poor by reducing socio-economic risks and stabilizing income flow, ultimately resulting in enhanced livelihoods.

Objective 2: Livelihood Diversification Strategies for Rural Women in the Textile Industry in Faisalabad, Pakistan

Empirical evidence from this study showed that the textile sector, through women's economic contribution, continued to impact farm families significantly. Yet, there is no documentary evidence to show that the government has made/implemented policies that could help to improve their productivity. There might be a need to

formulate relevant policies as well as implement rural development programs aimed at guiding the decision of rural women on existing opportunities for livelihood activities and potential impacts on the rural economy of Pakistan. Table 2 shows that a substantial part of the respondents adopted and engaged in multiple sources of income, including daily-based wages (85.2%), poultry (63.5%), livestock (55.7%), farming (40.9%) and casual labor (40.0%).

The findings are in consonance with previous research works where it had been established those women not only engaged in farming activities but also diversified into various off-farm and non-farm activities (Aderinoye-Abdulwahab et al., 2015; Kassa, 2019; Shah et al., 2021). It is also in line with the theory of choice, where, in this instance, diversification represents the initiative to select from a range of livelihood diversification strategies rather than maintaining a monotonous means of living (farming) (De Giorgi & Mahmoud, 2016). Therefore, it becomes highly pertinent to understand and acknowledge the household diversification strategies, existing opportunities, and potential impacts of rural women in textile industries in the selected rural economy of Pakistan.

Objective 3: Factors Influencing the Decision-making Ability of Women to Diversify Livelihood Strategies in the Textile Industry in Faisalabad, Pakistan

Findings showed that certain indices, such as family pressures, inter-role conflicts, and stereotypic influences, triggered the decision-making abilities of farm women in this study. Data specified the factors responsible for women's decision to diversify their livelihood strategies, which pushed them to adopt multiple sources of income (Table 3).

Table 2

Livelihood diversification strategies of rural women in the textile industry in Faisalabad

Source of income	Yes		No	
	f	*%*	*f*	*%*
Farming	47	40.9	68	59.1
Agricultural labor	19	16.5	96	83.5
Casual labor	46	40.0	69	60.0
Government employment	37	32.2	78	67.8
Daily based wages	98	85.2	17	14.8
Business	9	7.8	106	92.2
Remittances	8	7.0	107	93.0
Cottage industries	16	13.9	99	86.1
Livestock	64	55.7	51	44.3
Poultry	73	63.5	42	36.5
Others	18	15.7	97	84.3

Note. Multiple responses were allowed
Source: Authors' work

Table 3

Factors affecting rural women's livelihood diversification in the textile industry in Faisalabad, Pakistan

Constraints	W.S.	Mean	Rank
Gender-based discrimination	442	3.84	1
Lack of education	435	3.78	2
Colleagues are non-cooperative	433	3.77	3
Low wages	424	3.69	4
Lack of training	421	3.66	5
Lack of communication skills	415	3.61	6
Transport problem	414	3.60	7
Lack of self-confidence	405	3.52	8
Lack of skills	402	3.50	9
Unpaid employees	390	3.39	10

Note. W.S.: Weighted score
Source: Authors' work

In light of the findings, inter-role conflict (mean=3.78±0.88), family pressure (mean=3.71±0.81) and stereotypic influence (mean=3.68±0.96) were the leading factors responsible for respondents' decision to adopt some other sources of income (Table 4). In agreement with these findings, studies have reported that stereotypic instances usually influenced decisions to engage in livelihood diversification (Chimgonda-Nkhoma et al., 2019; Villanueva-Moya & Francisca, 2021). These findings are equally in line with those of Kosec et al. (2021), who found that neighbors and family members greatly influence women's decision-making ability for livelihood diversification.

Further analysis to assess the influence of educational status on the decision of the women to undertake multiple sources of income indicated that inter-role conflict (χ^2=20.475; p<0.05), stereotypic influence (χ^2=23.544; p<0.01), less appreciation (χ^2=22.077; p<0.01), lack of awareness (χ^2=21.631; p<0.05), male dominance (χ^2=22.837; p<0.05), and illiteracy (χ^2=23.544; p<0.01) were significantly influenced by respondents' educational status (See Table 4); thereby suggesting that women's decision to engage in livelihood diversification is a factor of their level of education. It meant that women with inadequate levels of education might decide not to diversify but rather continue to work in the textile industry even when they experience inter-role conflict, stereotypic influence, less appreciation, lack of awareness, and male dominance in the textile industry. Studies have shown that male dominance, less appreciation of the female folk, and illiteracy have become reasons why males continue to dominate over women in developing societies (Carlana, 2019; Kosec et al., 2021; Larsson & Alvinius, 2019). Therefore, women might not consider these factors necessary to take additional jobs if they do not have the required educational qualifications to diversify their livelihoods or search for alternative sources of income.

Table 4

Factors affecting rural women's decision to adopt livelihood diversification in the textile industry and the influence of educational status in Faisalabad, Pakistan

Factors					Educational status	
	W. S	Mean	S. D	Rank	χ^2	p-value
Inter-role conflict	435.00	3.78	.88	1	20.475	0.015*
Family pressure	427.00	3.71	.81	2	19.451	0.78[NS]
Stereotypes	424.00	3.68	.96	3	23.544	0.005**
Less appreciation	418.00	3.63	.97	4	22.077	0.009**
Lack of awareness	410.00	3.56	.96	5	21.631	0.042*
Male dominance	402.00	3.49	1.02	6	22.837	0.029*
Illiteracy	401.00	3.48	1.04	7	23.544	0.005**

Note. W.S.: Weighted Score; S.D.: Standard Deviation; χ^2-Chi-Square; NS = not-significant, * = Significant (at 5% level), ** = Significant (at 1% level)
Source: Authors' work

Munaza Zafar, Rana Muhammad Amir, Hafiz Ali Raza, Sidiqat Aderinoye-Abdulwahab, Inam Irshad and Tawalkalitu Abiola Dolapo

However, family pressure (χ^2= 19.451; p>0.05), a leading factor for the decision to take more jobs, indicated no significant association with the level of education. It implied that educational qualifications are never considered for women's decision to search for and be engaged in more jobs if family pressure existed. It is an indication that literate and illiterate women working in textile industries undertook more jobs to diversify their livelihood when family pressure existed, though mostly for financial support. It further implied that family pressure is a strong push factor for livelihood diversification among women working in the textile industry in the study.

Hypothesis Testing: Relationship Between Socio-economic Factors and the Rural Women's Decision to Diversify Livelihood in Textile Industry in Faisalabad, Pakistan

The Spearman correlation analysis (Table 5) showed the socio-economic factors affecting the decision to diversify livelihood strategies among women working in textile industries. Spearman's correlation

coefficient (r = -.225; p = .016) showed a significant but inverse relationship between age and the decision-making ability of rural women in the textile industry (Table 5). Age negatively but significantly influences women's decision-making abilities in the textile industry. It could be that the older women were still not able to make informed decisions regarding household incomes and the textile industry despite having garnered more experience that could help them navigate through the industry with fewer issues. In contrast, younger women with higher education qualifications might have adjusted to the intrigues of the industry and, hence, were better able to make more informed decisions. It also implied that younger women tend to undertake additional jobs/occupations in the textile industry more than older women.

The Spearman correlation coefficient of education (r = .342; p = .000) showed a significant positive relationship. The implication is that women with higher levels of education were better able to make informed decisions that resulted in paid jobs in the textile industry compared

Table 5
Relationship between socio-economic factors and the rural women's decision-making ability in the textile industry in Faisalabad, Pakistan

Socio-economic characteristic	Pearson correlation coefficients	p-value
Age of respondents	-.225	.016[*]
Education of respondent	-.342	.000[**]
Marital status of respondents	-.228	.014[*]
Household size of respondents	.046	.626[NS]
The landholding of respondent	.032	.736[NS]
Women's total income	-.308	.001[**]

Dependent variable = Constraints
Source: Authors' work

to women with lower levels of education (the illiterates) who might likely remain in farming occupations where little or no literacy is needed. It was similarly reported that households with low education levels were affected by poverty more than those with higher levels of education in Nepal (Acharya et al., 2022).

Furthermore, a significant but negative relationship existed between marital status and the decision to diversify livelihood strategies in the textile industry (r = -.228; p = .014) (Table 5). Thus, unmarried women had better decision-making abilities than their married counterparts in the textile industry. Consequently, unmarried women in the textile industries had better chances to adopt livelihood diversification strategies than those who were married. Again, this result may not be out of place as married women could rely on their spouses and other family members to make decisions. In contrast, unmarried women might be more independent in terms of decision-making. The correlation coefficient of income (r = .308; p = .001) also showed that the women who had higher incomes had a higher probability of diversifying their livelihoods than those with lower incomes. Lower income translates into less diversification of livelihood in the textile industry (Table 5).

On the decision to diversify livelihood, there was no significant relationship between household size (r = .046; p = .626) and size of land holdings (r = .032; p = .736) of the women working in the textile industry. The interpretation here is that the size of land holdings and number of people in the women's households did not play any significant role in the way decisions were made and the kind of livelihood diversification made by the women. Hence, women were able to make decisions in whatever direction they chose, regardless of their household and land sizes. Overall results showed that age, education, marital status, and income had significant relationships with the decision-making abilities of women working in the textile industry to diversify their livelihood. This finding is corroborated by Omang et al. (2020), who found that human assets, education, and skill transform the capability of rural households to diversify their incomes. These findings are equally in line with Nusrat (2021), who reported that the prospects for women to adopt livelihood diversification vary considerably due to social, cultural and economic norms of the society. Moreover, acknowledging the presence and associations of rural leadership will greatly influence and substantially impact decision-making ability that will enhance the planned change toward community development (Hou & Wu, 2020).

CONCLUSION

The study concluded that non-farm employment is critical in boosting the revenue of rural women households, with nearly three-quarters of the women combining daily wages, poultry farming, and livestock production. At the same time, it is also a viable method for supplementing the living standards of rural

women in Faisalabad, Pakistan. Factors responsible for women's decision-making in livelihood diversification were illiteracy, inter-role conflict, stereotypic influence, less appreciation, and lack of awareness. Rural women's farm families, therefore, diversified into the textile industry as one of the daily occupations that could be used to complement their livelihoods. Despite diversifying into the textile industry, rural women are still faced with challenges such as gender-based discrimination, lack of education, non-cooperation of colleagues, low wages, lack of training, lack of communication skills, transport problems and lack of self-confidence in the textile industry. It, therefore, means there is still a dire need to boost farm women's capacity by strengthening them in the identified inadequacies. It was also established that certain socio-economic factors such as age, education, land holdings, household sizes and household incomes had a bearing on the decision-making ability of rural women in the textile industry. Consequently, women with higher levels of education, larger land holdings, and higher incomes might have firmer capacities to improve their livelihoods.

Implication and Recommendation for the Study

Given the findings of this study, it was shown that women working in the textile industry diversified their livelihood into both farming and off-farm income sources. Thus, relevant governmental agencies should provide an enabling environment for women and men to equally participate in workplaces of their choice, and this should be made possible by enacting appropriate legislative policies geared toward ending gender-based discrimination and low wages among women. It was also established that women's lack of formal education posed a challenge to livelihood diversification decisions. The Pakistani government should, therefore, mount education programs targeted to favor adults so rural women could benefit. More importantly, family pressure is a strong push factor for livelihood diversification among women in the textile industry. Consequently, rural agro-based small-scale industries should be encouraged through policy reforms to diversify agricultural resources. Moreover, rural agro-based industries would also provide jobs for rural women so they could financially support their families for enhanced livelihoods.

ACKNOWLEDGEMENT

REFERENCES

Acharya, K. P., Khanal, S. P., & Chhetry, D. (2022). Factors affecting poverty in Nepal - A

binary logistic regression model study. *Pertanika Journal Social Sciences & Humanities, 30*(2), 641 – 663. https://doi.org/10.47836/pjssh.30.2.12

Aderinoye-Abdulwahab, S. A., Nwachukwu, S. C., Salawu, O. L., & Popoola, P. O. (2015).

Assessment of livelihood activities of rural farmers in Kwara State, Nigeria. *Ethiopian Journal of Environmental Studies and Management, 8*(2), 120-129. https://doi.org/10.4314/ejesm.v8i2.2

Asravor, R. K. (2018). Livelihood diversification strategies to climate change among smallholder farmers in Northern Ghana. *Journal of International Development, 30*(8), 1318-1338. https://doi.org/10.1002/jid.3330

Baghal, M. G. U., Tunio, S., & Sheikh, M. (2018). An analytical study of the critical situation faced by women at working place and legislation policies in Pakistan perspective. *The Government-Annual Research Journal of Political Science, 7,* 170-177.

Bayeh, E. (2017). The role of empowering women and achieving gender equality to the sustainable development of Ethiopia. *Pacific Science Review B: Humanities and Social Sciences, 2*(1), 37-42. https://doi.org/10.1016/j.psrb.2016.09.013

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *Quarterly Journal of Economics, 134*(3), 1163–1224. https://doi.org/10.1093/qje/qjz008

Chambers, R., & Conway, G. (1992). *Sustainable rural livelihood practical concepts for the 21st century.* https://www.ids.ac.uk/publications/sustainable-rural-livelihoods-practical-concepts-for-the-21st-century/

Chimgonda-Nkhoma, J.J., Garforth, C., Cardey, S. Aderinoye-Abdulwahab, S. A., & Fawolde, B. (2019). Perception of concept and practice of social power in development interventions in Malawi. *Journal of Agricultural Extension, 23*(3), 147–155. https://doi.org/10.4314/jae.v23i3.13

De Giorgi, E. G., & Mahmoud, O. (2016). Diversification preferences in the theory of choice. *Decisions in Economics and Finance.* 35. http://doi.org/10.2139/ssrn.2627970

Dinku, A. M. (2018). Determinants of livelihood diversification strategies in Borena pastoralist communities of Oromia regional state, Ethiopia. *Agriculture & Food Security, 7*(1), 1-8. https://doi.org/10.1186/s40066-018-0192-2

Etuk, E. A., Udoe, P. O., & Okon, I. I. (2018). Determinants of livelihood diversification among farm households in Akamkpa local government area, cross river state, Nigeria. *Agrosearch, 18*(2), 101-112. https://doi.org/10.4314/agrosh.v18i2.8

Gautam, Y., & Andersen, P. (2016). Rural livelihood diversification and household well-being: Insights from Humla, Nepal. *Journal of rural studies, 44*, 239-249. https://doi.org/10.1016/j.jrurstud.2016.02.001

Habib, N., Alauddin, M., Cramb, R., & Rankin, P. (2022). A differential analysis for men and women's determinants of livelihood diversification in rural rain-fed region of Pakistan: An ordered logit model (OLOGIT) approach. *Social Sciences & Humanities Open, 5*(1) Article 100257. https://doi.org/10.1016/j.ssaho.2022.100257

Hafeez, A., & Ahmed, H., (2002). Factors determining the labor force participation decision of educated married women in a district of Punjab. *Pakistan Economic Social Review, 40* (1), 75–88.

Haq, M. A. U., Victor, S., & Akram, F. (2021). Exploring the motives and success factors behind female entrepreneurs in India. *Quality & Quantity, 55*(3), 1105–1132. https://doi.org/10.1007/s11135-020-01046-x

Haq, M. A. U., Jali, M. F. M., & Islam, G. (2017). Decision-making ability as a source of empowerment among rural women of Pakistan. *Global Social Welfare, 4,* 117–125. https://doi.org/10.1007/s40609-017-0091-7

Hou, C., & Wu, H. (2020). Rescuer, decision maker, and breadwinner: Women's predominant leadership across the post-Wenchuan earthquake

efforts in rural areas, Sichuan, China. *Safety Science, 125,* Article 104623. https://doi. org/10.1016/j.ssci.2020.104623

Iqbal, M., Ping, Q., Abid, M., Ahmed, U. I., Nazir, A., & Rehman, A. (2016). Adoption of off-farm diversification income sources in managing agricultural risks among cotton farmers in Punjab Pakistan. *Journal of Applied Environmental and Biological Sciences, 6*(8), 47-53.

Islam, M. S., Islam, S., Kanij, F., Khanum, R. (2022). Rural women participation in farm and off-farm activities and household income in Bangladesh. *Heliyon, 8*(9), Article e101618. https://doi. org/10.1016/j.heliyon.2022.e10618

Javaid, Q. U. A, Khan, Z. U., & Siddiq, U. (2020). Societal challenges faced by working women in Pakistan. *PalArch's Journal of Archaeology of Egypt/Egyptology, 17*(7), 10275-10288.

Kassa, W. A. (2019). Determinants and challenges of rural livelihood diversification in Ethiopia: Qualitative review. *Journal of Agricultural Extension and Rural Development, 11*(2), 17-24. https://doi.org/10.5897/JAERD2018.0979

Khan, M., Muhmood, K., Noureen, S., Mahmood, H. Z., & Amir-ud-Din, R. (2020). Epidemiology of respiratory diseases and associated factors among female textile workers in Pakistan. *International Journal of Occupational Safety and Ergonomics*, *28*(1), 184-198. https://doi.or g/10.1080/10803548.2020.1751973

Kongla, N. L., & Akoh, N. R. (2019). Demographic factors influencing rural women's diversification of rural activities for sustaining rural households' livelihoods in Cameroon-Case study of Bui Division. *Saudi Journal of Humanities and Social Sciences, 4*(1), 1-9. https://doi.org/10.21276/ sjhss.2019.4.1.1

Kosec, K. Mo, C. H., Schmidt, E., & Song J. (2021). Perceptions of relative deprivation and women's empowerment. *World Development,*

138, Article 105218. https://doi.org/10.1016/j. worlddev.2020.105218

Larsson, G., & Alvinius, A. (2019). Comparison within gender and between female and male leaders in female-dominated, male-dominated and mixed-gender work environments. *Journal of Gender Studies, 29*(7). 738-750. https://doi.or g/10.1080/09589236.2019.1638233

Mahboob, A. M., & Anita, M. (2016). A poor country clothing the rich countries: Case of garment trade in Bangladesh. *Экономика региона, 12*(4), 1178-1193.

Naudiyal, N., Arunachalam, K., & Kumar, U. (2019). The future of mountain agriculture amidst continual farm-exit, livelihood diversification and outmigration in the Central Himalayan villages. *Journal of Mountain Science, 16*, 755-768. https://doi.org/10.1007/s11629-018-5160-6

Nusrat, H. (2021). Climate change, livelihoods and gender dynamics of mountainous communities in Pakistan. *Sarhad Journal of Agriculture, 37,* 1269–1279. https://doi.org/10.17582/journal. sja/2021/37.4.1269.1279

Oludipe, B. (2019). Socio-economic effects of livelihood diversification on rural women farmers in Odeda local government area of Ogun State, Nigeria. *International Journal of Economic Development Research and Investment, 10*(3), 11-21.

Omang, T. A., Okpa, J. T., Okoi, O. N., & Iniama, J. (2020). Women barriers and empowerment opportunities in the nigerian context. *Pertanika Journal Social Sciences & Humanities, 28*(4), 3167 – 3186. https://doi.org/10.47836/ pjssh.28.4.37

Pakistan Economic Survey. (2015). *Government of Pakistan survey report (2015)*. https://www. finance.gov.pk/survey_1516.html

Sample Size Calculator (n.d.). Calculator.net. https:// www.calculator.net/sample-size-calculator.html

Sayginer, C., & Kurtsan, K. (2022). An extended decision-making model of coastal recreational area use during the COVID-19 through goal-directed behavior and perceived benefits framework. *Pertanika Journal Social Sciences & Humanities, 30*(2), 541– 556. https://doi.org/10.47836/pjssh.30.2.07

Shah, A. A., Gong, Z., Khan, N. A., Khan, I., Ali, M., & Naqvi, S. A. A. (2021). Livelihood diversification in managing catastrophic risks: Evidence from flood-disaster regions of Khyber Pakhtunkhwa Province of Pakistan. *Environmental Science and Pollution Research, 28*(30), 40844-40857. https://doi.org/10.1007/s11356-021-13598-y

Singh, I., Squire, L., & Strauss, J. (1986). A survey of agricultural households, models receipt findings and policy implications. *The World Bank Economics, 1*(1), 149-179.

Villanueva-Moya, L., & Francisca E. F. (2021). Gender differences in decision-making: The effects of gender stereotype threat moderated by sensitivity to punishment and fear of negative evaluation. *Behavioral Decision Making, 34*(5), 706-717. https://doi.org/10.1002/bdm.2239

APPENDIX

Questionnaire (Interview Schedule)

Title: An Analysis of Livelihood Diversification of Women of Farm Families working in the Textile Industries in District Faisalabad, Pakistan

Part A:

To investigate the socio-economic characteristics of respondents

General survey data

Date of interview : _____

Name of interviewer : _____

Socio-economic characteristics

1. Age (years): ☐ 19–23   ☐ 24–28   ☐ 29–33   ☐ 34–38   ☐ >38
2. Education level: ☐ Illiterate   ☐ Primary (1–5)   ☐ Middle (6–8)   ☐ High school (9-12) ☐ Graduate and above
3. Marital status: ☐ Married   ☐ Divorce   ☐ Widow   ☐ Un-married
4. Household size: ☐ 3–4   ☐ 5–6   ☐ 7–8   ☐ ≥9
5. Who is the head of the household?   ☐ Father   ☐ Husband   ☐ Respondent
6. Type of family: ☐ Nuclear   ☐ Joint
7. Predominant occupation
8. Daily wages
9. Farming
10. Small business
11. Government job
12. Others
13. Do you have landholding?   ☐ Yes   ☐ No. If 'YES,' then how much? _____
14. Women's total income
    ☐ 5000–8000 (PKR)
    ☐ 8001–10000 (PKR)
    ☐ 10001–14000(PKR)
    ☐ Above 14000 (PKR)
15. Doing overtime
    ☐ Yes
    ☐ No

Part B:

To identify the available sources of income adopted by rural farm families.

Livelihood diversification strategies of rural women in the textile industry in Faisalabad

S/N	Source of income	Yes	No
1	Source of income		
2	Farming		
3	Agricultural labor		
4	Casual labor		
5	Government employment		
6	Daily base wages		
7	Business		
8	Remittance		
9	Cottage industries		
10	Livestock		
11	Poultry		

Part C:

To identify the access of farm women towards decision-making of choosing their alternative source of income

1. Factors that affect women's decision-making ability

S/N	Factors that affect women's decision-making ability	1	2	3	4	5
1	Illiteracy					
2	Family pressure					
3	Inter-role conflict					
4	Stereotype					
5	Male dominance					
6	Lack of awareness					
7	Less appreciation					

2. Perception of women towards decision-making of choosing their alternative source of income

S/N	Statements	Yes	No
1	Do you think that jobs raise the social status of women?		
2	Do you think that the families encourage their daughters to seek jobs?		
3	Do you think that jobs have an impact on the empowerment of women?		
4	Do you think that the job provides financial independence to women?		
5	Do you think the job enables women to play a leading role in the family?		
6	Do you think that the job develops confidence in the daily life of women?		
7	Do you think the job enables women to make personal life decisions freely?		
8	Do you think that the social status of men and women is equal in their workplace?		
9	Do you think that job has an impact on the marriage of women?		
10	Do you think that job opportunities for women are equal to men?		

Scale= 1=Very low, 2=Low, 3=Medium, 4=High, 5=Very high

Part D:

To explore the factors/challenges faced by farm families working in textile industries.

S/N	Factors	1	2	3	4	5
1	Lack of skills					
2	Lack of communication skills					
3	Lack of education					
4	Lack of self-confidence					
5	Lack of training					
6	Unpaid employees					
7	Low wages					
8	Gender-based discrimination					
9	Colleagues are non-cooperative					
10	Transport					

Scale: 1= Very low, 2=Low, 3=Medium, 4= High, 5=Very high

Part E:

To provide research-based suggestions to rural working women

How can rural women play a role in improving their family livelihood?

_____
_____
_____
_____

Do you think that industries give sufficient opportunities for employment, which helps rural women and their family life to uplift their lives?

_____
_____
_____
_____

Review Article

Exploring the Roots and Solutions of Maladministration, Power Abuse, or Corruption in Contemporary Indonesian Villages

Rifai Rifai*, Kamaluddin Kamaluddin and Rahmad Hidayat

Department of Administrative Science, Universitas Mbojo Bima, 84119 Bima, Indonesia

ABSTRACT

This article summarises previous studies on corruption in the rural sector after Indonesia's Village Law was passed. It uses a non-systematic multidisciplinary survey of forty peer-reviewed social science literature from 2014–2023 in the Scopus and Google Scholar databases to explore various complex factors that lead to village fund corruption and the set of conceptually proposed intervention measures to fight it. Public officials have frequently misused the granted autonomy, as proved by the rampant corruption in recent years. Beyond direct financial motives and low salaries, inadequate information provision and institutional weaknesses emerge as additional contributors to corruption. Various tactics, including manipulation of budgets, creation of fictitious reports, hindrance of public facility construction, misappropriation through unproductive enterprises, establishment of "ghost" villages, and solicitation of bribes, are employed by corruptors. Previous authors suggested that effective village governance can prevent and address such a worst maladministration by implementing robust practices like internal control, accountability, transparency, and compliance with regulations. Other proposed solutions to combat village fund corruption are empowering village officials, using advanced technologies, intensifying external audits, getting the community more involved, and enforcing strict legal sanctions. This overview urges evidence-based policies, collaboration, and systemic efforts for inclusive development and equitable outcomes in Indonesia's rural areas. While macro-level studies and specific effective formulas still need to be improved, we call for future research to address this

E-mail addresses:
rifaiahmad69unmbo@gmail.com (Rifai Rifai)
al.kamal6921@gmail.com (Kamaluddin Kamaluddin)
rahidsmart@gmail.com (Rahmad Hidayat)
* Corresponding author

gap by conducting a rigorous exploration of village corruption and producing the best ways to stop it.

INTRODUCTION

Maladministration, corruption, and abuse of power are concepts with intertwined meanings. Maladministration works as a mode of procedural deviation or intentional abuse of power by elected leaders and any level of government apparatus to expand influence and increase personal wealth by neglecting public interests. Maladministration reflects a condition in which decisions taken by government officials deviate from the public interest so that they fall into the category of corruption (Romano et al., 2021) because it always intersects with a series of actions ranging from misdemeanours to professional misconduct, negligence, moral decay, legal violations, and criminal activity (Milley & Dulude, 2021); as a result of the administrator's ignorance of the regulatory regime and ethical dimensions related to corruption problems (Onyango, 2022). Maladministration becomes a device of acts of corruption that reflects the application of arbitrary power in it, and corruption itself can be placed in the category of arbitrary power because it always presupposes the use of force to achieve particular intentions that are contradictory to the original ideal goals (Breakey, 2017).

Abuse of power is the main distinguishing feature of corruption from other maladministration behaviours because corruption is often based on a person's positional factors. Once the individual holds a position of power, corruption can arise when the individual intends to do corrupt work (Modesto & Pilati, 2020). In other words, anyone in power in a specific public position may be tempted to utilise that power for a different goal than that dictated by laws and standards (Pozsgai-Alvarez, 2020). Thus, acts of maladministration, corruption, and especially abuse of power originate from a discrepancy between the ideal goals implied by regulations for administrative authorities and the personal goals pursued deliberately by actors working in government institutions (Hayati, 2019).

The optimism of many parties towards the potential contribution of post-decentralisation village regulation in Indonesia and the large disbursement of village funds for the realisation of the independence of this lower-level government area is increasingly being eroded by the emergence of the many cases of worst maladministration, abuse of power, or village fund corruption (Harun et al., 2021; Novrizal & Podger, 2014; Permana et al., 2020; Sofyani et al., 2020, 2022, 2023; Srirejeki & Putri, 2023; Yuliastina et al., 2021). Since village autonomy was rolled out in 2015 as a manifestation of putting the Village Law in effect, the trend of corruption in the village sector has increased. During 2015-2022, the Central Government disbursed around four hundred seventy

trillion rupiahs of village funds. Still, in their management, it was revealed that 601 corruption cases ensnared 686 village heads from all over the country (Pribadi, 2022). These cases have been reported in various national and regional media, where different court sentences have been handed down to the perpetrators involved. Based on this fact, village fund corruption has affected numerous local administrations throughout Indonesia. The increasing trend of corrupt behaviour reflects the personal needs and interests of the perpetrators, especially the Village Heads, to accumulate wealth and pay off political debts while they are still in office, which is made possible by the opportunity to take such actions (Lucas, 2016).

Corruption is not limited to undermining the quality of transparency and accountability of the village government and other executive, legislative, and judiciary institutions. That is why Indonesia suffers from endemic and entrenched corruption (Tambunan, 2023). The rampant corruption committed by government officials in this country has further strengthened the seeds of public distrust, which is not only caused by the government's low formal legitimacy in accommodating public demands (Hidayat, 2019). In other words, corruption has destroyed the public's belief that government institutions will always act for the common good, not the interests of individuals or groups.

It is undeniable that inefficiency in the economy, acute socioeconomic injustice, and the weakening of democratic functions in developed and developing countries are also negative consequences triggered by acts of corruption (Fisman & Golden, 2017). Because there is no efficient system of accountability, government entities abuse their authority, which leads to corruption (Bonsu et al., 2023). It is the most prominent type of deviation by local governments (Yates & Graycar, 2020), including in Indonesia. Corrupt behaviour, as the worst maladministration, is inherent in unclear bureaucratic processes. The lack of capacity of bureaucrats to comply with financial procedures (Onyango, 2022), so the primary characteristics of acts of violation of the law always lie in the use of authority and formal power clandestinely under the pretext of legality (Breakey, 2017), to obtain something that is termed a "personal unlawful benefit" (Putri et al., 2024).

Corruption worsens social inequality by disadvantaging vulnerable populations, exacerbating existing disparities in Indonesian villages. By tackling corruption, policymakers can work towards a more equitable distribution of resources and opportunities and reduce social inequality. Corruption obstructs progress on various goals, and understanding the local manifestations is vital to meeting these global targets. Addressing corruption also aligns with Indonesia's commitment to achieve the SDGs (Sustainable Development Goals). Consequently, a targeted approach to combating corruption can contribute significantly to Indonesia's SDG progress. Understanding and addressing corruption and power abuse at the village level is essential for building strong communities and, by

extension, a stronger nation. Transparent and accountable village governance can serve as a model for development (Hidayat et al., 2019; Taufiq & Hidayat, 2022). By enhancing local governance, Indonesia can foster sustainable development and community empowerment, benefiting the country.

An extensive collection of substantive but scattered literature discussing the issue of village corruption in Indonesia, especially post-decentralisation (during the period 2014–2023), has so far increased, in which the authors used forty peer-reviewed journal articles indexed in the Scopus and Google Scholar databases as a basis for analysis, out of sixty-nine literature listed in the bibliography. Nevertheless, other writings outside the theme and period of the search are also used to present the formulation of the meaning of the concepts of maladministration, corruption, and abuse of power. Understanding the complex interplay of factors that contribute to corruption and power abuse in Indonesian villages is crucial for both policymakers and concerned citizens.

A complete literature study needs to be conducted on the reasons for, how, and other complicated factors of village fund corruption in Indonesia, especially after the 2014 decentralisation. The necessity for a comprehensive synthesis that brings all the findings is still a research gap because most research is fragmented and only looks at certain aspects. That is why we need a full review that profoundly examines the motives, modes, and complicated factors

that led to village corruption after the Village Law granted autonomy. Village corruption prevention efforts have also been studied, but the literature must fully document these measures. This collection of past research can show village fund corruption, fill in the blanks where we do not know enough, and help us devise better ways to stop it.

The core idea to be conveyed about the above narrative is that the decentralisation of village governance in Indonesia, especially after the issuance of the Village Law, has led to a significant increase in corruption cases as the worst maladministration practice that reflects the abuse of power of the village government officials, which harms socioeconomic justice, weakening democratic functions, and making unbalanced distribution of development resources (Antlöv et al., 2016; Ash-Shidiqqi & Wibisono, 2018). Large allocations of village funds are vulnerable to corruption by public officials due to the absence of efficient accountability systems. Abuse of power is the main characteristic of corruption compared to other maladministration behaviours, which often arise due to a person's ability or political position where they are tempted to use it to achieve goals contrary to legal provisions.

This literature review presents a narrative of the close relationship between corruption as the worst maladministration practice and the abuse of power by the village government officials to expand influence and increase personal wealth without regard to the public interest. It delves into the heart of this issue, shedding light on the root causes,

consequences, and efforts to combat these challenges in the contemporary Indonesian context. By examining specific cases and the broader socio-political landscape based on previous research findings, we aim to provide a comprehensive perspective on a matter that has significant implications for the well-being and development of local communities in Indonesia.

The research aims to address two main questions: (i) what are the motives and modes employed by individuals involved in village fund corruption in Indonesia following the implementation of the Village Law and (ii) what interventions have previous researchers proposed to combat it. This literature review seeks to uncover the motivations and methods employed by individuals involved in village budget corruption in Indonesia following the implementation of the Village Law. Another objective of this study is to outline the strategies for combating corruption in village budgets, drawing on prior research undertaken by scholars. This research aims to enhance comprehension of the underlying causes of village fund corruption and inform the formulation of productive ways to combat it in the future.

This article delves into the motives, modes, and consequences of the village fund corruption and the prior proposed intervention measures. Understanding the dynamics of corruption is crucial for policymakers, government officials, and stakeholders involved in rural development. Corruption diverts resources from essential projects and undermines public trust

in government institutions, hindering socioeconomic progress and perpetuating poverty. Our study sheds light on the factors contributing to corruption and the tactics applied by corruptors, providing valuable insights for designing targeted interventions. Addressing corruption at the village level aligns with Indonesia's commitment to achieving the Sustainable Development Goals and can contribute to broader efforts to foster inclusive development and reduce inequality. By synthesising findings from previous studies and identifying gaps in knowledge, this article lays the groundwork for future investigations into village fund corruption and the most effective strategies for combating it.

The Historical Foundations of Village and the Ongoing Problem of Corruption

The issue of corruption in Indonesian villages has a longstanding historical foundation, dating back to pre-colonial and colonial periods. The historical settings have had a crucial role in the emergence and perpetuation of corruption within contemporary village governance. During the pre-colonial period, Indonesia comprised numerous kingdoms and sultanates, each distinguished by its unique system of government and taxation. Despite the presence of diversity and decentralisation throughout the kingdoms, incidences of corruption were common. Local administrations often charge excessive taxes, misuse funds, and participate in nepotism. The persistence of corrupt

practices remained prevalent during the colonial era, as evidenced by the Dutch's development of a centralised administrative framework, which consolidated power within a selected group of colonial officials. Corruption and bribery were prevalent in the Dutch government, including at the grassroots level of government, such as within villages (Hannigan, 2015; Kroeze, 2021).

Following its independence in 1945, Indonesia faced the enormous task of establishing a unified nation. The enduring presence of corruption within village governance can be ascribed to a convergence of factors, including resource limitations, insufficient institutional capacities, and a scarcity of effective oversight mechanisms. Village officials' discretionary authority often influenced resource allocation, enabling occurrences of embezzlement and misuse of monies. Moreover, a cultural framework centred around patronage and clientelism endured, hence fostering corrupt practices in which those in positions of authority granted favours and allocated resources to their loyal followers. The period of Suharto's governance, spanning from 1967 to 1998, was marked by a notable prevalence of corruption, signifying a solemn episode in the historical account of Indonesia. The consolidation of power under the leadership of Suharto's authoritarian administration had a significant impact, leading to the widespread prevalence of corruption at many levels of government, including among local communities. Government officials are frequently involved in unlawful activities, including embezzlement, extortion, and bribery, fostering a pervasive corruption culture (Vickers, 2013).

Despite the political demise of Suharto, corruption persisted in rural areas of Indonesia. The subsequent decentralisation endeavours to enhance the authority of local governments presented novel obstacles. The village authorities, who possess increased autonomy, must prepare to manage their augmented duties effectively and consistently. The presence of inadequate supervision measures and inconsistent enforcement of accountability mechanisms resulted in the emergence of chances for corrupt practices. Corruption continues to be a pressing issue in rural areas of Indonesia, characterised by financial misappropriation, bribery, and inadequate administration of public resources. It is imperative to recognise that endeavours to address corruption have been continuously underway, accompanied by increasing awareness of its adverse effects on development and government. Understanding this historical context and the prevalence of corruption is essential for crafting effective strategies to address this deeply entrenched issue in Indonesian villages (Fürstenberg et al., 2023; Mukartono & Rustamaji, 2019).

The adoption of the Village Law marks the stretching of fiscal and political decentralisation for villages as regions and lower-level government entities in Indonesia. This regulation reflects the efforts of village autonomy to bridge the accelerated increase in public welfare in the various areas of the country. In line with

this framework, village decentralisation can also trigger initiatives to increase village income (Lewis, 2015). Through the decentralisation scheme, the village has been acknowledged as a distinct legitimate entity with unique rights and powers, especially in the domain of development management as well as determining its budget expenditure (Vel et al., 2017) for community empowerment and poverty alleviation (Putri et al., 2024). The Village Law gives greater autonomy to village governments to control and oversee local activities under community interests while promoting residents' active involvement in identifying issues and opportunities and producing solutions (Sofyani et al., 2022). This regulation provides adequate authority and disbursement of funds to village governments for administering the government, putting in place development, community empowerment, and community development while requiring each village administration to abide by and put those ideals, like anti-corruption, openness, participation, and accountability, into practice (Susan & Budirahayu, 2017).

Villages in Indonesia occupy a pivotal position in the country's administrative hierarchy, and their importance is multifaceted. They serve as the foundation of governance, facilitating a direct connection between the government and the people. This proximity enables localised and responsive management, catering to the unique needs of individual communities. Moreover, it is imperative to acknowledge the significant contribution of villages in

facilitating vital public services, such as education, healthcare, and infrastructure development, especially in geographically isolated regions. Beyond administrative functions, they function as custodians of Indonesia's diverse cultural heritage, preserving and promoting Indigenous traditions.

Additionally, they empower local communities by encouraging civic engagement and offering a degree of autonomy in decision-making through elected village leaders and councils. They are central to rural development efforts, fostering progress in rural areas, and their bottom-up planning approach ensures alignment with community needs. Implementing decentralisation strategies has given villages a certain degree of local autonomy, allowing them to govern their affairs while operating within the boundaries of national legislation. Recognition and enhancement of their role are imperative for Indonesia's comprehensive development and citizens' welfare (Annahar et al., 2023).

The realisation of the ideal desire for village autonomy is currently prone to be tripped up by an overflow of village fund corruption cases, which has sparked high pessimism among various actors. Instead, the government's concentration was drained on the problem of rampant corruption in certain villages (Nugraha et al., 2019), conducted by the village head and his apparatus (Srirejeki & Putri, 2023). Non-accountable officials misused the allocation of development funds, so this large budget contributed little to advancing village life

(Sofyani et al., 2020). These sad facts gave rise to the pessimism of many parties that the village autonomy policy only provides new space for inefficiency and corruption of the state finance at the village level instead of bringing significant benefits to increasing the group of people's welfare.

The trend of increasing corrupt behaviour in village funds management is contrary to the formulation of the lofty ideals of decentralisation or granting autonomy through the passing of the Village Law in which the Village Government is given great discretion in managing local affairs based on the public interest (Sofyani et al., 2022), increasing the performance of providing services and reducing social inequality (Lewis, 2015), and maintaining a democratic control room that allows active participation of citizens in overseeing village development budget spending (Vel et al., 2017). Currently, the Village Government has complete control over the management of government affairs and has the right to receive direct allocations of village funds from the Central Government (Pratolo et al., 2022). Village funds are an additional substantive fiscal resource for villages to accelerate development programs and improve public welfare by intensifying poverty alleviation programs (Harun et al., 2021).

In summary, corruption in Indonesian villages has long historical roots, starting from the pre-colonial, colonial, and post-independence periods, where limited resources, weak institutional capacity, a culture of patronage, and the lack of effective monitoring mechanisms reinforce the roots of corruption in village governance. Decentralisation has given autonomy to villages, but unfortunately, corruption remains a serious problem that reduces the benefits of this politico-administrative discretion to accelerate the realisation of the welfare of village communities.

METHODS

This article is a narrative overview, a particular type of narrative literature review that presents a summary or narrative synthesis of a concise version of previously relevant information published through simple steps in the form of (a) performing a preliminary search of the literature, (b) selecting some of the best evidence synthesis, and (c) presenting a narrative review as objective as possible (Ferrari, 2015). The primary purpose of such a literature review is only to present a summary of prior knowledge where the scope of questions is broad; there is a selective search strategy and primary materials' conceptual-empirical nature; and a need for a straightforward study selection process and quality assurance. It synthesises or analyses findings using a narrative summary (Paré et al., 2015). Our article aims to provide a concise but comprehensive overview of the available literature on motives, modes, and other relevant aspects surrounding all cases of the worst maladministration, abuse of power, or corruption in the rural sector, as has occurred in Indonesia. We compiled a narrative overview adopting the steps of previous works such as Monteduro et al.

(2016), Kok et al. (2018), Clemente-Suárez et al. (2022), Leh et al. (2023), and Yeo et al. (2023), which cover corruption and non-corruption issues, respectively.

This article undertakes a non-systematic multidisciplinary survey of forty peer-reviewed social science literature spanning the period from 2014 to October 2023 in the Scopus and Google Scholar databases to comprehensively summarise previous studies on maladministration, corruption, and power abuse in the rural sector following the enactment of Indonesia's Village Law. We prefer the Scopus database for collecting relevant literature to emphasise the characteristics of peer-reviewed journal articles to be reviewed. However, we also found several pertinent documents in the Google Scholar database that are closely related to the issue of village corruption or governance after the issuance of the Village Law. These documents are journal articles, book chapters, and conference papers. However, the first type fills the highest composition in our overview.

The deliberate choice of a non-systematic approach is designed to capture diverse perspectives from various disciplines within the social sciences. Beyond the confines of the search period, the study also incorporates relevant writings to provide a holistic understanding of the three core concepts. This inclusive strategy acknowledges that insights from contemporary and older literature contribute to a nuanced and comprehensive overview of maladministration, corruption, and power abuse in the context of the Village Law's

implementation. By extending the scope beyond the search period, the research methodology aims to capture the evolution and continuity of ideas surrounding these critical issues in Indonesia's rural sector.

We employed search terms such as "village maladministration," "village corruption," "power abuse in the village," "Indonesian villages," and "Village Law" in the Scopus and Google Scholar databases to locate relevant literature. We established inclusion and exclusion criteria to filter through the search results. The inclusion criteria functioned as filters, selecting articles from peer-reviewed social science journals covering the specified topics and published between 2014 and October 2023. On the other hand, we used exclusion criteria to remove sources lacking peer review, articles irrelevant to the research topics, and those published beyond the specified timeframe.

We then examined the keywords and abstracts of a few articles to determine if they passed the first screening. We were further analysed to identify significant information and ideas aligned with the study goals. We applied systematic thematic analysis to organise the collected data logically. That facilitated the categorisation of common themes and trends in the literature. Furthermore, we synthesised the compiled results to create a narrative-like summary of the selected literature topics. At the final step of the review process, we compiled the most significant findings, trends, and insights from the examined literature. This thorough review process has

yielded a comprehensive understanding of the research on village governance issues in Indonesia, enabling further study and informed discussion in this area (Figure 1).

The thematic analysis used for interpreting the data in this study involved identifying, organising, and understanding the main thematic patterns that emerged from the forty selected scientific literature. This thematic analysis approach is a qualitative method of inquiry that enables researchers to explore and understand the essence of the information contained in the texts (Naeem et al., 2023). The process of thematic analysis began with the data collection stage, where scholarly literature relating to

maladministration, corruption, and abuse of power in Indonesia's rural sector, particularly after the enactment of the Village Law, was collected from the Scopus and Google Scholar databases. After data collection, we read thoroughly to understand the context and substance of each piece of literature.

The next step involved labelling or categorising pieces of text reflecting main themes related to maladministration, corruption, and abuse of power. This process allowed us to identify common patterns or differences that emerged in the literature. Furthermore, we organised these themes systematically to portray a holistic and comprehensive picture. This
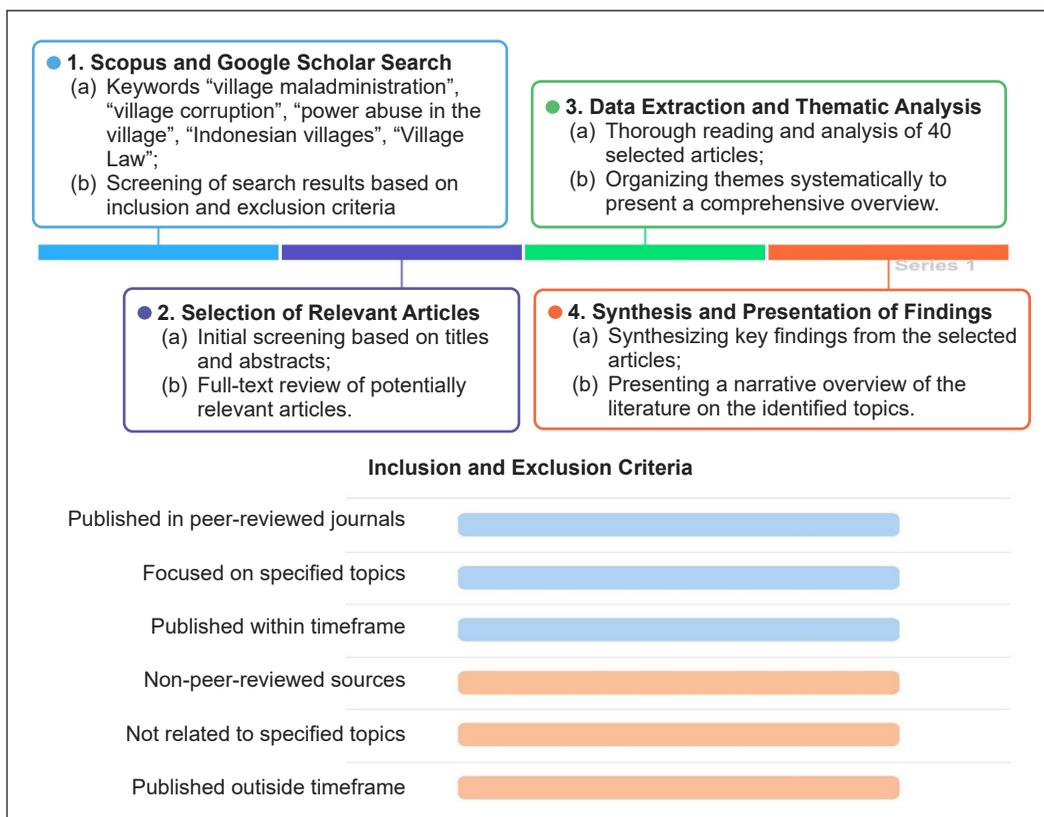
Figure 1. The process of reviewing articles
Source: Authors' work

research did not limit itself to a systematic approach, allowing for the inclusion of diverse perspectives from different disciplines. This thematic analysis also reflects on the evolution of arguments around maladministration, corruption, and abuse of power by incorporating literature outside the data search period. As such, it provides an in-depth understanding of the shifts and continuities of ideas over time. Overall, this thematic analysis is an approach that allows the researchers to explore the deeper meanings of relevant scholarly literature, create space for concept development, and provide a comprehensive picture of the issues studied in the context of Village Law in Indonesia's rural sector.

RESULTS AND DISCUSSION

We explore the complex dynamics surrounding village fund corruption in Indonesia following the enactment of the Village Law to address two fundamental questions: the motives and modes of the perpetrators' corruption and the interventions proposed by previous researchers to combat it. Corrupt individuals employ various methods to evade detection and exploit loopholes in governance systems. These include manipulation of budgets, fabrication of reports, obstruction of public projects, diversion of funds through fictitious enterprises, establishment of "ghost" villages, and solicitation of kickbacks. Perpetrators strategically execute these tactics to circumvent oversight mechanisms, perpetuate a culture of impunity, and exacerbate socioeconomic disparities.

Meanwhile, previous researchers have proposed a range of interventions to address village fund corruption and promote transparency and accountability in governance. By implementing these interventions in a coordinated and holistic manner, stakeholders can work effectively to combat village fund corruption, foster inclusive development, and restore public trust in governance institutions in Indonesian villages.

Motives, Modes, and Other Complex Factors of Village Fund Corruption

Factors causing corruption in the village budget, apart from the aspects mentioned earlier in this article, are rooted in the greed of village government officials (Buchan & Hill, 2014), low salaries of civil servants (Atmadja et al., 2019), and certain aspects in the Fraud Hexagon theory in the form of opportunity, ability, arrogance, and collusion available to the Village Head, accompanied by low community participation in conducting supervision (Sofyani et al., 2023); which is also marked by the lack of provision of information about the administration of village funds (Triyono, 2020). At the institutional level, corruption occurs because of weak internal controls, the absence of comprehensive accounting standards, conflicting village governance regulations, and poor development of information technology (Putri et al., 2024). Poor management or lousy governance in village development, especially in village funds management, is often the main trigger for the emergence of corruption cases.

The above motivating factors for corruption can be divided into different categories, as presented in Table 1.

Meanwhile, Table 2 presents several modes of corruption in Indonesia's village fund system. These modes are usually in the form of budget mark-ups and other misuses of funds utilised to construct physical structures (Harun et al., 2021), making fictitious reports on the use of the budget, and reducing the volume of public facility development projects (Triyono, 2020); buying office inventory for personal use (Purba et al., 2022); embezzling the budget through the formation of non-productive BUMDesa and creating fictitious or ghost villages (Sofyani et al., 2023); as well as

seeking rents (Vel et al., 2017) through bribery schemes as "speed money" in providing public services and issuing concessions. Directly theft of state funds and public property by the village head and other irresponsible actors is the primary mode when they commit corruption in village funds.

A complex interplay of factors influences corruption and power abuse in Indonesian villages. Inadequate oversight procedures, frequently hindered by constraints in capacity and resources, prove ineffective in surveilling village officials' activities, enabling unchecked wrongdoing. As a result, the presence of accountability measures needs to be improved, granting authorities

Table 1

Factors behind village fund corruption in Indonesia

Categories	Indicators	Descriptions
Internal factors	The greed of village government officials	Related to the village government officials' motivation and corrupt behaviour (Buchan & Hill, 2014).
External factors	The low salary of the apparatus	Related to inadequate financial incentives stimulating corruptive behaviour to earn extra income (Atmadja et al., 2019).
	Lack of information provision on village fund administration	Related to the lack of transparency and accessibility of information that can influence corrupt actions (Triyono, 2020).
Institutional factors	Fraud Hexagon Theory	Related to the institutional structures and dynamics that enable corruption, such as opportunity, ability, arrogance, and collusion (Sofyani et al., 2023).
	Weak internal controls, a lack of comprehensive accounting standards, conflicting village government regulations, and poor information technology development	Related to weaknesses in internal control systems, policies, and institutional infrastructure that facilitate corruption (Putri et al., 2024).
	Poor governance in rural development	Related to improvement in management and better governance in the context of village development, particularly in managing village funds (Harun et al., 2021).

Source: Authors' work

Table 2

The modes of village fund corruption in Indonesia

Categories	Descriptions
Budget mark-ups	Corruptors often mark up budgets for personal gain, manipulating them to obtain more funds than they should (Harun et al., 2021).
Fictitious reports	Corruptors create false reports on budget utilisation to cover up their corrupt acts. These reports do not reflect the actual use of the budget (Triyono, 2020).
Reducing the volume of public facility development projects	Corruptors lessen the volume of public facility development projects that should be conducted using village funds. They divert the funds for personal use or misuse them (Triyono, 2020).
Buying office inventory for personal use	Corruptors use village funds to buy office inventory that should be used for the public good, but they use it for personal use (Purba et al., 2022).
Embezzling the budget through non-productive village-owned enterprises	Corruptors create unproductive or fictitious village-owned enterprises (*BUMDesa*) to steal village funds. They use the *BUMDesa* to commit acts of corruption (Sofyani et al., 2023).
Creating fictitious or ghost villages	Corruptors create fictitious villages or "ghost villages" to divert village funds. They make false records of these villages and divert village funds to where they are not supposed to go (Sofyani et al., 2023).
Seeking rents through bribery schemes	Corruptors seek to profit through bribery schemes, known as "fast money," in delivering public services and granting licenses. They demand bribes to provide public services that should be freely available (Vel et al., 2017).

Source: Authors' work

the capacity to partake in corrupt activities without facing consequences. Furthermore, the complex bureaucratic framework in Indonesia, characterised by elaborate administrative procedures and numerous agencies, gives rise to potential corruption in resource distribution and project execution. The presence of cultural and customary factors that emphasise loyalty and patronage may foster nepotism and favouritism.

Consequently, this phenomenon further amplifies the prevalence of corruption. Moreover, the restricted availability of information inside rural communities poses a significant obstacle to the villagers' capacity to recognise and report corruption, hence facilitating the perpetuation of such misconduct. Limited resources in specific communities tempt officials to engage in unlawful activities, such as embezzlement or misuse of cash (Pozsgai-Alvarez, 2020).

Meanwhile, political patronage networks can play a significant role, with officials using their positions to reward supporters, sometimes at the cost of engaging in corrupt practices. Furthermore, limited education and awareness among villagers about their rights and governance practices make them vulnerable to manipulation. Economic pressures on officials, stemming from low salaries, may drive them to demand bribes or engage in corrupt practices. Lastly, resource allocation and service delivery may be manipulated, with officials prioritising specific individuals or communities while marginalising others. Understanding this

intricate web of factors is pivotal for effectively crafting targeted strategies and reforms to combat corruption and power abuse in Indonesian villages (Lucas, 2016; Shaleh et al., 2021).

Village fund corruption, as found in several pieces of literature reviewed by the authors, attaches corruption to actions of individuals (personal crime) and groups (collective corruption) because village heads often conduct budget corruption as a negative consequence of the "too-strong power" they have in which this great power is used solely to form evil and deviant conspiracies (Sofyani et al., 2023). Beyond that, subordinates of the village head also conducted embezzlement of village funds, either singly or through conspiracy with the village head. Moreover, most corruption cases involve the village head and Supra-Village government officials assigned to control plans for development activities and the disbursement of village funds (Harun et al., 2021).

Thus, the dominant analytical model used to explain the issue of village fund corruption in Indonesia is based on the micro-level and meso-level perspectives. The first perspective discusses the intra-individual dimension of corrupt behaviour, including positional (political abuse of power) and dispositional (individual characteristics) factors. The second perspective reviews interactions between individuals to obtain illegal collective benefits (how individuals are linked together in corrupt networks) with a background of intra-individual aspects according to a micro-level perspective. It is scarce to find writings that review the causes

and consequences of village fund corruption based on a macro-level perspective, namely contextual settings in cultural dimensions and political-socio-economic norms of deterioration, such as the influence of specific political systems and cultures.

If translated through the conception of Li and Yang (2023), village fund corruption in Indonesia can be classified as transaction-type corruption in the form of bribery and auto-corruption, including embezzlement and misuse of public finances. Village government is a type of public sector organisation in Indonesia, and village budget corruption that sticks out to the surface is often conducted by political leaders and administrative staff within it. The transaction-type and auto-corruption they commit reflect the actions of public officials who abuse their authority and government employees for personal gain, hindering the realisation of the wider community's interests.

Based on that, it can be asserted that corruption and abuse of power in Indonesian villages are influenced by a variety of complex factors, including the low salaries of the apparatus, the greed, arrogance, and conspiracy surrounding the Village Head, and the low participation of the community in conducting supervision. In addition, weak internal controls, the absence of comprehensive accounting standards, conflicting village government regulations, and poorly developed information technology contribute to corruption at this institutional level. The misappropriation of village budgets is done through various means, such as falsifying budget utilisation

reports, reducing the volume of public facility construction projects, purchasing office inventory for personal use, embezzling budgets through the establishment of unproductive institutions and the creation of fictitious villages, and seeking personal gain through bribery schemes. Corruption and abuse of power in Indonesian villages are also influenced by inadequate oversight procedures, a complex bureaucratic framework, cultural and customary factors that encourage nepotism and favouritism, limited information in rural communities, economic pressures on village officials, and manipulation of resource allocation.

As our narrative overview explores the findings of previous researchers, there are significant differences in their views on village fund corruption in Indonesia. A critical review of the studies reveals interesting interpretative shifts, indicating the complex dynamics behind this issue. One notable difference is the approach to the causal factors of village fund corruption. While some scholars emphasise individual motivations, such as the greed of village officials (Buchan & Hill, 2014), others highlight economic conditions, such as the low salaries of civil servants (Atmadja et al., 2019). A closer look at these differences indicates that the causes of corruption of village funds may be multifactorial, involving a combination of internal and external factors.

In addition, there is variation in the emphasis on institutional aspects. Some studies highlight weaknesses in internal controls and a lack of comprehensive accounting standards as triggers for corruption (Putri et al., 2024), and others accentuate regulatory conflicts within village governance as a significant factor (Triyono, 2020). This comparison suggests that to understand village fund corruption, we must look at the individual level and the institutional and regulatory structures surrounding it. It is primary to note the shift in research focus over time. Some early research may have emphasised individual characteristics and corruptive behaviour, while more recent studies may involve structural and systemic analysis. These changes reflect an increased understanding of the complexity of village fund corruption and a push to address its root causes.

This review also leaves room to consider contextual and regional factors. Can differences in findings be attributed to variability across regions in Indonesia? Are there any differences in socioeconomic conditions and local government regulations that may affect corruption levels? The possible causes of these differences involve the complexity of the individual, institutional, and structural factors. Future research can enrich our understanding by further exploring the interactions between these factors and the changing dynamics of village fund corruption over time.

Addressing Corruption in Indonesian Villages: Strategies and Lessons Learned

A deep understanding of the complex factors mentioned above is essential to designing appropriate strategies and reforms to combat

corruption and abuse of power in Indonesian villages. Efforts are needed to improve oversight, strengthen internal governance, increase transparency and accountability, and involve communities in decision-making and oversight. It is also necessary to increase community education and awareness of their rights, as well as improve economic conditions and equitable resource allocation to reduce financial pressure on the village officials. It is mandatory to strengthen oversight mechanisms, implement clear accounting standards, increase community participation, and develop information technology to improve village governance to prevent corruption, as well as address these complex challenges. Such corrective measures can be implemented as soon as possible to overcome the problems of village funds implementation and achieve better results in rural development and poverty alleviation in post-decentralised Indonesia.

Based on previous research, creating good village governance has become a general recommendation from several authors to prevent and overcome village fund corruption through the creation of strong governance practices by fully implementing the principles of compliance with regulations, internal control, accountability, transparency, and public engagement (Darsono et al., 2021; Permana et al., 2020; Putri et al., 2024; Sofyani et al., 2022, 2023; Sugiharti et al., 2021; Susan & Budirahayu, 2017). That is in line with influential arguments from observers of global corruption issues, such as Ceschel et al. (2022), who assert that effective results in preventing corruption can be achieved through improving control systems over government power holders, limiting monopoly conditions, and increasing the scope of transparency of their actions. In other words, implementing strategies to reduce the extent of corruption through policy changes, increasing the cost of bribery through external monitoring and sanctions, and designing self-control systems have proven capable of reducing crime in government organisations. Multiple frameworks are needed to understand a complex problem like corruption instead of trying to find a single perfect answer (Martinsson, 2021).

The above studies highlight the urgency of presenting effective and efficient village governance as a preventive measure against potential corruption. The importance of regulatory compliance creates a profound legal foundation for budget management. Villages must ensure that their budgeting process is within applicable legal provisions. The principle of internal control is a primary foundation for ensuring that every stage in village budget management is closely monitored, from planning to reporting, to detect potential irregularities early and address them quickly.

The concept of accountability plays a primary role in building public trust in village financial management. Villages must practice good accountability and organise clear and transparent accountability for using the budget. These are the mechanisms to ensure every budget-related decision

can be accounted for by the community. In addition, transparency is also profound in building quality village governance. By giving the public broad access to information on village budgets, policies, and decisions, transparency creates a basis for better external and internal oversight.

Community involvement is also a crucial element in creating good village governance. By involving the community in the budget decision-making process, village governments will gain various views and inputs and increase public awareness and responsibility towards the village budget management. Continuous implementation of these principles is primary to creating effective, efficient, and corruption-free village governance, maintaining sustainable development, and improving the welfare of village communities.

Table 3 presents another set of intervention measures to address corruption in the village fund in Indonesia. These measures provide conceptual efforts to address corruption, covering various aspects such as capacity building, technology utilisation, external oversight, community engagement, and legal enforcement.

Previous researchers have proposed significant interventions to address village fund corruption that mark advancements in strategic thinking. However, these proposals primarily exist at a conceptual level, demanding rigorous testing in real-world conditions. Careful evaluation becomes imperative to ensure the positive impact of interventions such as utilising internal control and compliance with regulations, increasing accountability and transparency, empowering village officials, leveraging information technology, conducting independent audits, fostering community participation, and implementing legal sanctions.

Table 3

Intervention measures to addressing village fund corruption in Indonesia

Categories	Descriptions
Empowerment of village officials	Empowering village officials through training and capacity building is a primary step in preventing corruption at the village level (Sarawati, 2019).
The use of information technology	Information technology can effectively minimise the risk of corruption of village funds by strengthening electronic reporting and monitoring mechanisms (Fikri et al., 2018; Magdalena et al., 2020; Rosyidin et al., 2023; Wijayanti et al., 2020).
Independent audit	External parties play a crucial role in uncovering potential budget fraud and improving accountability at the village level (Putri et al., 2024; Srirejeki & Faturokhman, 2020).
Increasing community participation	Active community participation creates legitimacy and provides effective social control over village budget management (Diansari et al., 2023; Fadhal et al., 2021; Savitri et al., 2023; Susilowati et al., 2018).
Strict enforcement of legal sanctions	Strict legal sanctions against corruption are critical in ensuring the effectiveness of corruption eradication efforts at the village level (Farid et al., 2022; Puasanto et al., 2023; Triana & Mulya, 2022).

Source: Authors' work

To truly gauge their effectiveness and adaptability across diverse Indonesian villages, in-depth field research is paramount. Through this field research, these conceptual formulations can be proven, refined, and widely adopted, thereby instigating genuine change in the fight against corruption at the village level. Future studies must converge on specific formulas recommended by all prior researchers as the most effective strategy for combating corruption, especially within the village sector. There needs to be a consensus regarding strategies that present a research gap that future scholars must address. Further research is needed to conduct comprehensive syntheses based on experimentation and academic exploration. Bridging this research gap is crucial to formulating targeted and effective strategies for combating corruption in the unique context of Indonesian villages.

The comprehensive understanding of the widespread problem of corruption and the abuse of power in Indonesian villages carries significant consequences for policymaking and governance. This understanding presents prospects for substantial change and enhanced control at both the local and national scales. An essential prerequisite for evidence-based policymaking is a thorough understanding of these obstacles. As a result, policymakers possess the ability to design focused efforts that effectively tackle the underlying factors contributing to corruption and maladministration. This approach ensures that policies are customised to suit Indonesian villages' requirements and circumstances. Moreover,

a comprehensive comprehension of the pervasiveness and intricacies of corruption facilitates policymakers in formulating more efficacious approaches toward prevention, enforcement, and remediation.

The significance of tackling corruption in Indonesian villages must be considered. Corruption leads to misallocating crucial resources meant for necessary services, undermines public confidence in governmental institutions, and poses obstacles to socioeconomic progress. Moreover, it sustains a recurring pattern of impoverishment and disparity, exacerbating the marginalisation of already vulnerable groups. Furthermore, corruption erodes the foundations of the rule of law and democracy, undermining the principles of justice and fairness. Additionally, it stifles innovation and economic growth. Therefore, it is imperative to confront this issue head-on to unlock the full potential of Indonesia's rural areas. Addressing corruption in villages is not just a matter of compliance with anti-corruption laws; it is a moral imperative that resonates with integrity, transparency, and accountability. By doing so, Indonesia can pave the way for more equitable and prosperous communities where every citizen can enjoy the benefits of good governance and inclusive development.

The long narrative above reflects several primary points: (1) the importance of deeply understanding the complex factors that intersect with each other to design appropriate reform strategies to fight corruption and abuse of power in Indonesian villages, (2) the negative impact

of corruption on resource allocation, public trust, socioeconomic progress, equity, innovation, and economic growth cannot be ignored, (3) systemic efforts are needed to address village corruption by improving oversight, strengthening internal governance, and increasing transparency and accountability. It is also essential to involve the community in decision-making and oversight, (4) the existence of evidence-based policies and a robust governance framework is increasingly important to address village corruption in Indonesia, and (5) collaboration between the government, civil society organisations, international partners, and local communities in developing anti-corruption initiatives is also an essential factor that should be intensified. These ideas underscore the need for comprehensive and collaborative action to address the issue of village corruption in Indonesia and realise inclusive development with equitable outcomes.

CONCLUSION

The democratic pattern of state management will have a significant potential to narrow the space for rampant corruption. Nevertheless, the validity of this statement is contingent upon adopting such measures in specific countries, as there is no generally applicable way to combat corruption. Instead, analysing the local environment and other influential elements is imperative. Like Indonesia, the post-reform political system's democratisation through the implementation of decentralisation not only results in improvements to people's

lives but also contributes to an increase in the misappropriation of public finances conducted by government officials at the regional and village levels. Village decentralisation is supposed to improve public services, reduce poverty as the dominant form of social inequality, and recognise village rights and obligations. However, its implementation still faces complex challenges, such as unfair allocation of village funds, lack of clear boundaries of responsibility for public service delivery, chaotic financial management, and inadequate control and accountability mechanisms. Lack of transparency and accountability can lead to corruption, while good transparency and accountability can encourage village governments to provide better services to the community.

This literature review comprehensively explores maladministration, corruption, and abuse of power in Indonesian villages post-decentralisation. It highlights the challenges, consequences, and the need for a holistic understanding of the factors contributing to corruption in village governance. The novelty of our overview lies in the nuanced exploration of individual, institutional, and structural factors contributing to village fund corruption and the set of intervention measures to address rural corruption proposed conceptually by previous researchers to provide a holistic perspective beyond conventional analyses. Through collecting and synthesising preexisting research, case studies, and expert perspectives, this narrative overview serves as a valuable addition to the existing

body of knowledge concerning corruption and governance matters in Indonesia. Some literature reviewed in this paper has provided descriptions of the motives, modes, and impacts of corruption in development funds by certain actors, especially those holding the highest authority at the village level.

Given the intricate network of factors contributing to corruption, there is an urgent requirement for extensive governance reforms across various levels. Establishing effective village governance emerges as a pivotal preventive measure against corruption, encompassing adherence to regulations, internal controls, accountability, and transparency. Additional recommended interventions comprise the empowerment of village officials, the application of information technology, external audits, heightened involvement of the community, and the rigorous enforcement of legal sanctions. Furthermore, continuous assessment and adjustment of these strategies through meticulous field research are imperative to assess their efficacy and make well-informed modifications.

Our primary recommendations for policymakers and stakeholders addressing corruption in village funds in Indonesia are similar to or align with earlier researchers' suggestions. These include establishing robust whistleblower protection mechanisms that will encourage individuals who know about someone's corrupt practices to come forward without fear of retaliation. Additionally, adopting advanced technological solutions such as blockchain and other emerging technologies can enhance transparency in financial transactions and ensure the traceability of funds.

This narrative overview underscores the urgency of addressing corruption in Indonesian villages, emphasising the need for evidence-based policies, collaboration, and systemic efforts. It highlights the interconnectedness of factors and the need for comprehensive strategies that involve local communities, strengthen governance structures, and promote transparency and accountability. The moral imperative of addressing corruption aligns with the values of integrity and justice, aiming for inclusive development and equitable outcomes in Indonesia's rural areas. While there are no macro-level studies and specific but most effective recommended formulas for tackling corruption in the village sector, this literature review encourages future researchers to fill this gap by unravelling the intricacies of village fund corruption and then finding the best strategies for mitigation.

This literature review has several notable limitations. First, despite briefly mentioning the background of village corruption in the pre-colonial, colonial, and post-independence periods, its dominant content is about the findings of previous research that are closely related to village corruption after the issuance of the Village Law, it looks pretty superficial to explain the issue of village corruption in Indonesia comprehensively. Ignoring an in-depth analysis of the background of village corruption before the issuance of the Village

Law may reduce understanding of the historical factors that influence the existence of village corruption in the present. Second, this review needs to identify the social and cultural factors that underlie the different meanings of the subjects of village budget corruption that may be discussed in each reviewed paper, even though these two aspects significantly influence perceptions and actions related to corruption in various contexts. Third, the authors merely reiterate the proposed interventions for combatting village fund corruption from previous researchers without comprehensively analysing the effectiveness of each solution. Fourth of the forty primary documents included in this overview, only some met the standards for peer-reviewed literature as they had an index in the Scopus database. The presence of non-peer-reviewed literature may create uncertainty regarding the quality of methodology and the validity of their findings, impacting the results' reliability and discussion in our narrative review.

ACKNOWLEDGEMENT

REFERENCES

Annahar, N., Widianingsih, I., Muhtar, E. A., & Paskarina, C. (2023). The road to inclusive decentralized village governance in Indonesia. *Sustainability (Switzerland)*, *15*(11), 1–23. https://doi.org/10.3390/su15118616

Antlöv, H., Wetterberg, A., & Dharmawan, L. (2016). Village governance, community life, and the 2014 village law in Indonesia. *Bulletin of Indonesian Economic Studies*, *52*(2), 161–183. https://doi.org/10.1080/00074918.2015.1129047

Ash-Shidiqqi, E. A., & Wibisono, H. (2018). Corruption and village: Accountability of village fund management on preventing corruption (problems and challenges). *Journal of Indonesian Legal Studies*, *3*(2), 195–212. https://doi.org/10.15294/jils.v3i02.27524

Atmadja, A. T., Saputra, K. A. K., & Manurung, D. T. H. (2019). Proactive fraud audit, whistleblowing, and cultural implementation of Tri Hita Karana for fraud prevention. *European Research Studies Journal*, *XXII*(3), 201–214. https://doi.org/10.35808/ersj/1466

Bonsu, A. B., Appiah, K. O., Gyimah, P., & Owusu-Afriyie, R. (2023). Public sector accountability: Do leadership practices, integrity and internal control systems matter? *IIM Ranchi Journal of Management Studies*, *2*(1), 4–15. https://doi.org/10.1108/irjms-02-2022-0010

Breakey, H. (2017). Arbitrary power, arbitrary interference and the abuse of power: Corruption, natural rights and human rights. *Research in Ethical Issues in Organizations*, *17*, 125–145. https://doi.org/10.1108/S1529-209620170000017013

Buchan, B., & Hill, L. (2014). *An intellectual history of political corruption*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137316615

Ceschel, F., Hinna, A., & Homberg, F. (2022). Public sector strategies in curbing corruption: A review of the literature. *Public Organization Review*, *22*(3), 571–591. https://doi.org/10.1007/s11115-022-00639-4

Clemente-Suárez, V. J., Navarro-Jiménez, E., Simón-Sanjurjo, J. A., Beltran-Velasco, A. I., Laborde-Cárdenas, C. C., Benitez-Agudelo, J. C., Bustamante-Sánchez, Á., & Tornero-Aguilera, J. F. (2022). Mis–dis information in COVID-19 health crisis: A narrative review. *International Journal of Environmental Research and Public Health*, *19*(9), Article 5321. https://doi.org/10.3390/ijerph19095321

Darsono, I., Saputra, M., & Abdullah, S. (2021). The effect of village fund management accountability in Aceh Singkil District, Indonesia. *Journal of Research in Business, Economics and Management*, *16*(1), 1–8.

Diansari, R. E., Musah, A. A., & Binti Othman, J. (2023). Factors affecting village fund management accountability in Indonesia: The moderating role of prosocial behaviour. *Cogent Business and Management*, *10*(2), 1–25. https://doi.org/10.1080/23311975.2023.2219424

Fadhal, M., Lubis, Z., Zulkifli, Z., Ismail, R., Sitorus, H., & Tobing, R. L. (2021). Village fund management: Pattern of community participation for village development. *International Journal of Multicultural and Multireligious Understanding*, *8*(3), 65–72. https://doi.org/10.18415/ijmmu.v8i3.2377

Farid, M., Huda, M., Darmawan, D., & Zakki, M. (2022). Eradicate village fund corruption: Efforts to strengthen law enforcement. *Bulletin of Science, Technology and Society*, *1*(2), 22–26.

Ferrari, R. (2015). Writing narrative style literature reviews. *Medical Writing*, *24*(4), 230–235. https://doi.org/10.1179/2047480615z.000000000329

Fikri, H., Suharto, D. G., & Nugroho, R. A. (2018). The utilization of electronic government in realizing transparency and accountability of village government: Synergy of implementation of electronic village budgeting and electronic monitoring system by Banyuwangi government. *International Journal of Multicultural and Multireligious Understanding*, *5*(4), 453–469. https://doi.org/10.18415/ijmmu.v5i4.425

Fisman, R., & Golden, M. A. (2017). *Corruption: What everyone needs to know*. Oxford University Press. https://libgen.is/book/index.php?md5=B0092657FA475E8C27DFDACA7671254B

Fürstenberg, A., Starystach, S., & Uhl, A. (2023). Culture and corruption: An experimental comparison of cultural patterns on the corruption propensity in Poland and Russia. *European Journal of Criminology*, *20*(5), 1719–1739. https://doi.org/10.1177/14773708221081017

Hannigan, Tim. (2015). *Brief history of Indonesia: Sultans, spices, and tsunamis: The incredible story of Southeast Asia's largest nation*. Tuttle Publishing. https://libgen.is/book/index.php?md5=D3BB8CC463BCAB7AE97723DED9EAC599

Harun, H., Graham, P., Kamase, H. P., & Mir, M. (2021). A critical analysis of the impacts of financial literacy and NPM on village funds initiative in Indonesia. *International Journal of Public Administration*, *44*(4), 336–345. https://doi.org/10.1080/01900692.2020.1722165

Hayati, T. (2019). Abuse of authority by government officials: Controversy between administrative and criminal sanctions. *Journal of Legal, Ethical and Regulatory Issues*, *22*(5), 1–7.

Hidayat, R. (2019). Public distrust and environmental citizenship: Primary impetuses of radical protest in Bima District, 2011-2012. *Jurnal Ilmu Sosial dan Ilmu Politik*, *22*(3), 187–200. https://doi.org/10.22146/JSP.35981

Hidayat, R., Hendra, H., & Iptidaiyah, M. (2019). The steep road to institutionalizing negotiable governance in the management of village development. *Journal of Governance*, *4*(1), 45–67. https://doi.org/10.31506/jog.v4i1.5369

Kok, J. K., Pheh, K.-S., & Hor, G. L. (2018). Psychological and social factors of depression

recovery: A narrative review. *Pertanika Journal of Social Sciences and Humanities*, *26*(1), 41–58.

Kroeze, R. (2021). Colonial normativity? Corruption in the Dutch–Indonesian relationship in the nineteenth and early twentieth centuries. In R. Kroeze, P. Dalmau & F. Monier (Eds.), *Corruption, empire and colonialism in the modern era* (pp. 173–208). Palgrave Macmillan. https://doi.org/10.1007/978-981-16-0255-9_7

Leh, F. C., Isa, N. K. M., Ibrahim, M. H., Ibrahim, M., Yunos, M. Y. M., & Ibrahim, J. A. (2023). A literature review of low-carbon urban tourism indicators and policy. *Pertanika Journal of Social Sciences and Humanities*, *31*(2), 655–681. https://doi.org/10.47836/pjssh.31.2.10

Lewis, B. D. (2015). Decentralizing to villages in Indonesia: Money (and other) mistakes. *Public Administration and Development*, *35*(5), 347–359. https://doi.org/10.1002/pad.1741

Li, H., & Yang, X. (2023). Who steals? Who bribes? Multilevel determinants of corruption types in China. *Crime, Law and Social Change*, *80*(2), 125–152. https://doi.org/10.1007/s10611-022-10070-x

Lucas, A. (2016). Elite capture and corruption in two villages in Bengkulu province, Sumatra. *Human Ecology*, *44*(3), 287–300. https://doi.org/10.1007/s10745-016-9837-6

Magdalena, H., Santoso, H., & Septryanti, A. (2020). Web-based village fund management monitoring system. *Proceedings of the 2020 8th International Conference on Cyber and IT Service Management (CITSM)*, 1–6. https://doi.org/10.1109/CITSM50537.2020.9268923

Martinsson, J. (2021). Combatting institutional corruption: The policy-centered approach. *Crime, Law and Social Change*, *75*(3), 267–280. https://doi.org/10.1007/s10611-021-09934-5

Milley, P., & Dulude, É. (2021). On maladministration in higher education: Towards a multilevel theoretical framework for understanding its emergence and persistence in an era of neoliberal, managerial and corporatist reforms. *International Journal of Leadership in Education*, *24*(5), 588–612. https://doi.org/10.1080/13603124.2020.1757158

Modesto, J. G., & Pilati, R. (2020). Why are the corrupt, corrupt? The multilevel analytical model of corruption. *Spanish Journal of Psychology*, *28*(e5), 1–13. https://doi.org/10.1017/SJP.2020.5

Monteduro, F., Hinna, A., & Moi, S. (2016). Governance and corruption in the public sector: An extended literature review. *Studies in Public and Non-Profit Governance*, *5*, 31–51. https://doi.org/10.1108/S2051-663020160000005002

Mukartono, A., & Rustamaji, M. (2019). The development of corruption in Indonesia (is corruption a culture of Indonesia?). *Proceedings of the 3rd International Conference on Globalization of Law and Local Wisdom (ICGLOW 2019), 358,* 139–141. https://doi.org/10.2991/icglow-19.2019.36

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods*, *22*, 1–18. https://doi.org/10.1177/16094069231205789

Novrizal, M., & Podger, O. (2014). Corruption prevention in village finance governance. *Proceedings of the International Conference on Eradicating Corruption: An Inter-Disciplinary Perspective, 24*(4), 60-67.

Nugraha, A., Novianto, W. T., & Riwanto, A. (2019). The law enforcement practices on corruption cases by the heads of villages in Indonesia. *International Journal of Academic Multidisciplinary Research*, *3*(10), 28–34.

Onyango, G. (2022). Bureaucratic corruption and maladministration in Kenya: A bureaucratic

analysis. *International Journal of African Renaissance Studies*, *17*(1), 171–189. https://doi.org/10.1080/18186874.2022.2084696

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information and Management*, *52*(2), 183–199. https://doi.org/10.1016/j.im.2014.08.008

Permana, I., Wulandari, S., & Komara, A. (2020). Developing good village governance to prevent corruption of village fund. *Systematic Reviews in Pharmacy*, *11*(8), 635–639.

Pozsgai-Alvarez, J. (2020). The abuse of entrusted power for private gain: Meaning, nature, and theoretical evolution. *Crime, Law and Social Change*, *74*(4), 433–455. https://doi.org/10.1007/s10611-020-09903-4

Pratolo, S., Sofyani, H., & Maulidini, R. W. (2022). The roles of accountability and transparency on public trust in the village governments: The intervening role of covid-19 handling services quality. *Cogent Business and Management*, *9*(1), 1–15. https://doi.org/10.1080/23311975.2022.2110648

Pribadi, B. (2022, September 26). 'KPK sebut sudah ada 686 oknum kades terjerat korupsi dana desa' [KPK mentions that 686 village heads have been caught in village fund corruption]. *Republika*. https://news.republika.co.id/berita/ritizy430/kpk-sebut-sudah-ada-686-oknum-kades-terjerat-korupsi-dana-desa

Puasanto, P., Rosmaniar, R., & Atika, A. (2023). Law enforcement corruption of village funds. *International Journal of Social Science Research and Review*, *6*(3), 200–206.

Purba, R. Br., Aulia, F., Tarigan, V. C. E., Pramono, A. J., & Umar, H. (2022). Detection of corruption in village fund management using fraud analysis. *Quality - Access to Success*, *23*(190), 120–128. https://doi.org/10.47750/QAS/23.190.14

Putri, C. M., Argilés-Bosch, J. M., & Ravenda, D. (2024). Creating good village governance: An effort to prevent village corruption in Indonesia. *Journal of Financial Crime*, *31*(2), 455–468. https://doi.org/10.1108/JFC-11-2022-0266

Romano, G., Masserini, L., & Lombardi, G. V. (2021). Environmental performance of waste management: Impacts of corruption and public maladministration in Italy. *Journal of Cleaner Production*, *288*, 1–10. https://doi.org/10.1016/j.jclepro.2020.125521

Rosyidin, A., Pangestu, A. Y., Iskandar, A. F., & Darusalam, D. (2023). Social media and e-government to prevent corruption in Indonesia villages. *Bulletin of Social Informatics Theory and Application*, *7*(1), 57–62. https://doi.org/10.31763/businta.v7i1.212

Sarawati, R. (2019). *Village capacity building through strong, developed, independent, and democratic village governance* [Paper presentation]. First International Conference on Islamic Development Studies 2019, Bandar Lampung, Indonesia. https://doi.org/10.4108/eai.10-9-2019.2289429

Savitri, E., Diyanto, V., & Herda Nik Abdullah, N. (2023). The effect of internal control systems, community participation, and village financial system on the accountability of village fund management: Prosocial behavior as moderation. *Journal for Re Attach Therapy and Developmental Diversities*, *6*(9s2), 772–789.

Shaleh, K., Sinaga, O., & Roespinoedji, D. (2021). Village budget and patron-client social relations: Implications for fraud: A geographical review. *Review of International Geographical Education Online*, *11*(1), 720–725. https://doi.org/10.33403/rigeo.800611

Sofyani, H., Pratolo, S., & Saleh, Z. (2022). Do accountability and transparency promote community trust? Evidence from village government in Indonesia. *Journal of Accounting*

and Organizational Change, *18*(3), 397–418. https://doi.org/10.1108/JAOC-06-2020-0070

Sofyani, H., Simali, M. F. I. S., Najda, T., & Al-maghrebi, M. S. (2020). The role of budgetary participation and environmental uncertainty in influencing managerial performance of village government. *Journal of Accounting and Investment*, *21*(2), 258–276. https://doi.org/10.18196/jai.2102148

Sofyani, H., Yaya, R., & Widiastuti, H. (2023). The story of rising corruption post-village government reform - A view of three theories: fraud, managerial hegemony, and culture. *Journal of Accounting and Investment*, *24*(1), 101–119. https://doi.org/10.18196/jai.v24i1.16462

Srirejeki, K., & Faturokhman, A. (2020). In search of corruption prevention model: Case study from Indonesia village fund. *Acta Universitatis Danubius, Œconomica, 16*(3), 214–229.

Srirejeki, K., & Putri, P. K. (2023). Local government officials' susceptibility to corrupt behavior: Some Indonesian evidence. *Journal of Financial Crime*, *30*(6), 1517–1533. https://doi.org/10.1108/JFC-06-2022-0132

Sugiharti, D. K., Muttaqin, Z., & Ramadhani, R. H. (2021). The supervision of village fund management to prevent corruption. *Padjadjaran Jurnal Ilmu Hukum*, *8*(3), 356–378. https://doi.org/10.22304/pjih.v8n3.a3

Susan, N., & Budirahayu, T. (2017). Village government capacity in the implementation of village law no. 6 of 2015 in Indonesia. In B. McLellan (Ed.), *Sustainable future for human security: Society, cities, and governance* (pp. 17–27). Springer. https://doi.org/10.1007/978-981-10-5433-4_2

Susilowati, N., Herdiani, A., & Widhiastuti, R. (2018). Village community participation model in village funds management to exteriorize the accountability. *KnE Social Sciences*, *3*(10), 1024–1038. https://doi.org/10.18502/kss.v3i10.3190

Tambunan, D. (2023). Indonesia under threat: The danger of corruption to political legitimacy. *Asian Journal of Comparative Politics*, *8*(1), 112–140. https://doi.org/10.1177/20578911221124965

Taufiq, M., & Hidayat, R. (2022). The constraining dynamics to promote the inclusive village governance in an autocratic leadership setting. *Journal of Governance and Local Politics*, *4*(1), 55–69. https://doi.org/10.47650/jglp.v4i1.442

Triana, M. A., & Mulya, K. B. (2022). Synchronization of sanctions against corruption by village heads. *Interdisciplinary Journal on Law, Social Sciences and Humanities*, *3*(2), 169–186. https://doi.org/10.19184/idj.v3i2.33051

Triyono, A. (2020). Framing analysis of village funding corruption in media suaramerdeka.com in central java, Indonesia, 2019. *International Journal of Criminology and Sociology*, *9*, 1154–1163. https://doi.org/10.6000/1929-4409.2020.09.136

Vel, J., Zakaria, Y., & Bedner, A. (2017). Law-making as a strategy for change: Indonesia's new village law. *Asian Journal of Law and Society*, *4*(2), 447–471. https://doi.org/10.1017/als.2017.21

Vickers, A. (2013). *A history of modern Indonesia* (2nd ed). Cambridge University Press. https://doi.org/10.1017/CBO9781139094665

Wijayanti, P., Nurhidayati., & Hanafi, R. (2020). Fraud prevention on village government: The importance of digital infrastructure supervision. In L. Barolli, F. Hussain & M. Ikeda (Eds.), *Complex, intelligent, and software intensive systems* (pp. 959–965). Springer. https://doi.org/10.1007/978-3-030-22354-0_91

Yates, A., & Graycar, A. (2020). Death by a thousand cuts: Recognizing, reporting, and responding to corruption in local government. *Public Integrity*, *22*(1), 85–103. https://doi.org/10.1080/10999922.2018.1534924

Yeo, O. T. S., Yusof, M. J. M., Maruthaveeran, S., Saito, K., & Kasim, J. A. (2023). A review of policies and regulations of green infrastructure establishment in Kuala Lumpur, Malaysia. *Pertanika Journal of Social Sciences and Humanities*, *31*(2), 561–584. https://doi.org/10.47836/pjssh.31.2.06

Yuliastina, M., Hudaya, M., Nor, W., & Ainun, B. (2021). New remuneration system for village government apparatus (apd): Can it halt fraud incident(s)? *Journal of Accounting and Investment*, *22*(1), 1–22. https://doi.org/10.18196/jai.v22i1.8936

Green Consumption: Behavior of Young Indonesian Consumers — Role of Environmental Knowledge, Responsibility, and Attitudes

Megawati Simanjuntak* and Ismaul Fitri

Department of Family and Consumer Sciences, Faculty of Human Ecology, IPB University, Bogor 16680, Indonesia

ABSTRACT

Younger consumers are important in addressing environmental problems caused by consumption activities. It is mediated through green behavior, which is influenced by Internal factors. Thus, this study examines how young Indonesian consumers' environmental attitudes, knowledge, and responsibility relate to their green consumption habits. Using the voluntary sampling approach, 328 respondents were chosen, and data were examined using SEM and descriptive analysis. The findings demonstrated that views about the environment and green consumer behavior were significantly influenced by environmental knowledge. A person's attitude greatly influences their green consumption habits. These findings support the cognitive-affective-behavioral (CAB) hypothesis, which can be applied to similar study subjects or topics. An important factor influencing perceptions is environmental responsibility. It follows the Norm Activation Theory (NAM). In contrast to green consumption behavior, the effect of environmental responsibility was not significant. It is not in line with the NAM, so it is recommended that the theory be modified or other theories be used to examine similar variables. Environmental responsibility is significantly impacted by environmental knowledge. While attitudes were in the high category, knowledge, responsibility, and green consumption behavior were in the medium category. The study suggests marketers offer and promote eco-friendly products to provide more choices to consumers. These results also provide good insights for policymakers to achieve sustainable consumption targets by considering knowledge enhancement through environmental campaigns on social media.

Keywords: Environmental attitudes, knowledge, responsibility, green consumption behavior

INTRODUCTION

Unsustainable consumption is detrimental to the environment in different ways. The primary driver of these negative consequences is consumption volume, which significantly contributes to greenhouse gas emissions (Bengtsson et al., 2018). Furthermore, this excessive consumption is linked to other environmental issues, such as ozone layer depletion, soil and water pollution, and ecosystem damage (Ekasari, 2017). Humans, as consumers, generate waste through daily activities, leading to pollution of air, water, and soil (Sunarsih, 2014). In Indonesia, 39.63% of waste comes from household waste. Based on the type, the majority of national waste generation is food and plastic, with a proportion of 40.64% and 18.08% (Ministry of Environmental and Forestry, 2022). Moreover, the increased use of chemical-based products has amplified hazardous waste generation (Utami, 2020).

A previous study stated that people's behavior strongly influenced the success of programs to address issues in selecting and consuming environmentally friendly products (Utami, 2020). Individual green behavior can largely be explained by internal factors like knowledge (Kostadinova, 2016). Prior studies by Chekima et al. (2015), Paul et al. (2016), Peattie (2010), and Shamsi and Siddiqui (2017) examined the influence of a person's demographics on their consumption patterns of green products. Environmentally conscious consumers have a greater desire to adopt friendly consumption habits (Peattie, 2010). Meanwhile, awareness and responsibility positively affect attitudes toward environmentally friendly products, and consumers concerned about the environment will likely pay more for renewable energy. Individuals with socially responsible attitudes toward issues tended to purchase friendly products (Oyewole, 2001). Attitudes are a key prerequisite of green consumption intentions and behaviors (Wang et al., 2021). Other research in sustainable tourism suggests that there is an attitude-behavior gap. The cause is the cognitive dissonance that occurs in the context of environmental sustainability (Juvan & Dolnicar, 2014).

Consumers play an important role in degradation (Berglund & Matti, 2014). Hence, good conduct must be adopted to address the issues. It can be achieved by actively engaging in eco-friendly practices and promoting green consumption (Zuraidah et al., 2012). Young consumers are concerned with the present situation and pay attention to future consequences (Hume, 2010). In Indonesia, the young consumer is primarily populated by millennials and Z generations. There were 75.49 million members of Generation Z (born 1997–2012), or 27.94% of the population. Meanwhile, the second most prevalent population was the millennial generation (born in 1981–1996), with 69.38 million people or 25.87% (Central Bureau of Statistics, 2020).

According to recent research, Generation Z—the generation living in the age of digital transformation—is more conscious of the environment, values diversity in ethics, and feels that equality and justice are important. They also value sustainability

in all forms (Gomes et al., 2023). At the moment, youthful customers who have been designated as a distinct market category and form influential consumer spending groups are the ones who buy friendly items (Yadav & Pathak, 2016). Since they have been exposed to the advancements in information technology since they were small children, even young customers now comprise a market segment that demands everything quickly (Hendrawan & Zorigoo, 2019). Sustainable consumption is an important, essential approach to addressing global environmental issues. Effective digital communication strategies, such as the AISAS model, can raise awareness and influence people's consumption behavior, especially among teenagers, to shift to more sustainable practices (Yuliati & Simanjuntak, 2024). Previous research on green behavior at the generational level has been conducted by Fabiola and Mayangsari (2020), Marmaya et al. (2019), Natakoesoemah and Adiarsi (2020), and Rizkalla and Erhan (2020). However, most previous studies are limited to the purchasing perspective of green consumption, excluding use and disposal. The consumption of young consumers of green products is interesting to analyze in describing the predominant consumer behavior. Meanwhile, the lack of interest in purchasing products can increase environmental damage.

Multiple parties have considered the concern about environmental damage caused by unsustainable consumption behavior. The attempt was to build on previous results by examining the effect of individual factors such as knowledge, responsibility, and attitudes as mediating variables on green consumption behavior in young Indonesian consumers, specifically, the millennial generation and Generation Z. This study defines consumption as green purchasing, usage, and recycling behavior. It is important to see the consistency of the behavior by looking at green consumption behavior from pre-, during, and post-consumption.

The Cognitive-Affective-Behavior (CAB) theory was adopted to explain the connection between cognitive (i.e., knowledge), emotional (i.e., attitudes), and behavior (Anuar et al., 2017). The cognitive component is knowledge or beliefs about an object, while the emotional component is the reaction or feeling associated with an attribute. Finally, behavior, often referred to as conative, is related to the intention or motivation to take action (Ojiaku et al., 2018). The fundamental theory in this study is the Norm Activation Model (NAM). According to the NAM model, an individual's pro-environmental behavior is determined by their responsibility, which is reflected in personal norms. Additionally, this study used the concepts of sustainable consumption and theories to describe green consumption behavior.

Ensuring sustainable consumption and production patterns is the 12th goal of the Sustainable Development Goals (SDGs). SDG 12 aims to reduce food waste in half worldwide, both at the retail and consumer levels. Development plans at the national and regional levels have also included

these objectives. For instance, the National Medium-Term Development Plan for 2020–2024 is governed by the Presidential Regulation of the Republic of Indonesia Number 18 of 2020.

The practical description of green consumption habits is expected to be an input for government policies in achieving SDG 12, namely responsible consumption and production. Additionally, this supports the national medium-term development plan as a form of government commitment to implementing the SDG. The extent of the role and dominance of internal factors is also reported in influencing green consumption behavior. This study proves whether consumers' decisions to behave in consuming green products follow the standard order of the learning hierarchy under the CAB theory. Finally, the results analyze the influence of the responsibility variable in NAM on green consumption behavior.

LITERATURE REVIEW

Cognitive-Affective-Behavior Theory (CAB)

The CAB model is based on the idea that consumer decisions are hierarchical, starting with cognition (personal beliefs, thoughts, perceptions, meanings, or attitudes) related to a particular issue or object. The cognitive and affective aspects are included in behavior as an intention to act or an actual action. The cognitive model is the individual's perception of the outcome of knowledge, while the affective model is an emotional assessment with preferences. The behavioral component

is the actions dedicated to the desire of individuals to consume a product (Chou et al., 2020). This CAB model is acceptable for all attitudes (Gigerenzer, 2020) but should be used to develop perceptual attitudes recognizing consumer decision-making (Gursoy et al., 2018).

Other studies have analyzed internal factors such as knowledge and attitudes based on the hierarchical order in the CAB theory model. The findings validate the CAB theory, which holds that attitudes are influenced by knowledge and consequently affect green consumer behavior (Anuar et al., 2017).

Norm Activation Model Theory (NAM)

NAM is a theoretical model widely used in several previous studies to determine the factors resulting in individual behavior (Onwezen et al., 2013). This theory was first proposed in 1977 by Schwartz and NAM to analyze the problem of pro-environmental behavior (Fang et al., 2019). The theory states that individuals show pro-environmental behavior due to the possession of personal values, understanding the consequences, and responsibility. Individuals are more committed to conservation through a connection with the negative consequences of their actions (Fang et al., 2019; Van der Werff et al., 2013).

Sustainable Consumption

A pattern of obtaining products and services to meet fundamental human requirements without having a detrimental effect on the environment is known as sustainable

consumption. Sustainable consumption results from a decision-making process by consumers as a responsibility to the environment in accordance with their needs. The process involves employing goods and services to provide necessities and promote a greater quality of life while minimizing the use of chemicals, natural resources, and the disposal of waste and pollutants to prevent harmful future generations (Seyfang & Smith, 2007). Some fundamental principles are understanding what products are consumed, knowing the impact of consumption, and recognizing the danger to communities, the national economy, and local industry (Sari, 2017). Thus, optimizing consumption's effects on the environment, society, and economy while also taking into account the demands of present and future generations is what is meant by sustainable consumption (Brix-Asala et al., 2016).

Green Consumption Behavior

Green consumption is a personal choice driven by environmental awareness. This behavior is shown by the individual who seeks, purchases, uses, evaluates, and discards products (Siringi, 2012). Consuming environmentally friendly products will be influenced by how someone acts (Andrew & Slamet, 2013). Meanwhile, awareness of issues is a multi-faceted construction that includes cognitive, emotional, and behavioral components (Tantawi & Shaughnessy, 2009). Consumer environmental behavior can be complicated for a variety of reasons, including cost (the scarcity of reasonably priced eco-friendly products), convenience (separating rubbish for recycling and keeping it at home instead of taking it to a collection point), and time (losing time) (Widayat et al., 2022).

Environmental Knowledge and Attitudes

Environmental knowledge should be possessed by promoting behavior commitments to purchase friendly products (Lee, 2011). Environmental knowledge also plays an important role in shaping environmental attitudes and encouraging more sustainable consumption behavior (Simanjuntak et al., 2023). Additional research has revealed that knowledge positively impacts opinions towards green items (Barber et al., 2010; Haryanto, 2018). This variable influences environmental care behavior through attitudes and motivation (Vicente-Molina et al., 2013).

Consumer attitudes are influenced by knowledge about the environment, which also promotes amiable conduct (Taufique et al., 2016). It inspires care for the environment and perceived validity to some extent. According to Lin and Niu (2018), environmental knowledge plays a crucial role in shaping environmental attitudes. The following is the theory that this study is founded on:

H1: Environmental knowledge has a significant effect on attitudes.

Environmental Knowledge and Responsibility

Environmental issues are associated with human subjective knowledge, behavior,

and attitudes (Slavoljub et al., 2015). Furthermore, sustainability can be achieved through knowledge, social behavior, and altered individual and group attitudes (Schutte & Bhullar, 2017). According to earlier studies, this variable has been linked to environmentally conscious actions. Habibi (2020) stated that more responsible behavior correlates with a higher awareness level. Thus, the hypothesis proposed in this study is:

H2: Environmental knowledge has a significant effect on responsibility.

Environmental Knowledge and Green Consumption Behavior

Knowledge influences consumer purchasing behavior (Yuliati & Simanjuntak, 2024). Knowledge of the natural world is often considered one of the primary reasons for environmentally friendly consumption. Consumers with better knowledge of environmental issues tend to be more aware of the impact of their product choices and are more likely to choose environmentally friendly products (Simanjuntak et al., 2023). Meanwhile, awareness directly affects behavior (Zareie & Navimipour, 2016), which is influenced by knowledge concerning the existence and degradation of quality, management issues, and the effectiveness of prevention methods. Awareness of degradation leads to individual and collective concerns (Hamiyati et al., 2020). A study conducted by Law et al. (2017) found that thorough environmental knowledge enables an individual to determine positive environmental behaviors.

Environmental knowledge is also strongly correlated with the intention to purchase green products. The following is the hypothesis in this study based on earlier research:

H3: Environmental knowledge has a significant effect on green consumption behavior.

Environmental Responsibility and Attitudes

Environmental responsibility is often considered a personal obligation that makes people feel guilty after failing (Middlemiss, 2010). However, studies have shown that consumers with environmentally conscious attitudes tend to purchase friendly products (Joshi & Rahman, 2015). This variable is associated with attitudes (Lee, 2008). Thus, the hypothesis proposed is:

H4: Environmental responsibility has a significant effect on attitudes.

Environmental Responsibility and Green Consumption Behavior

People with a positive understanding can positively affect the environment, have individual responsibility, and participate in sustainable development (Fryxell & Lo, 2003). A different study found that green consumption behaviors and products had moderate environmental responsibility. (Al-Haziazi & Muthuraman, 2019). For instance, awareness of consumers can affect the desire to purchase environmentally friendly products at a higher price. Similarly, an individual who takes environmental responsibility seriously will increase their

commitment to actualize the purchases of products (Junaedi & Fatmawati, 2016). Thus, the hypothesis proposed in this study is:

> H5: Environmental responsibility has a significant effect on green consumption behavior.

Environmental Attitudes and Green Consumption Behavior

An individual dedicated to protecting the environment should understand the value as an example of complete action. This behavior is expected to endure for an extended duration, thereby prompting all individuals in attendance to engage in identical actions (Septian et al., 2016). Many studies showed that consumer attitudes directly influence the intention to consume environmentally friendly products or behavior (Chen et al., 2012; Fielding et al., 2008; Paul et al., 2016). A global study on sustainability also showed a positive correlation between people's environmental awareness, green attitudes, and green consumption (Law et al., 2017). Thus, the hypothesis proposed in this study is:

> H6: Environmental attitudes have a significant effect on green consumption behavior.

METHODOLOGY

This study applied quantitative and qualitative methods. Quantitative procedures were conducted using online surveys to investigate the effects of environmental knowledge, environmental responsibility, environmental attitude, and green consumption behavior. Google Forms was used to gather data, and WhatsApp, Instagram, Twitter, and Telegram were used to share information. The questionnaire was completed independently by the respondents. Three sections made up of structured questions were included in the questionnaire: screening, profile, and research variable questions. Screening was conducted to ensure the respondents' criteria met the research needs. In the meantime, profiling was done to find out the respondents' gender, age, location of residence, level of education, occupation of work, and income.

The sample included 328 people, 160 of whom were millennials and 168 of whom were members of Generation Z. These participants were chosen voluntarily. They included Indonesian citizens between 17 and 41 who had purchased or used environmentally friendly products.

In the qualitative analysis, in-depth interviews conducted via Zoom Meeting involved six respondents (i.e., two millennials and four Generation Z) selected by purposive sampling. Physical retail observations were conducted by visiting modern retailers such as supermarkets in Jakarta. Digital observations were made by visiting online shop websites such as Zero Waste Indonesia, Demi Bumi, and Tokyo Fashionku. The previously discussed theory examines the important relationship between knowledge and attitudes, responsibility, and green consumer behavior (see Figure 1).

Knowledge of issues, environmental conditions, and solutions to protect the
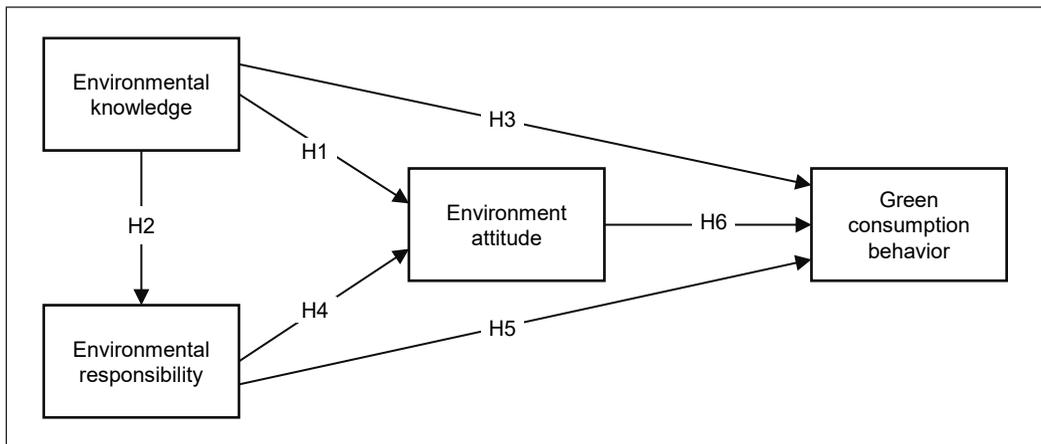
Figure 1. Study model
Source: Authors' work

environment was assessed using five different indicators (Safari et al., 2018). Furthermore, responsibility is a mindset or behavior that involves improving the environment. Responsible consumption was assessed by six modified versions of the ECO-SCALE (Environmentally Responsible Consumers Scale) (Stone et al., 1995). Attitudes are the feelings, judgments, and actions contributing to problems, measured by four different versions of inventory (Milfont & Duckitt, 2010). Additionally, green consumption is a behavior that includes the conservation of the environment, measured by ten indicators developed by some researchers (Barr et al., 2001; Straughan & Roberts, 1999; Quoquab et al., 2019). In the in-depth interview process, the questions were derived from Barr et al. (2001), Milfont and Duckitt (2010), Safari et al. (2018), Stone et al. (1995), Straughan and Roberts (1999) and Quoquab et al. (2019), to obtain more incisive information.

The variables were assessed using a five-point Likert scale with a weighted value of 1-5. The Likert scale assesses environmental knowledge: very aware, aware, somewhat aware, unaware, and ignorant. Furthermore, the scale for evaluating responsibility and attitudes is firmly in agreement, agree, entirely agree, disagree, and strongly disagree. For instance, the green consumption behavior variable comprises five scales: always, often, sometimes, rarely, and never.

Using SPSS 25.0, a descriptive analysis was carried out, and in the conclusion, the indicators are provided to help with the quantitative investigation. Descriptive statistics were used to identify each research variable. An index (i.e., 0–100 scale) was created by adding the average scores of all the indicators in a single variable. The goal is to compare every variable in the same way. The index is divided into three categories: low (i.e., index < 60.00), moderate (i.e., index 60.01–79.99), and high (i.e., index ≥ 80.00) to decide its classification. The following formula was used to determine the index:

$$\text{Index} = \frac{\text{Gained Score} - \text{Minimum Score}}{\text{Maximum Score} - \text{Minimum Score}} \times 100$$

Lastly, Lisrel 8.80 was utilized with Structural Equation Modeling (SEM) to investigate direct and indirect effects between variables. Covariance-based SEM (CB-SEM) analysis was selected to test the causal relationship between constructs or latent variables and confirm the two theories used as the basis of the study.

RESULTS

Respondents Profile

This study was primarily dominated by females (67.4%), and males only comprised only a third of the total respondents (32.6%), with millennials lower than Generation Z. Respondents from every province of Indonesia, with the most significant percentage (29.9%) were from the DKI Jakarta area. Most respondents had a high school diploma (47%) and were enrolled in college (41.5%). Furthermore, one-third had a monthly income of less than IDR 1 million (34.1%) and a monthly expenditure of less than IDR 1 million (38.1%).

Purchasing Behavior

Household goods are the most commonly used (63%) and purchased (57%) by respondents. Most purchases of environmentally friendly products are 1–2 pieces (48.5%), but the intensity of purchasing products in the last six months is twice (34.8%). The most environmentally friendly products are purchased through online platforms such as Tokopedia, Bukalapak, or Shopee (63.5%).

Green products are readily available online and in traditional retail outlets. Eco-friendly products are sold at almost all offline retailers, including supermarkets, such as tote bags, eco-friendly tissue, paper straws, Tumblr, and paper bags. Several online stores, including Demi Bumi, Zero Waste Indonesia, and Tokofashionku, sell environmentally friendly products. Online stores promote the zero-waste theme and sell products from brands. This store takes notice of packaging materials, and the merchandise at the Demi Bumi store includes accessories and equipment. Tokofashionku.com sells clothing made from bamboo fiber and organic cotton.

Description of Variable Indicators

Environmental knowledge is a person's knowledge about current environmental conditions and problems. Environmental knowledge was classified as moderate (42.4%), but there were also quite a lot of respondents in the high category (41.8%). In the distribution of respondents' answers based on the fifth indicator in the environmental knowledge variable, more than 80% of respondents know that consuming environmentally unfriendly products can cause environmental problems. It can be concluded that respondents already have a fairly good knowledge related to the environment.

The environmental responsibility in this study is moderate, at 56.1%, with an average index of 74.5%. Environmental responsibility is an individual's duty to the environment, making them feel guilty if the duty is not fulfilled. These data suggest that respondents are quite responsible for the environment. The attitude towards the environment is high, at 81.7%, with an average index of 89.3%. Respondents have high feelings, judgments and actions towards environmental issues. Furthermore, green consumption behavior in this study is moderate, at 54%, with an average index of 89.3%. It can be concluded that half of

the respondents prove their concern for the environment by demonstrating green consumption behavior. The results of the descriptive statistics are presented below (Table 1).

Evaluation of Model Fit Level

The overall results of the model fit test for the Normed Fit Index (NFI), Comparative Fit Index (CFI), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Incremental Fit Index (IFI), Goodness of Fit Index (GFI), and Root Mean Square Error of Approximation (RMSEA) are in Table 2.

Table 1
Descriptive statistics

Variables	Low (Index ≤ 60.00	Moderate (Index 60.01–79.99)	High (Index ≥ 80.00)	Range of Index	Mean ± SD
Environmental Knowledge (EK)	15.9	42.4	41.8	40.00–100.00	74.7 ± 11.6
Environmental Responsibility (ER)	7.6	56.1	36.3	25.00–100.00	74.5 ± 11.3
Environmental Attitude (EA)	8.8	9.5	81.7	40.00–120.00	89.3 ± 16.8
Green Consumption Behavior (GCB)	34.8	54.0	11.3	25.00–100.00	89.3 ± 11.4

Source: Author's work

Table 2
The overall model fit test

Goodness-of-fit	Cut-off-Value	Result	Description
Root Mean Square Error of Approximation (RMSEA)	RMSEA ≤ 0.08	0.037	Good fit
Root Mean Square Residual (RMR)	RMR ≤ 0.1	0.022	Good fit
Goodness of Fit Index (GFI)	0.80 ≤ GFI < 0.90	0.97	Good fit
Adjusted Goodness of Fit Index (AGFI)	0 < AGFI < 1; AGFI ≥ 0.9	0.95	Good fit
Incremental Fit Index (IFI)	0 < IFI < 1; IFI ≥ 0.9	0.99	Good fit
Normed Fit Index (NFI)	0 < NFI < 1; NFI ≥ 0.9	0.96	Good fit
Comparative Fit Index (CFI)	0 < CFI < 1; CFI ≥ 0.9	0.99	Good fit

Source: Author's work

The re-specification results demonstrate that the model's overall fit test is satisfactory and has a good fit. The root means square error of approximation is one of the absolute fit indices. The Chi-Square test can be offset in large samples using the RMSEA index. The model's RMSEA value needs to be less than or equal to 0.08 to be acceptable. The computed RMSEA score of 0.037 shows that the model is adequate and fits the required requirements. The model satisfies the good fit criteria, as evidenced by the RMR of 0.022, the GFI of 0.97, the AGFI of 0.95, the IFI of 0.99, the NFI of 0.96, and the CFI of 0.99.

The structural model's fit test was carried out following the general model's appropriateness assessment—the indicator variables' dependability on the latent variable forms the basis of the model's fit criterion. A valid indicator has a loading factor above 0.5 on a standard basis. Indicators EK3, EK5, ER1, ER2, ER3,

ER5, ER6, EA1, GCB2, GCB5, GCB6, GCB7, GCB9, and GCB10 are excluded because their standardized loading factor is considerably less than 0.5, indicating that not all indicators are valid (Figure 2).

Construct reliability (CR) and variance extracted (VE) measures for each construct were carried out to guarantee the validity and reliability of this model and result. Generally, a variance extracted (VE) value of ≥ 0.5 and an adequate construct reliability (CR) value of ≥ 0.7 are required. The study's findings indicated that every variable had CR values of at least 0.7, meeting the criterion for dependability. The variable extracted (VE) value of environmental responsibility is 0.98, which meets the requirements with VE $\geq$ 0.5. The variance extracted (VE) value of the other three variables is less than 0.5, namely 0.4. This value is still acceptable since convergent validity is still valid, VE is less than 0.5, and CR is greater than 0.6.
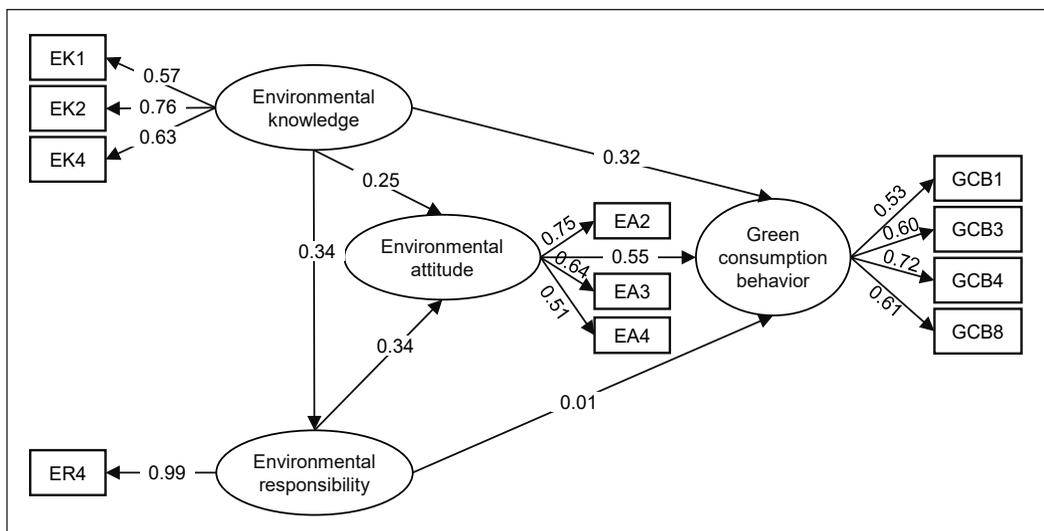
Figure 2. The result of the Structural Equation Modeling (SEM)
Source: Author's work

According to these findings, every indicator variable can assess its latent construct with a correct conclusion. Consequently, the measurement model evaluation findings indicate that the overall model fits the data (Table 3).

The structural equation model fit test is evaluated at a particular significance level. The t-value of each latent variable must be greater than 1.96 for the model to be considered fit; this is necessary because the hypothesis needs to be accepted at a significance level of 0.05 (95% confidence level). Five variables have t-values greater than 1.96. With a t-value of less than 1.96, H5 is regarded as an insignificant hypothesis (Figure 3).

Hypothesis Testing

Hypothesis testing examines the path coefficient and t-value of the structural equation model. The influence between variables is considered significant when the value is> 0.05 with a t-value value > 1.96. Conversely, the result is considered insignificant when the value is < 0.05 with a t-value of < 1.96. The estimated results of the SEM model based on the direct effect are shown below (Table 4).

The findings indicate that environmental attitude is positively and significantly influenced by environmental knowledge (β = 0.25; |t-value| = 3.20) and environmental responsibility (β = 0.34; |t-value| = 4.90). Environmental knowledge positively and significantly influences environmental

Table 3
Construct Reliability (CR) and Variance Extracted (VE)

Latent variables	CR	VE
Environmental Knowledge (EK)	0.70	0.40
Environmental Responsibility (ER)	0.98	0.98
Environmental Attitude (EA)	0.70	0.40
Green Consumption Behavior (GCB)	0.70	0.40

Source: Author's work
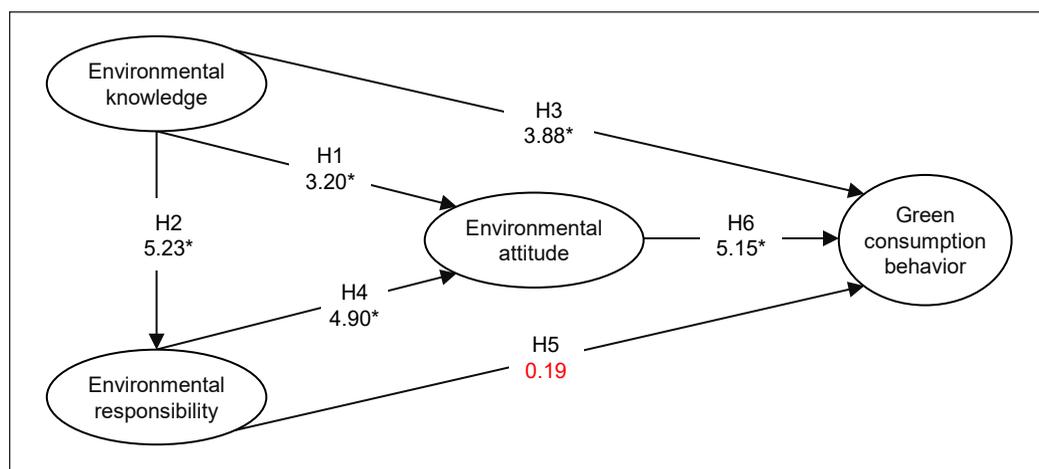
Figure 3. Results of the structural model fit test
Source: Author's work

Table 4
Hypothesis testing

Path			Path Coefficient	|*t-value*|	Results	
Environmental Knowledge (EK)	→	Environmental Attitudes (EA)	0.25	3.20	Significant	H1: Supported
Environmental Knowledge (EK)	→	Environmental Responsibility (ER)	0.34	5.23	Significant	H2: Supported
Environmental Knowledge (EK)	→	Green Consumption Behavior (GCB)	0.32	3.88	Significant	H3: Supported
Environmental Responsibility (ER)	→	Environmental Attitudes (EA)	0.34	4.90	Significant	H4: Supported
Environmental Responsibility (ER)	→	Green Consumption Behavior (GCB)	0.01	0.19	Insignificant	H5: Not Supported
Environmental Attitudes (EA)	→	Green Consumption Behavior (GCB)	0.55	5.15	Significant	H6: Supported

Source: Author's work

responsibility (β = 0.34; |t-value| = 5.23). The findings additionally indicate that green consumption behavior is positively and significantly influenced by environmental attitude (β = 0.55; |t-value| = 5.15) and environmental knowledge (β = 0.32; |t-value| = 3.88). Therefore, H5 is not supported by the fact that environmental responsibility (β = 0.01; |t-value| = 0.19) has a negligible impact on green consumption behavior. Finally, with the exception of H5, the directed hypotheses (H1, H2, H3, H4, and H6) that have been suggested have empirical evidence.

The results of the indirect effect test. The findings demonstrate that environmental responsibility (β = 0.12; |t-value| = 3.73) significantly mediates the connection between environmental knowledge and environmental attitude. Conversely, the connection between environmental knowledge and green consumption behavior is considerably mediated by environmental attitude (β = 0.20; |t-value| = 3.82) (Table 5).

In-depth Interview Results

In this study, millennials and Generation Z mentioned environmental issues related

Table 5
The indirect effect on the SEM model

The indirect path					Path Coefficient		*t-value*		Result
Environmental Knowledge (EK)	→	Environmental Responsibility (ER)	→	Environmental Attitudes (EA)	0.12	3.73	Significant
Environmental Knowledge (EK)	→	Environmental Attitudes (EA)	→	Green Consumption Behavior (GCB)	0.20	3.82	Significant

Source: Author's work

to waste, pollution, and disasters such as floods. Both generations learn about environmental issues through social media. They feel concerned about the issue of environmental damage. They think they can do nothing but try to introspect and improve their behavior. In this study, Millennials and Generation Z have shown their concern for the environment by implementing green consumption behaviors such as saving energy, throwing garbage in its place, using energy-efficient products, reducing the use of plastic by bringing their shopping bags when shopping, and bringing their own drinking water bottles. However, they do not have good recycling or waste management behaviors. Both generations find it quite difficult to reduce the use of plastic because businesses still provide products with plastic packaging. Nevertheless, they agree to use and encourage the availability of green products.

DISCUSSION

Environmental knowledge significantly influences attitudes (H1 supported). Therefore, there will be a significant increase in the attitudes of young Indonesian consumers with increasing knowledge. This study supports (Flamm, 2009; Gram-Hanssen, 2010; Polonsky et al., 2012) that consumers with high knowledge act positively on issues, and this variable is a significant predictor of attitudes and behavior. Other studies also align with this result, while knowledge positively affects attitudes toward environmentally friendly products (Barber et al., 2010;

Haryanto, 2018). Another study suggests that environmental knowledge negatively moderates the relationship between prosocial values and attitudes. Knowledge may affect the positive impact of prosocial values (Tamar et al., 2021). In this study, the relationship between environmental knowledge and attitude was significantly moderated by environmental responsibility.

Environmental knowledge influences behavior through attitudes and motivations (Vicente-Molina et al., 2013). This variable positively affects consumer attitudes and encourages environmentally friendly behavior (Taufique et al., 2016). This study found that environmental attitudes significantly mediate the relationship between environmental knowledge and green consumption behavior. The significant influence of knowledge on attitudes increases the CAB theory. As a result, individuals who understand the environment will attempt to have positive attitudes. Young consumers tend to develop more positive attitudes as their knowledge increases.

Environmental knowledge has a significant positive influence on responsibility (H2 supported). These results showed that this variable positively correlated with environmentally responsible behavior (Habibie, 2020). The product-moment and partial correlation show the association between knowledge and responsible behavior. According to the results, knowledge is the primary determinant of environmentally responsible behavior. The interpretation suggests that significantly increasing the variable enhances

environmentally responsible behavior (Habibie, 2020). High knowledge of the environment is directly proportional to its responsibility for the environment.

Knowledge of environmental science is associated with a change in green consumption behavior; hence, the third hypothesis (H3) is supported. Young Indonesians in this study will increase their green consumption significantly when the variable is improved. Previous studies showed that knowledge of a problem impacted decisions. Awareness is associated with behavior, and knowledge influences the variable. Consumers with a broad understanding of issues are motivated to purchase environmentally friendly products, adopt practices, and have a preference for purchasing behavior (Barber et al., 2010; D'Souza et al., 2006; Saleh et al., 2020). Therefore, knowledge tends to influence consumer behavior in a more environmentally friendly direction.

Environmental responsibility significantly affects caring attitudes (H4 supported). However, responsibility does not significantly influence green consumption behavior (H5 not supported). The insignificant influence is shown by a t-count value of 0.19, which falls below the threshold of 1.96. As a result, there is no discernible change in the green consumption behavior when the variable increases. The impact is shown by the route coefficient value, which is 0.01 and indicates that consumer behavior related to green consumption is 1% influenced by responsibility. Understanding that many

elements and a complex relationship exist between environmental responsibility and green consumption practices is crucial. This relationship is not always straightforward.

Although some studies have established a positive correlation between the two, there are several reasons why this impact might not be significant. Individual factors, such as values, beliefs, and personal circumstances, can significantly influence people's attitudes and behaviors, potentially diluting the direct impact of environmental responsibility. Other factors, such as price, convenience, and product availability, can also significantly influence consumer choices. Research suggests that people may perceive environmental issues as being psychologically distant from their daily lives, which can reduce the influence of environmental responsibility on actual consumption behavior. People may support eco-friendly practices in theory but may not always translate these beliefs into concrete actions due to the perceived lack of immediate consequences. In some cases, environmentally responsible alternatives may not be as readily available or as convenient as conventional options, which can discourage consumers from making green choices. Convenience often plays a crucial role in consumer decision-making.

This study contradicts previous results, in which individuals with responsibility affect sustainable consumption (Fryxell & Lo, 2003). Other studies have also reported that awareness can affect the desire to purchase products at a higher price and increase consumer dedication to

being environmentally friendly. Consumer commitment can be increased by improved responsibility to actualize purchases of environmentally friendly products (Junaedi & Fatmawati, 2016). However, having a role in the environment does not necessarily lead to developing friendly consumption.

Finally, attitudes significantly affect green consumption behavior (H6 supported), and the significant effect is derived from CAB theory. This theoretical framework states that behavior is derived from knowledge and attitudes. Meanwhile, the results follow the theoretical framework, and consumer behavior in purchasing products starts with knowledge. In this context, knowledge influences consumer attitudes toward environmental issues. The findings corroborate earlier research showing that customers are more likely to consume environmentally responsibly with positive attitudes and behaviors (Kim et al., 2012). The notion of attitudes also explains individual behavior and numerous earlier studies have looked at the connection between these factors.

Several studies have discovered a favorable correlation between worry or attitudes and actions related to the environment (Fraj & Martinez, 2006; Mostafa, 2007). As a result, perceptions greatly impact how people consume green. Additionally, this effect has a greater impact on green consumption behavior than on awareness and accountability.

This study provides reference and reinforces existing theories. First, the results show that knowledge significantly impacts attitudes and green consumption behavior. It supports the CAB theory, where behavior starts with knowledge and is followed by attitudes. The significant effect of responsibility on attitudes strengthens the NAM model, where responsibility promotes attitude formation. Finally, the negligible effect of the variable on green consumers' behavior can be incorporated into the model. Factors such as hopelessness can also prevent the activation of personal norms in action or behavior.

The results contribute to the concept of sustainable consumption, which is a pattern of goods and services without negative impact. They strengthen the concept of describing the respondents' environmental protection behavior. In addition, sustainable consumption is supported by good knowledge and the responsibility to develop positive attitudes. It motivates individuals to change their consumption behavior towards a more environmentally friendly direction.

In this study, Millennials and Generation Z can increase green consumption behavior by raising awareness of caring for and loving the environment. The process can be conducted by understanding issues and ways to protect the environment by adopting eco-friendly consumption habits. Furthermore, this green consumption habit includes three aspects. First, they can ensure their products are environmentally friendly by paying attention to the product's impact on the environment before purchasing. Second, the principle of thrift can also be followed when using goods or services. It means purchasing

only needed goods and services and being mindful of the impact.

Consumers can readily adapt to sorting waste by type and disposing it within their local communities. Furthermore, these individuals have the opportunity to harness their creativity to repurpose items suitable for independent recycling into new products. Another option is to use some items, such as cosmetic bottles, which are returned to the manufacturing companies for recycling. As a government, the Ministry of Commerce also motivated the supply of green products by promulgating regulations to encourage the development of industries that use goods or services of green raw materials. On the issue of waste management, the government needs to strengthen existing waste management regulations.

CONCLUSION

In conclusion, the green consumption behavior adopted by young Indonesian consumers in this study was energy-saving. It was practiced by disposing of garbage in its proper place, using energy-efficient products and shopping bags, reducing plastic consumption, and using drinking water bottles. Consumers also encountered challenges due to the prevalent availability of products packaged in plastic by numerous businesses. However, a promising trend was developed among younger consumers who showed a strong inclination toward adopting eco-friendly alternatives.

The findings demonstrated that knowledge has a strong positive impact on attitudes, behavior, and green consumption. It supported the CAB theory, which held that attitudes came after knowledge in shaping behavior. In the meantime, increasing knowledge improved attitudes, behaviors, and green purchasing. Socializing constructive campaigns on social media helped achieve this.

Environmental responsibility significantly positively affected attitudes but did not affect green consumption behavior. Attitudes had a significant positive effect on green consumption behavior, and this variable strengthened the NAM model, where responsibility encouraged the formation of attitudes. The young Indonesian consumers involved in this study have moderate environmental knowledge, responsibility, and green consumption behavior, while their attitudes are high. They improve their green consumption behavior by paying attention to 3 critical aspects: before buying products when using goods or services and waste management. The results also improved the concept of sustainable consumption, specifically in describing the respondents' behavior in preserving the environment.

Recommendations

In this study, millennials and Generation Z, who are young consumers, were expected to improve green consumption behavior by increasing their knowledge and responsibility for the environment. In this context, businesses motivated the growth of consumer green consumption

behavior by providing, campaigning, using environmentally friendly products, and reducing the use of plastics. The government also strengthened the regulations to encourage the growth of industries that provided and produced goods or services with environmentally friendly materials. It established stricter regulations and supervision of waste management.

This study was subjected to several limitations. First, the sample is unrepresentative of the general population. The purposive sampling was selected due to time and financial constraints. Second, during the in-depth interview phase, a number of potential sources meeting the specified criteria declined to participate in interview sessions, with a particular reluctance observed among millennial consumers. The following limitation was during field observation activities in DKI Jakarta. The future study was expected to include external factors such as the price of green products, the role of social media, government policies, or other factors contributing to the growth of green consumption behavior. Further studies could analyze eco-friendly products for comparison, such as modern and traditional retail levels, as well as urban and rural areas, to overcome the limitations.

ACKNOWLEDGEMENT

REFERENCES

Al-Haziazi, M., & Muthuraman, S. (2019). Consumer environmental responsibility towards green consumption in Sultanate of Oman. *Kuwait Chapter of Arabian Journal of Business and Management Review*, *8*(4), 24–37. https://doi.org/10.12816/0055340

Andrew, & Slamet, F. (2013). *The influence of environmental behavior on green purchasing behavior among young generation C in Jakarta* [Paper presentation]. Proceedings of The National Seminar and Call for Papers SANCALL 2013, Surakarta, Indonesia. https://publikasiilmiah.ums.ac.id/xmlui/bitstream/handle/11617/3841/02-Andrew.pdf?sequence=1

Anuar, M. M., Omar, K., & Ali, A. (2017). The influence of internal factors on consumer's green consumption behavior. *International Journal of Advanced and Applied Sciences*, *4*(12), 238–242. https://doi.org/10.21833/ijaas.2017.012.041

Barber, N., Taylor, D. C., Strick, S., Barber, N., & Taylor, D. C. (2010). Selective marketing to environmentally concerned wine consumers : A case for location, gender, and age. *Journal of Consumer Marketing*, *27*(1), 64–75. https://doi.org/10.1108/07363761011012967

Barr, S., Gilg, A. W., & Ford, N. J. (2001). Differences between household waste reduction, reuse and recycling behaviour: A study of reported behaviours, intentions and explanatory variables. *Journal of Environmental and Waste Management*, *4*(2), 69–82.

Bengtsson, M., Alfredsson, E., Cohen, M., Lorek, S., & Schroeder, P. (2018). Transforming systems of consumption and production for achieving the sustainable development goals: Moving beyond efficiency. *Sustainability Science*, *13*(6), 1533–1547. https://doi.org/10.1007/s11625-018-0582-1

Berglund, C., & Matti, S. (2014). Citizen and consumer : The dual role of individuals in environmental policy. *SHARP Research*

Programme, *15*(4), 550–571. https://doi.org/10.1080/09644010600785176

Brix-Asala, C., Hahn, R., & Seuring, S. (2016). Reverse logistics and informal valorisation at the base of the pyramid: A case study on sustainability synergies and trade-offs. *European Management Journal*, *34*(4), 414–423. https://doi.org/10.1016/j.emj.2016.01.004

Central Bureau of Statistics. (2020). *The results of the Population Census (SP2020) in September 2020 recorded a population of 270.20 million people.* https://www.bps.go.id/id/pressrelease/2021/01/21/1854/hasil-sensus-penduduk-sp2020-pada-september-2020-mencatat-jumlah-penduduk-sebesar-270-20-juta-jiwa-.html.

Chekima, B., Chekima, S., Azizi, S., Syed, W., Wafa, K., Igau, O. A., Laison, S., & Jr, S. (2015). Sustainable consumption: The effects of knowledge, cultural values, environmental advertising, and demographics. *International Journal of Sustainable Development & World Ecology, 23*(2). 210-220. https://doi.org/10.1080/13504509.2015.1114043

Chen, Y., Chang, C.-H., & Wu, F.-S. (2012). Origins of green innovations: The differences between proactive and reactive green innovation. *Management Decision*, *50*(3), 368–398. https://doi.org/10.1108/00251741211216197

Chou, S. F., Horng, J. S., Sam Liu, C. H., & Lin, J. Y. (2020). Identifying the critical factors of customer behavior: An integration perspective of marketing strategy and components of attitudes. *Journal of Retailing and Consumer Services, 55,* Article 102113. https://doi.org/10.1016/j.jretconser.2020.102113

D'Souza, C., Taghian, M., & Lamb, P. (2006). An empirical study on the influence of environmental labels on consumers. *Corporate Communications*, *11*(2), 162–173. https://doi.org/10.1108/13563280610661697

Ekasari, A. (2017). Purchase intention of reusable bags: The role of social marketing to achieve sustainable consumption. *Journal of Siasat Bisnis*, *21*(2), 119–130. https://doi.org/10.20885/jsb.vol21.iss2.art2

Fabiola, K., & Mayangsari, L. (2020). The influence of green skepticism, environmental knowledge and environmental concern on generation Z's green purchase intentions in Indonesia. *Journal of Social Sciences and Humanities*, *5*(8), 96–105. https://doi.org/10.47405/mjssh.v5i8.470

Fang, W.-T., Chiang, Y.-T., Ng, E., & Lo, J. C. (2019). Using the norm activation model to predict the pro-environmental behaviors of public servants at the central and local governments in Taiwan. *Sustainability (Switzerland)*, *11*(13), Article 3712. https://doi.org/10.3390/su11133712

Fielding, K. S., Mcdonald, R., & Louis, W. R. (2008). Theory of planned behaviour, identity and intentions to engage in environmental activism. *Journal of Environmental Psychology*, *28*, 318–326. https://doi.org/10.1016/j.jenvp.2008.03.003

Flamm, B. (2009). The impacts of environmental knowledge and attitudes on vehicle ownership and use. *Transportation Research Part D*, *14*(4), 272–279. https://doi.org/10.1016/j.trd.2009.02.003

Fraj, E., & Martinez, E. (2006). Environmental values and lifestyles as determining factors of ecological consumer behaviour: An empirical analysis. *Journal of Consumer Marketing*, *23*(3), 133–144. https://doi.org/10.1108/07363760610663295

Fryxell, G. E., & Lo, C. W. H. (2003). The influence of environmental knowledge and values on managerial behaviours on behalf of the environment: An empirical examination of managers in China. *Journal of Business Ethics*, *46*(1), 45–69. https://doi.org/10.1023/A:1024773012398

Gigerenzer, G. (2020). How to explain behavior? *Topics in Cognitive Science*, *12*(4), 1363–1381. https://doi.org/10.1111/tops.12480

Gomes, S., Lopes, J. M., & Nogueira, S. (2023). Willingness to pay more for green products: A critical challenge for Gen Z. *Journal of Cleaner Production, 390,* Article 136092. https://doi.org/10.1016/j.jclepro.2023.136092

Gram-Hanssen, K. (2010). Standby consumption in households analyzed with a practice theory approach. *Journal of Industrial Ecology*, *14*(1), 150–165. https://doi.org/10.1111/j.1530-9290.2009.00194.x

Gursoy, D., Ouyang, Z., Nunkoo, R., & Wei, W. (2018). Residents' impact perceptions of and attitudes towards tourism development: A meta-analysis. *Journal of Hospitality Marketing and Management*, *28*(3), 306–333. https://doi.org/10.1080/19368623.2018.1516589

Habibie, A. (2020). Relationship between self-efficacy and environment knowledge with responsibility environment behavior. *BIOEDUSCIENCE: Jurnal Pendidikan Biologi Dan Sains*, *4*(1), 21–26. https://doi.org/10.29405/j.bes/4121-264805

Hamiyati, Nadiroh, & Neolaka, A. (2020). Pengaruh pengetahuan pro environmental terhadap perilaku ramah lingkungan [The influence of pro-environmental knowledge on environmentally friendly behavior]. *Prosiding Seminar Nasional Pascasarjana Universitas Negeri Jakarta, 1*(01)*.* https://journal.unj.ac.id/unj/index.php/semnas-ps/article/view/16636/9119

Haryanto, B. (2018). The Influence of ecological knowledge and product attributes in forming attitude and intention to buy green product. *International Journal of Marketing Studies*, *6*(2), 83-91. https://doi.org/10.5539/ijms.v6n2p83

Hendrawan, D., & Zorigoo, K. (2019). Trust in website and its effect on purchase intention for young consumers on C2C e-commerce business.

Jurnal Aplikasi Manajemen, *17*(3), 391–399. https://doi.org/10.21776/ub.jam.2019.017.03.02

Hume, M. (2010). Compassion without action: Examining the young consumers consumption and attitude to sustainable consumption. *Journal of World Business*, *45*(4), 385–394. https://doi.org/10.1016/j.jwb.2009.08.007

Joshi, Y., & Rahman, Z. (2015). Factors affecting green purchase behaviour and future research directions. *International Strategic Management Review*, *3*(1–2), 128–143. https://doi.org/10.1016/j.ism.2015.04.001

Junaedi, D. K., & Fatmawati, I. (2016). Antecedents of purchasing green products. *Journal of Business Management and Marketing*, *9*(1), 175–186.

Juvan, E., & Dolnicar, S. (2014). The attitude-behaviour gap in sustainable tourism. *Annals of Tourism Research*, *48*, 76–95. https://doi.org/10.1016/j.annals.2014.05.012

Kim, S. Y., Yeo, J., Sohn, S. H., Rha, J. Y., Choi, S., Choi, A. Y., & Shin, S. (2012). Toward a composite measure of green consumption: An exploratory study using a Korean sample. *Journal of Family and Economic Issues*, *33*(2), 199–214. https://doi.org/10.1007/s10834-012-9318-z

Kostadinova, E. (2016). Sustainable consumer behavior: Literature overview. *Economic Alternatives*, *2*, 224–234.

Law, M. M. S., Hills, P., & Hau, B. C. H. (2017). Engaging employees in sustainable development – a case study of environmental education and awareness training in Hong Kong. *Business Strategy and the Environment*, *26*(1), 84–97. https://doi.org/10.1002/bse.1903

Lee, K. (2008). Opportunities for green marketing: Young consumers. *Marketing Intelligence and Planning*, *26*(6), 573–586. https://doi.org/10.1108/02634500810902839

Lee, K. (2011). The green purchase behavior of Hong Kong young consumers: The role of peer

influence, local environmental involvement, and concrete environmental knowledge. *Journal of International Consumer Marketing*, *23*(1), 21–44. https://doi.org/10.1080/08961530.2011.524575

Lin, S. T., & Niu, H. J. (2018). Green consumption: Environmental knowledge, environmental consciousness, social norms, and purchasing behavior. *Business Strategy and the Environment*, *27*(8), 1679–1688. https://doi.org/10.1002/bse.2233

Marmaya, N. H., Zakaria, Z., & Desa, M. N. M. (2019). Gen Y consumers' intention to purchase halal food in Malaysia: A PLS-SEM approach. *Journal of Islamic Marketing*, *10*(3), 1003–1014. https://doi.org/10.1108/JIMA-08-2018-0136

Middlemiss, L. (2010). Reframing individual responsibility for sustainable consumption: Lessons from environmental justice and ecological citizenship. *Environmental Values, 19*(2), 147–167. https://doi.org/10.3197/096327110X12699420220518

Milfont, T. L., & Duckitt, J. (2010). The environmental attitudes inventory: A valid and reliable measure to assess the structure of environmental attitudes. *Journal of Environmental Psychology*, *30*(1), 80–94. https://doi.org/10.1016/j.jenvp.2009.09.001

Ministry of Environment and Forestry. (2022). *National waste management information system*. https://sipsn.menlhk.go.id/

Mostafa, M. M. (2007). Gender differences in Egyptian consumers' green purchase behaviour: The effects of environmental knowledge, concern and attitude. *International Journal of Consumer Studies*, *31*(3), 220–229. https://doi.org/10.1111/j.1470-6431.2006.00523.x

Natakoesoemah, S., & Adiarsi, G. R. (2020). The Indonesian millenials consumer behaviour on buying eco-friendly products: The relationship between environmental knowledge and perceived consumer effectiveness. *International Journal of*

Multicultural and Multireligious Understanding, *7*(9), 292–302. http://doi.org/10.18415/ijmmu.v7i9.2047

Ojiaku, O. C., Achi, B. E., & Aghara, V. O. (2018). Cognitive-affective predictors of green purchase intentions among health workers in Nigeria. *Management Science Letters*, *8*(10), 1027–1038. https://doi.org/10.5267/j.msl.2018.8.001

Onwezen, M. C., Antonides, G., & Bartels, J. (2013). The norm activation model: An exploration of the functions of anticipated pride and guilt in pro-environmental behaviour. *Journal of Economic Psychology*, *39*, 141–153. https://doi.org/10.1016/j.joep.2013.07.005

Oyewole, P. (2001). Social costs of environmental justice associated with the practice of green marketing. *Journal of Business Ethics*, *29*(3), 239–251. https://doi.org/10.1023/A:1026592805470

Paul, J., Modi, A., & Patel, J. (2016). Predicting green product consumption using theory of planned behavior and reasoned action. *Journal of Retailing and Consumer Services*, *29*, 123–134. https://doi.org/10.1016/j.jretconser.2015.11.006

Peattie, K. (2010). Green consumption: Behavior and norms. *Annual Review of Environment and Resources*, *35*(1), 195–228. https://doi.org/10.1146/annurev-environ-032609-094328

Polonsky, M. J., Vocino, A., Grau, S. L., Garma, R., & Ferdous, A. S. (2012). The impact of general and carbon-related environmental knowledge on attitudes and behaviour of US consumers. *Journal of Marketing Management*, *28*(3–4), 238–263. https://doi.org/10.1080/0267257X.2012.659279

Quoquab, F., Mohammad, J., & Sukari, N. N. (2019). A multiple-item scale for measuring "sustainable consumption behaviour" construct: Development and psychometric evaluation. *Asia Pacific Journal of Marketing and Logistics*, *31*(4), 791–816. https://doi.org/10.1108/APJML-02-2018-0047

Rizkalla, N., & Erhan, T. (2020). Sustainable consumption behaviour in the context of millennials in indonesia – can environmental concern, self-efficacy, guilt and subjective knowledge make a difference? *Journal of Sustainable Businnes and Management Solutions in Emerging Economies*, *25*(3), 43–54. https://doi.org/10.7595/management. fon.2020.0001

Safari, A., Salehzadeh, R., Panahi, R., & Abolghasemian, S. (2018). Multiple pathways linking environmental knowledge and awareness to employees' green behavior. *Corporate Governance (Bingley)*, *18*(1), 81–103. https:// doi.org/10.1108/CG-08-2016-0168

Saleh, R. M., Anuar, M. M., Al-Swidi, A. K., & Omar, K. (2020). The effect of awareness, knowledge and cost on intention to adopt green building practices. *International Journal of Environment and Sustainable Development*, *19*(1), 33–58. https://doi.org/10.1504/ijesd.2020.105468

Sari, M. E. P. (2017). The role of society in achieving sustainable consumption patterns. *Trias Politika*, *1*(2), 1–15. https://doi.org/10.33373/jtp. v1i2.1057

Schutte, N. S., & Bhullar, N. (2017). Approaching environmental sustainability: Perceptions of self-efficacy and changeability. *Journal of Psychology: Interdisciplinary and Applied*, *151*(3), 321–333. https://doi.org/10.1080/0022 3980.2017.1289144

Septian, Y., Ruhimat, M., & Somantri, L. (2016). Learners' environmentally friendly behavior of high school students in the city of Bandung. *Social Science Education Journal*, *16*(2), 71–81. https://doi.org/10.17509/gea.v17i1.4678

Seyfang, G., & Smith, A. (2007). Grassroots innovations for sustainable development: Towards A new research and policy agenda. *Environmental Politics*, *16*(4), 584–603. https:// doi.org/10.1080/09644010701419121

Shamsi, M. S., & Siddiqui, Z. S. (2017). Green product and consumer behavior: An analytical study. *Pertanika Journal of Social Sciences and Humanities, 25*(4), 1545-1554.

Simanjuntak, M., Nafila, N. L., Yuliati, L. N., Johan, I. R., Najib, M., & Sabri, M. F. (2023). Environmental care attitudes and intention to purchase green products: Impact of environmental knowledge, word of mouth, and green marketing. *Sustainability, 15*(6), 5445. https://doi.org/10.3390/su15065445

Siringi, R. (2012). Determinants of green consumer behavior of post graduate teachers. *Journal of Business and Management*, *6*(3), 19–25. https:// doi.org/10.9790/487X-0631925

Slavoljub, J., Zivkovic, L., Sladjana, A., Dragica, G., & Zorica, P. S. (2015). To the environmental responsibility among students through developing their environmental values. *Procedia - Social and Behavioral Sciences*, *171*, 317–322. https://doi.org/10.1016/j.sbspro.2015.01.128

Stone, G., Barnes, J. H., & Montgomery, C. (1995). Ecoscale: A scale for the measurement of environmentally responsible consumers. *Psychology & Marketing*, *12*(7), 595–612. https://doi.org/10.1002/mar.4220120704

Straughan, R. D., & Roberts, J. A. (1999). Environmental segmentation alternatives: A look at green consumer behavior in the new millennium. *Journal of Consumer Marketing*, *16*(6), 558–575. https://doi.org/10.1142/ S0217751X15410031

Sunarsih, E. (2014). Concept of household waste in environmental pollution prevention efforts. *Journal of Community Health Sciences*, *5*(3), 162–167.

Tamar, M., Wirawan, H., Arfah, T., & Putri, R. P. S. (2021). Predicting pro-environmental behaviours: The role of environmental values, attitudes and knowledge. *Management of Environmental Quality: An International*

Journal, *32*(2), 328–343. https://doi.org/10.1108/MEQ-12-2019-0264

Tantawi, P., & Shaughnessy, N. O. (2009). Green consciousness of consumers in a developing country: A study of Egyptian consumers. *Contemporary Management Research*, *5*(1), 29–50. https://doi.org/10.7903/cmr.1149

Taufique, K. M. R., Siwar, C., Chamhuri, N., & Sarah, F. H. (2016). Integrating general environmental knowledge and eco-label knowledge in understanding ecologically conscious consumer behavior. *Procedia Economics and Finance*, *37*(16), 39–45. https://doi.org/10.1016/s2212-5671(16)30090-9

Utami, K. S. (2020). Green consumers behavior: Consumer behavior in purchasing environmentally friendly products. *Management, Cooperatives, and Entrepreneurship*, *9*(2), 208–223. https://doi.org/10.30588/jmp.v9i2.499

Van der Werff, E., Steg, L., & Keizer, K. (2013). It is a moral issue: The relationship between environmental self-identity, obligation-based intrinsic motivation and pro-environmental behaviour. *Global Environmental Change*, *23*(5), 1258–1265. https://doi.org/10.1016/j.gloenvcha.2013.07.018

Vicente-molina, M. A., Fernández-sáinz, A., & Izagirre-olaizola, J. (2013). Environmental knowledge and other variables affecting pro-environmental behaviour: Comparison of university students from emerging and advanced countries. *Journal of Cleaner Production*, *61*, 130–138. https://doi.org/10.1016/j.jclepro.2013.05.015

Wang, J., Shen, M., & Chu, M. (2021). Why is green consumption easier said than done? Exploring the green consumption attitude-intention gap in China with behavioral reasoning theory. *Cleaner and Responsible Consumption, 2,* Article 100015. https://doi.org/10.1016/j.clrc.2021.100015

Widayat, W., Praharjo, A., Putri, V. P., Andharini, S. N., & Masudin, I. (2022). Responsible consumer behavior: Driving factors of pro-environmental behavior toward post-consumption plastic packaging. *Sustainability (Switzerland)*, *14*(1). https://doi.org/10.3390/su14010425

Yadav, R., & Pathak, G. S. (2016). Intention to purchase organic food among young consumers: Evidences from a developing nation. *Appetite*, *96*, 122–128. https://doi.org/10.1016/j.appet.2015.09.017

Yuliati, L. N., & Simanjuntak, M. (2024). Digital communication innovation of food waste using the AISAS approach: Evidence from Indonesian adolescents. *Sustainability*, *16*(2), 488. https://doi.org/10.3390/su16020488

Yuliati, L. N., & Simanjuntak, M. (2024). Indonesian consumer complaint behavior: The role of information seeking, knowledge, purchase behavior, and tendency to talk. *Global Business and Finance Review*, *29*(1), 57-71. http://doi.org/10.17549/gbfr.2024.29.1.57

Zareie, B., & Navimipour, N. J. (2016). The impact of electronic environmental knowledge on the environmental behaviors of people. *Computers in Human Behavior*, *59*, 1–8. https://doi.org/10.1016/j.chb.2016.01.025

Zuraidah, R., Hashima, H. N., Yahya, K. W., & Mohamad, S. A. (2012). Environmental conscious behaviour among male and female Malaysian consumers. *OIDA International Journal of Sustainable Development*, *4*(8), 55–64.

APPENDIX

Research Questionnaire

Symbol	Indicators
Environmental Knowledge	
EK1	Know the problem of environmental pollution
EK2	Know about environmental issues
EK3	Knowing that environmental conditions are deteriorating
EK4	Know how to protect the environment from pollution
EK5	Consuming environmentally unfriendly products causes environmental problems
Environmental Responsibility	
ER1	The amount of energy used has no significant effect on environmental
ER2	Engaging in environmental activities helps save the environment for future generations
ER3	Worrying about the environment is useless
ER4	Environmentally responsible person
ER5	Do not buy products that are known to be bad for the environment
ER6	Excessive packaging sources of pollution can be avoided if producers are environmentally conscious
Environmental Attitude	
EA1	Believe that humans are currently exploiting environmental
EA2	Save water
EA3	Save electricity
EA4	Convince others that the environment is important
Green Consumption Behavior	
GCB1	Reduce the purchase of plastic packaging products
GCB2	Reducing the use of redundant items
GCB3	Avoiding single-use products
GCB4	Minimizing over-consumption
GCB5	Recycling plastic bottles
GCB6	Buying products at the lowest price. Regardless of their impact on the environmental
GCB7	Using lights with less power
GCB8	Buying products that cause the least pollution
GCB9	Buying products that are less harmful to environmental
GCB10	Using environmentally friendly products

Assessing Determinants of Technical Efficiency in Livestock Production: A Case Study from Shaanxi, China

Yujie Li, Ee Shiang Lim and Ghee-Thean Lim*

School of Social Sciences, Universiti Sains Malaysia, 11800 Gelugor, Penang, Malaysia

ABSTRACT

The demand for livestock products is rising, and China is actively encouraging farmers to increase their livestock production to meet this growing demand. At Shaanxi Province's livestock industry's current production output and growth rate, it appears unfeasible to meet the government's production target for 2025. Inefficiencies within livestock production can significantly impede the development of this industry. Therefore, this research employs the Data Envelopment Analysis (DEA) technique, considering Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS) assumptions, to assess the technical efficiency of the livestock industry in Shaanxi Province. The data utilised are secondary data from 2010 to 2019. The findings reveal that the Shaanxi livestock industry has an average technical efficiency of 0.84 (CRS) and 0.92 (VRS), suggesting that there is room for further production growth with the current inputs, breeding scales and technology. Although dairy cows, cattle and goats have achieved full technical efficiency. Technical and scale inefficiencies still exist in hog and layer farming practices, which can be improved to increase production. Notably, hog farming demonstrated the lowest technical efficiency, scoring 0.68. The results of factors affecting inefficiency suggest that increasing spending on disease prevention and raising the selling price can both improve technical efficiency. Additionally, reducing death loss has the potential to improve technical efficiency. Thus, the government is expected to promote farm consolidation and expansion while actively advocating for establishing livestock production cooperatives.

Keywords: Data Envelopment Analysis (DEA), inefficiency, livestock industry, technical efficiency, Tobit regression

INTRODUCTION

The global demand and production of livestock products are increasing rapidly, especially in China, due to population growth, rising income, and changes in lifestyle and dietary habits (Food and Agriculture Organization, 2021). In 2022, the National Bureau of Statistics of China reported that China's livestock industry produced 52.959 million tons of pork, 6.975 million tons of beef, 5.141 million tons of mutton, 24.825 million tons of poultry, 34.088 million tons of eggs, and 36.827 million tons of milk (National Bureau of Statistics of China, 2023). China is a major producer of livestock products, ranking first in the world in producing pork, mutton, broiler and eggs and third in dairy production (Food and Agriculture Organization, 2022a). Despite the increasing annual production of livestock products, China remains the world's top importer of livestock products (Food and Agriculture Organization, 2022b). Empirical research shows that China's current pork, beef, and mutton production is inadequate to meet domestic demand (Shi et al., 2015). In 2020, the proportion of imported meat in China's total meat production reached 12.7%, equivalent to 9.91 million tons, indicating a significant increase over the previous five years (AskCI Consulting, 2022). Amid the COVID-19 pandemic, China intensified restrictions on imported food products after detecting the virus in frozen food (Cadell, 2020). It has caused challenges not only for food suppliers and supply chains but also for China's dependence on imported meat and dairy products. The China Government's 14th Five-Year Plan (2021–2025) includes a target to increase livestock production, with an expected growth of 15% in meat production (Patton, 2022; Shaanxi Provincial Department of Agriculture, 2022). If all producers operate at full technical efficiency, the production of the livestock industry would more easily align with the output targets expected by the government, consequently reducing the demand for imported products.

The livestock industry plays a significant role in Shaanxi. Because it contributes to 80% of China's goat milk production, and its egg production is crucial to meeting the demand of surrounding provinces (National Bureau of Statistics of China, 2021). Moreover, pork and milk provide the most protein for Shaanxi residents. However, this industry confronts obstacles such as outbreaks of diseases, including African swine fever for hog (Wang, Zhao, et al., 2021) and avian influenza for layer. All these obstacles may reduce the production enthusiasm of farmers and lead to a scaling down of farming operations.

As of 2021, Shaanxi's livestock industry ranked 20th out of 31 provinces in mainland China, contributing 21.3% of the agricultural production value. The province produced 1.274 million tons of meat, 634,000 tons of eggs, and 1.619 million tons of dairy products. The most recent "14th Five-Year Plan for Livestock and Veterinary Development in Shaanxi" (Shaanxi Provincial Department of Agriculture,

2022) anticipated a rise in production, with an expected production of 1.8 million tons of meat, 0.8 million tons of eggs and 3 million tons of dairy products by 2025. Maintaining the current breeding practices will make it challenging for the livestock industry to achieve this government objective in production.

The output of the livestock industry in Shaanxi increased annually, but the growth rate has shown a decreasing trend year on year (Figure 1). In comparison to the national average growth rate, Shaanxi exhibits a lower growth rate (National Bureau of Statistics of China, 2021). It is possibly due to technical inefficiencies, which consequently impact the production of Shaanxi's livestock industry (Terry et al., 2021). Thus, it is important to know whether inefficiency exists in the Shaanxi livestock industry and identify the factors affecting it. Such an understanding would enable producers in the livestock industry to take necessary steps to improve efficiency

and enhance the overall performance of the livestock industry.

Many studies in the field of livestock do not distinguish between animal species or focus solely on specific animal species (Kuhn et al., 2020; Wang, Han, et al., 2021; Zhou et al., 2015), whereas this study adopts a more comprehensive approach. This study not only individually examines the primary livestock species within Shaanxi's livestock industry but also conducts a comparative analysis of technical efficiency across different scales of production. Moreover, a significant portion of research on factors affecting inefficiency focuses on emphasising the factors related to livestock farmers, such as farmers' gender or their experience in the livestock industry (Tian et al., 2015; Wang, Han, et al., 2021). In contrast, this paper extends the analysis to explore how medical and epidemic prevention, death loss, and selling prices affect inefficiency in the Shaanxi livestock industry.
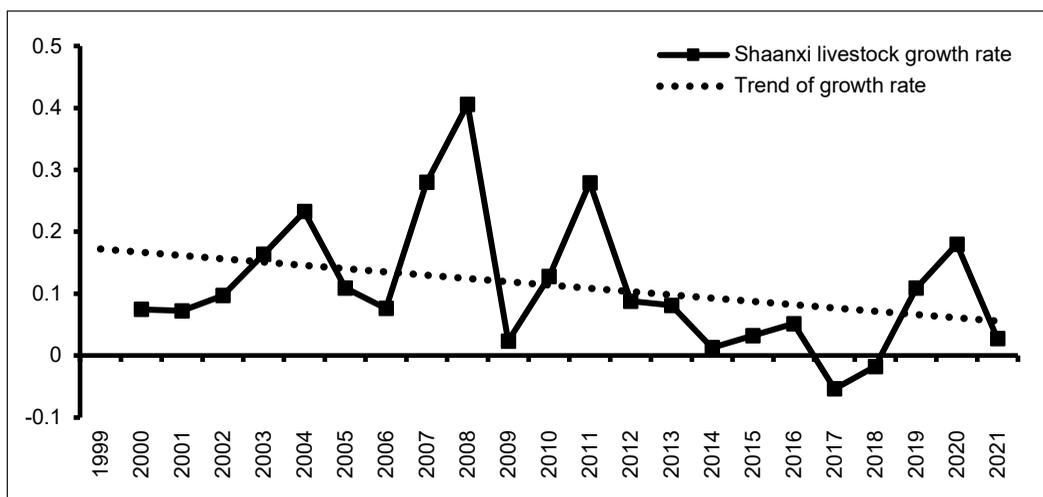
Figure 1. Shaanxi livestock growth rate
Source: National Bureau of Statistics of China (2021)

In light of the abovementioned considerations, this study seeks to accomplish two objectives: (1) to evaluate the level of technical efficiency and (2) to identify factors that affect technical inefficiency in the Shaanxi livestock industry. The findings of this study will provide valuable insights for policymakers on how to increase livestock production and modify relevant policies to promote sustainable development in the industry.

LITERATURE REVIEW

Production efficiency refers to a firm's ability to achieve the maximum output with a given set of inputs and contemporary technology (Farrell, 1957). However, efficiency cannot be directly observed. Therefore, appropriate methods are needed to measure it. There are two main techniques for measuring technical efficiency: non-parametric Data Envelopment Analysis (DEA) and parametric Stochastic Frontier Analysis (SFA). The choice of technique can influence the technical efficiency results, and there is no consensus on which technique is most appropriate for agricultural technical efficiency (Heshmati et al., 1995). DEA cannot require a specific functional form to be imposed on the data and can easily be adapted to multiple outputs. Additionally, DEA is deterministic and attributes all deviations from the frontier to inefficiency, making it sensitive to measurement errors and other statistical noise in the data. Unlike SFA, DEA is more inclusive of small samples (Zhu, 2009). SFA represents a parametric approach, also known as an econometric approach, which involves fitting an assumed structure of the observed data (Aigner et al., 1977; Meeusen & van Den Broeck, 1977). The main advantage of SFA is its ability to handle random noise. However, SFA requires a specific functional form to be imposed on the underlying technology and a distributional assumption to be imposed on the inefficiency term. Upon comparison, the DEA method has been selected as the preferred technique for measuring efficiency in this study.

DEA is a non-parametric technique to measure the efficiency of firms by comparing their production set with a production frontier. Based on the fundamentals of efficiency, DEA was developed into two assumptions: CCR (Charnes, Cooper and Rhodes) and BCC (Banker, Charnes and Cooper). The CCR assumption is also known as the CRS (constant returns to scale) assumption. Assuming that the Decision-Making Units (DMUs) are operating using the given inputs and technology, the results generated by DEA would indicate the technical efficiency (TE) score. When the score equals 1, the DMUs function efficiently and operate at optimal production levels. Conversely, a score lower than 1 suggests that the DMUs are inefficient (Zhang et al., 2017). Banker et al. (1984) introduced the BBC assumption (also known as VRS, variable returns to scale) as an improvement over the CRS assumption, which envelops the data points more tightly than the CRS assumption. Technical efficiency (TE) is the result of CRS DEA, while pure technical efficiency (PTE) is the

result of VRS DEA. The difference between these two is the scale efficiency (SE), which represents the ratio of the actual output of a DMU to its optimal output at the efficient scale of operation. If there is a difference between the TE and PTE scores, it indicates that the firm is operating at a suboptimal size. Scale efficiency equals one only when the scores of TE and PTE are equal (Färe & Lovell, 1978).

Efficiency is widely used in the livestock industry, with DEA commonly employed. Such studies have identified inefficiencies in several countries in the dairy industry, including Australia (Fraser & Cordina, 1999), Sweden (Hansson & Öhlmér, 2008), and Estonia (Luik-Lindsaar et al., 2019). Some studies have utilised both CRS DEA and VRS DEA for efficiency analysis. Notably, the VRS efficiency scores are relatively higher than the CRS efficiency scores, as VRS can capture the efficiency from scale benefits. For example, previous studies have found that in Hawaii's pig farming industry, the VRS efficiency score was 0.726, whereas the CRS score was 0.644 (Sharma et al., 1997). Similarly, Lansink and Reinhard (2004) reported a VRS score of 0.90 and a CRS score of 0.89 in the Netherlands. Mugera and Featherstone (2008) found VRS and CRS scores of 0.41 and 0.33 in the Philippines. Besides, Galluzzo (2019) investigated dairy farms in Iceland and reported a CRS score of 0.881 and a VRS score of 0.946. İlkikat et al. (2020) studied hair goat farms in Turkey and found a CRS score of 0.67 and a VRS score of 0.76. These findings suggest that the DEA technique can be applied to various livestock categories, regardless of geographic location. Also, scale inefficiency is prevalent in most livestock farming, making it important to analyse CRS and VRS to enhance farming scale.

Most previous studies on efficiency in the livestock industry in China have mostly concentrated on the hog industry (Kuhn et al., 2020; Somwaru et al., 2003; Tian et al., 2015; Wang, Zhao, et al., 2021; Yang et al., 2008; Zhou et al., 2015). For instance, Yang et al. (2008) surveyed 39 hog farmers to assess their technical efficiency in Taiwan between 2003 and 2004. The results demonstrated that farms could increase their output by an average of 52.8% while maintaining the same input levels. In an earlier study, Somwaru et al. (2003), who used a non-parametric technique, discovered that Shaanxi experienced technical inefficiency in the livestock industry with technical efficiency scores of 0.75, higher than the national average of 0.24 in 1996. Similarly, several studies, such as Tian et al. (2015), Zhou et al. (2015), and Wang, Zhao, et al. (2021) used parametric methods to examine the efficiency of the hog industry in China and found it to be inefficient. Moreover, these studies revealed that the technical efficiency of the northern provinces, including Shaanxi, is lower than that of other provinces.

The variance in technical efficiency among farms can be utilised to discover the factors that affect inefficiency. The standard approach entails conducting a regression analysis of efficiency scores

against a series of explanatory variables (Lansink & Reinhard, 2004). In economics, determining the impact of exogenous factors on production involves converting the technical efficiency score into technical inefficiency, which is obtained by subtracting the technical efficiency score from 1 (Coelli et al., 2005; Farrell, 1957). Then, factors affecting inefficiency can be identified through Tobit regression analysis applied to truncated data (Dogan et al., 2018; Liu et al., 2021; Zhang et al., 2017). Several studies examining efficiency in the hog farming industry across different countries, including China (Tian et al., 2015), the Philippines (Mugera & Featherstone, 2008), and Turkey (İkikat et al., 2020), found that a higher level of education among farmers leads to increased efficiency in hog farming, likely due to improved knowledge of scientific breeding techniques. Similarly, Wang, Han, et al. (2021) conducted a survey involving 449 herders within the Inner Mongolia grassland area of China in 2017. Their findings revealed that the existing policies have facilitated the expansion of livestock farming scales, leading to the departure of inefficient farmers from the industry. Additionally, Jo et al. (2021) surveyed farms in Heilongjiang Province, determining that a reduction in death losses contributes to enhanced technical efficiency within the livestock industry.

This study's selection of factors affecting inefficiency draws upon theoretical frameworks and previous empirical investigations. Within the Keynesian theory, governmental financial aid to farmers for epidemic prevention is perceived to safeguard their operations and stimulate heightened production. It aligns with the objectives outlined in the 10th to the 14th Shaanxi Five-Year Development Plans (spanning from 2001 to 2025) within the livestock industry, which accentuate the importance of bolstering epidemic control measures to enable farmers to achieve enhanced incomes (Crop Farming Management Office of Shaanxi Province, 2022; Shaanxi Government, 2008, 2022; Shaanxi Provincial Department of Agriculture, 2018; Shaanxi Statistics Bureau, 2015). Moreover, according to risk management theory, death loss has adverse impacts on production outcomes by introducing uncertainty and potential disruptions to agricultural operations, thereby diminishing efficiency. The law of supply implies that as prices of agricultural products escalate, farmers exhibit a heightened inclination to expand production and augment supply to the market. Conversely, lower prices might lead to diminished production or market exit, resulting in reduced quantities supplied.

METHODOLOGY

Variable Selection and Description

To obtain precise results of technical efficiency, the number of input variables should be neither too few nor too many (Reinhard et al., 1999). Based on livestock features and plenty of readings, such as Sharma et al. (1997), Fraser and Cordina (1999), Galluzzo (2019) and Soh et al. (2021), this study selected one output variable, three

input variables and three factors that are likely to have an impact on inefficiency. The description of the variables is as follows:

(1) Output: The output value measurement is based on each animal's farm price. In the case of hogs, cattle, and goats, the primary output is determined by the live weight price of each animal at the time of sale. In the case of layer, the output value is the farm price of total eggs from 100 layers.

(2) Labour: In livestock production, labour input is quantified as the number of employees' working days required per animal. However, when evaluating layers specifically, the measurement unit employed is the working days required for the management of 100 layers. These working days are calculated based on 8 hours per day. For instance, if one goat needs six workdays of labour, an employee would spend 48 hours on one goat.

(3) Young: This variable is measured as the average purchase price of each young animal. The expenses of young animals vary across different livestock categories. The expenses for cattle, hogs, and goats consist of the acquisition cost of each young animal. While the layer is small livestock, the expenses are the purchase price of 100 chicks.

(4) Feed: The weight of feed is the total weight of grains, beans, fodder and additives consumed by each animal or 100 layers.

(5) MEP: Medical and epidemic prevention (MEP) expenses include immunising livestock, preventing epidemics, testing, quarantine, eradicating infectious diseases and government-enforced controlling measures. MEP measures the individual expenditures for each animal, while for layers, it quantifies the expenses of 100 layers. As it protects farmers' output, this variable is expected to have a negative impact on technical inefficiency.

(6) DL: Livestock death loss (DL) occurs when livestock dies due to various causes, such as disease, disaster, nutritional deficiencies and inadequate management. Calculating the death loss per animal involves dividing the number of animal deaths by the initial number of animals and multiplying the result by the farm price per animal. The loss of livestock can lead to reduced production, increased costs and decreased profitability for farmers.

(7) SP: The selling price (SP) represents the average selling price at which 50 kg of livestock products are sold in the Shaanxi agricultural wholesale market. This factor is expected to have a negative effect on technical inefficiency since high selling prices can encourage farmers to increase their breeding activities.

Data

A two-stage analysis was used in this study to examine the technical efficiency of the ten categories of animals in the Shaanxi livestock industry (Liu et al., 2021). In the

first stage, Data Envelopment Analysis (DEA) is employed to calculate the technical efficiency score of each category in the Shaanxi livestock industry. In the second stage, Tobit regression is used to estimate the factors that affect inefficiency in each category. Variables such as inputs (including labour, young animals and feed), output, and factors affecting inefficiency (medical and epidemic prevention, death loss and selling price) are selected. The data utilised in this study constitute secondary data sourced from the China Agricultural Product Cost–Benefit Compilation (Price Department of the National Development and Reform Commission & Price Cost Research Centre of the National Development and Reform Commission, 2020). This dataset, characterised as panel data, encompasses ten categories of animals within the Shaanxi livestock industry from 2010 to 2019. A linear interpolation method was applied to generate missing data points to minimise the impact of missing data. This study's price data is anchored to 2010, with adjustments made using the Shaanxi Livestock Consumer Price Index (CPI) or the Shaanxi Young Animal CPI. The CPI data are sourced from the Shaanxi Statistical Yearbook. The data

source also classifies livestock into four categories based on their breeding scale: backyard, small-, medium- and large-scale. Backyard farming refers to a practice that involves individual households or small farmers raising livestock in a small pen or backyard. This study utilises ten categories of DMUs: backyard and medium-scale dairy cows, backyard, small-, medium-, and large-scale hogs, small- and medium-scale layers, and backyard goats and cattle. The classification standard of the breeding scales for each category is shown in Table 1.

Models

Essentially, the DEA technique is built on the technical assumptions of Constant Returns to Scale (CRS), which assumes that the DMUs are operating at an optimal scale. Banker et al. (1984) extended this method to incorporate technologies with Variable Returns to Scale (VRS), which would envelop the data points more tightly than the CRS assumption. Prior to computing technical efficiency, it is essential to select the orientation of minimising the inputs or maximising the outputs based on the variable (input or output) that the manager needs to control the most. Substantively,

Table 1
Standard classification of breeding scale based on quantity

Livestock	Backyard	Small-scale	Medium-scale	Large-scale
Dairy Cow	Q≤30	-	50<Q≤500	-
Hog	Q≤30	30<Q≤100	100<Q≤1000	Q>1000
Layer	-	300<Q≤1000	1000<Q≤10000	
Goat	Q≤100	-	-	-
Cattle	Q≤50	-	-	-

Notes. Q = Quantity, signifying the number of animals bred on the farms
Source: Price Department of the National Development and Reform Commission (2019)

the Shaanxi Provincial Government has been encouraging the enhancement of livestock production. Thus, this study adopts an output-oriented approach. Following Coelli et al. (2005), the output-oriented VRS assumption in Equation 1 and the CRS assumption in Equation 2 for measuring technical efficiency are given:

$$max_{\phi,\lambda} \; \phi,$$

subject to

$$-\phi q_i + Q\lambda \geq 0, \qquad (1)$$
$$x_i - X\lambda \geq 0,$$
$$I'_{I\times 1}\lambda = 1$$
$$\lambda \geq 0,$$

$$max_{\phi,\lambda} \; \phi,$$

subject to

$$-\phi q_i + Q\lambda \geq 0, \qquad (2)$$
$$x_i - X\lambda \geq 0,$$
$$\lambda \geq 0,$$

where λ represents a $I \times 1$ vector of constants (weights). The parameter ϕ signifies the ratio between the distance from the optimum point to the origin of the coordinate axes and the distance from the observed point to the origin of the coordinate axes. This ratio is equivalent to calculating the inverse of technical efficiency, subject to the constraint that $1 \leq \phi \leq \infty$. The vectors X and Q are the observed inputs and outputs, respectively. The value of $\frac{1}{\phi}$ refers to the level of technical efficiency (TE) ranging from 0 to 1 of the i-th decision-making unit,

with a value of 1 indicating that the farm is technically efficient and on the production frontier (Farrell, 1957). The VRS DEA assumption involves three constraints. In the first constraint, the observed output (q_i) of the i-th farm is multiplied by ϕ and compared to the maximum output vector of the theoretically efficient farm ($\phi\lambda$). With the same quantity of inputs, the constraint indicates that the theoretically efficient farm produces more or the same volume output than the actual output produced by the i-th farm. The second constraint illustrates that the observed input (x_i) in the i-th farm is more than or equal to the input ($X\lambda$) of the theoretically efficient farm. The third constraint of $I'_{I\times 1}\lambda = 1$ signifies the inefficiency of a farm evaluated against other farms of similar size. Such constraint enables the evaluation of farm efficiency in terms of technical and scale efficiencies (Mohd Idris et al., 2013).

If the technical efficiency score turns out to be $\frac{1}{\phi} = 1$, this farm is technically efficient. Then, the output of this farm is as much as the production of the technically efficient farm using the same volume of inputs. If the efficiency score turns out to be $\frac{1}{\phi} < 1$, the farm is technically inefficient. It means that the farm's output can be increased to the level of $\phi\lambda$. Notably, the linear programming problem needs to be solved I times to obtain a value for each sample farm. Hence, a value of ϕ is calculated for each farm. As shown in Equation 2, the VRS DEA assumption can be transformed into the CRS DEA assumption by removing the constraint of $I'_{I\times 1}\lambda = 1$ (Siafakas et al., 2019).

Scale efficiency refers to the extent to which a DMU is operating at the most productive scale size. If a DMU is operating below its optimal scale size, it may be able to increase its efficiency by adjusting its practice scale. Dividing the CRS efficiency score by the VRS efficiency score allows the capture of the impact of the scale effect. Scale efficiency can be expressed as:

$$SE = \frac{TE_{CRS}}{TE_{VRS}} \tag{3}$$

After obtaining the technical efficiency scores through DEA in the first stage, the efficiency scores are converted into inefficiency scores. Subsequently, a regression analysis is conducted to examine the relationship between inefficiency scores and other exogenous variables in the second stage. The technical inefficiency scores are derived by subtracting the TE_{CRS} or TE_{VRS} scores obtained from the first stage from 1 (Coelli et al., 2005; Farrell, 1957). The Tobit regression handles the truncated inefficiency estimates at 0 and 1 (Greene, 1994). The explanatory variables in the regression model, such as medical and epidemic prevention, death loss and selling price, reflect the factors that affect inefficiency (the explained variable). The following equation expresses the Tobit regression:

$$Ineff_{i,t} = \beta_0 + \beta_1 MEP_{i,t} + \beta_2 DL_{i,t} + \beta_3 SP_{i,t} + \varepsilon_{i,t} \tag{4}$$

Where $Ineff_{i,t}$ refers to the technical inefficiency score of each category of

livestock (i-th) ranging between 0 and 1 for t periods. It is obtained from the reciprocal of results of the CRS or VRS DEA assumption subtracted by one; β_0 is the intercept; β_1 to β_3 are coefficients estimated for individual independent variables; $MEP_{i,t}$ denotes medical and epidemic prevention expenses; $DL_{i,t}$ indicates livestock death loss during farming; $SP_{i,t}$ is the average selling price in Shaanxi wholesale markets; and $\varepsilon_{i,t}$ refers to the error term.

RESULTS AND DISCUSSION

Technical Efficiency

Figure 2 indicates that there have been fluctuations and downward trends in the technical efficiency scores of the Shaanxi livestock industry in recent years. The average technical efficiency scores over the decade reached 0.84 (CRS) and 0.92 (VRS); the scale efficiency score was 0.91. The technical efficiency scores under the VRS assumption, which range from 0.70 to 1, are higher than those under the CRS assumption (0.54–1). Additional details regarding these findings are presented in Table 2. Geometric means are calculated for each category to compare the technical efficiencies across various categories between 2010 and 2019. Dairy cow, goat, and cattle farms are operating at full technical efficiency (1.00) during the study period. However, inefficiencies are mainly observed in hog and layer farming practices, indicating that enhancing the performance of these two species could lead to an overall improvement in the technical efficiency of the livestock industry in Shaanxi.
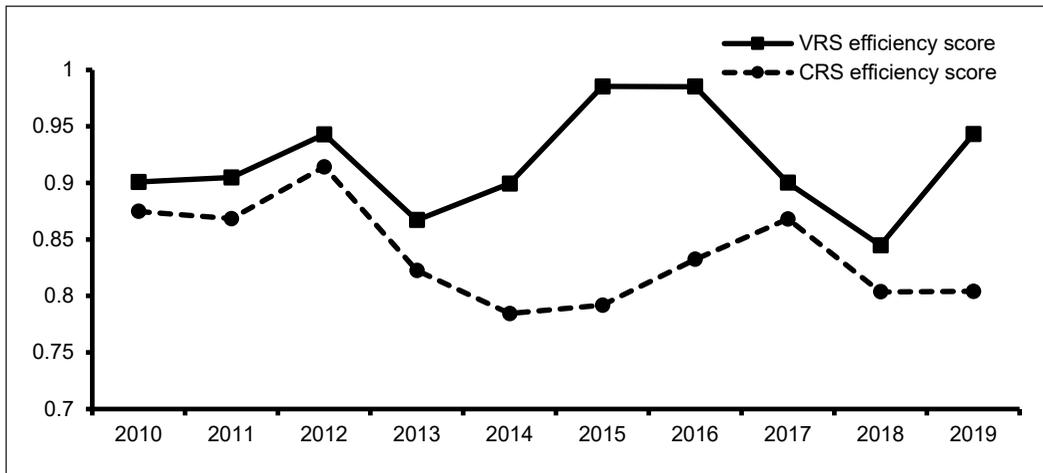
Figure 2. Technical efficiency score in the Shaanxi livestock industry
Source: Authors' work

Table 2
Technical efficiency and inefficiency in the Shaanxi livestock industry from 2010 to 2019

Products	Technical Inefficiency		SE	Technical Efficiency	
	VRS	CRS		VRS	CRS
Dairy Cow					1
Backyard	0	0	1	1	1
Medium-scale	0	0	1	1	1
Hog					0.68
Backyard	0.30	0.46	0.77	0.70	0.54
Small-scale	0.17	0.41	0.71	0.83	0.59
Medium-scale	0.13	0.32	0.78	0.87	0.68
Large-scale	0	0.04	0.96	1	0.96
Layer					0.89
Small-scale	0.17	0.21	0.96	0.83	0.79
Medium-scale	0	0	1	1	1
Goat	0	0	1	1	1
Cattle	0	0	1	1	1
Mean for all livestock	0.08	0.16	0.91	0.92	0.84

Notes. CRS: Constant Returns to Scale, SE: Scale Efficiency, VRS: Variable Returns to Scale
Source: Authors' work

From Table 2, it is evident that within the Shaanxi livestock industry, hog farms demonstrate the lowest efficiency scores (0.68), lower than those of cows (1.00), goats (1.00), and layers (0.89). Moreover, across all scales of hog farming and small-scale layer farming, the technical efficiency score obtained via VRS analysis is consistently higher than that obtained through CRS analysis, indicating the potential for

efficiency improvement through optimal scale management. Additionally, the results of the CRS and VRS assumptions provide evidence that the technical efficiency score increases with an expansion in the breeding scale.

In the case of layer, small-scale breeding demonstrates a technical efficiency score of 0.79, indicating a potential 21% increase in output. Meanwhile, medium-scale breeding operates at full technical efficiency. The findings are consistent with Zhong et al. (2021) and confirm the suitability of the government's five-year plan for livestock, which promotes farmers' increasing breeding scales in the layer sector.

The technical efficiency score of a hog (0.68) in Shaanxi is higher than the 0.58 reported by Tian et al. (2015) in Shaanxi and 11 other provinces but lower than the 0.75 measured by Somwaru et al. (2003) and 0.84 measured by Zhou et al. (2015). These findings are inconsistent with prior research and possibly attributable to divergent data sources where backyard farming is relatively more prominent in the surveyed data.

In Shaanxi, there has been a significant decrease in the proportion of households involved in backyard hog farming, which declined from 89% in 2010 to 43% in 2020 (National Bureau of Statistics of China, 2010, 2021). This shift has resulted in a corresponding increase in concentrated hog farming, encompassing small-, medium-, and large-scale farms, which has risen from 11% to 57% over the span of a decade (National Bureau of Statistics of China, 2010, 2021). The results indicate that hog farming across different scales does not demonstrate optimal technical efficiency. Yet, larger-scale operations tend to exhibit higher levels of technical efficiency. Backyard hog farms typically remain small-scale, with many still employing traditional labour-intensive feeding methods, such as using crop straw and swill for feed (Xiao et al., 2012). These practices contribute to low levels of technical proficiency and production efficiency. Due to the recent sharp rise in feed prices, backyard farmers have been compelled to reduce feed quantity, leading to an increased reliance on swill feeding. Regrettably, this tendency is anticipated to have adverse effects on hog health, including elevated mortality rates as well as lower farming efficiency (Xiao et al., 2012). Therefore, it is recommended that larger-scale breeding operations be adopted in the hog farming industry.

Factors Affecting Technical Inefficiency

The results of the Tobit regression analysis used to identify the factors affecting inefficiency in the Shaanxi livestock industry of each category are presented in Table 3. The expenses of medical and epidemic prevention have a negative impact on the overall technical inefficiency score, which implies that increasing prevention expenses might lead to a decrease in the inefficiency score, subsequently increasing the technical efficiency of Shaanxi's livestock industry. As discovered by Yan et al. (2023), compared to backyard farmers, large-scale hog farmers are more proactive in terms of biosecurity construction, aiming

Table 3
Results of the Tobit regression model on factors that affect technical inefficiency

Factors	Estimate		Std. Error		Prob.	
	VRS	CRS	VRS	CRS	VRS	CRS
Intercept	0.3678	0.8534	0.1857	0.1636	0.0477**	0.0000***
MEP	-0.0090	-0.0102	0.0021	0.0018	0.0000***	0.0000***
DL	0.0181	0.0136	0.0064	0.0054	0.0044***	0.0111**
SP	-0.0006	-0.0009	0.0002	0.0002	0.0068***	0.0000***

Note. *** and ** denote significance at the 1 per cent and 5 per cent levels, respectively
Source: Author's work

to reduce losses and enhance production efficiency. Medium- and large-scale farms exhibit enhanced financial capabilities, providing farmers with greater capital to invest in improved disease-preventive measures. In turn, it facilitates increased specialisation and reduces the susceptibility to disease outbreaks. However, increasing investment in disease prevention necessitates transitioning farms into medium- and large-scale operations. Therefore, policies that support the farm consolidation of backyard and small-scale farms into larger ones and facilitate the growth of medium-scale farms are necessary to increase production and technical efficiency (Crop Farming Management Office of Shaanxi Province, 2022). Consolidating and expanding farms would allow larger farms to allocate more resources to preventing outbreaks like avian influenza (Wang, Zhou, et al., 2021).

Additionally, it is observed that death loss positively correlates with technical inefficiency, consistent with findings by Jo et al. (2021). Death loss results in a decrease in inefficiency scores, subsequently enhancing the technical efficiency of the livestock industry in Shaanxi, particularly in hog and layer farming. The breeding environment for hogs and layers is typically less sanitary, which can more easily result in death loss. An elevated death loss is associated with a decline in technical efficiency, resulting in decreased output production and diminished farmer income. This trend prompts the departure of numerous backyard and small-scale farmers from the livestock industry (Wang, Han, et al., 2021).

On the other hand, the average selling price of livestock products shows a negative impact on technical inefficiency in the Shaanxi livestock industry, suggesting that a higher selling price might lead to a decrease in the inefficiency score and improve the technical efficiency of Shaanxi's livestock industry. The market selling price and revenue of livestock products are key factors determining the duration of livestock feeding days. Farmers can adjust the length of feeding days to manage their production costs and maintain their income during price fluctuations. When farmers face an increase in the prices of livestock products, they may choose to shorten the duration of feeding days. Farmers may increase feeding days when food prices fall to maintain their

income. It illustrates the relationship between higher prices and increased efficiency, as well as the inverse association between lower prices and declining efficiency due to prolonged feeding periods. It is not a healthy practice for farmers and consumers. Therefore, it is imperative for the government to implement measures to stabilise prices, enabling farmers to maintain a sustainable livestock feeding base.

CONCLUSION

This study employs Data Envelopment Analysis (DEA) to investigate technical efficiency. It utilises Tobit regression to examine the factors affecting inefficiency across various scales and species of farming operations in the Shaanxi livestock industry from 2010 to 2019.

The main findings of this study are as follows. Firstly, the livestock industry in Shaanxi exhibits inefficiencies in current farming practices. Goat, dairy cow and beef farming exhibit full technical efficiency. However, both hog and layer farming practices show the presence of technological inefficiency and scale inefficiency. Notably, the results indicate that the technical efficiency scores of hog and layer farming increase as the breeding scale increases. This study also employed the Tobit regression to explore the potential effect between technical inefficiency and three influencing factors. The findings indicated that increasing medical and epidemic prevention expenses, reducing death loss and raising selling prices are crucial improvements that can enhance industry performance.

From a policy perspective, the findings of this study provide valuable insights for policymakers. Firstly, the breeding of dairy cows, goats and cattle in Shaanxi has reached full technical efficiency, suggesting that the government should motivate more farmers to participate in these farming activities. Secondly, the technical efficiency of hog and layer farming increases with the expansion of the production scale. Therefore, the Shaanxi government could actively promote the establishment of livestock production cooperatives and incentivise farmers to participate by combining adjacent breeding facilities. This approach would enable farms to accumulate funds to enhance epidemic prevention measures. Additionally, reducing death losses could enhance technical efficiency within the livestock industry. Hence, the Shaanxi provincial authority could implement subsidies and encourage farmers to reduce stocking density. Lastly, raising the selling price of livestock products can enhance the technical efficiency of the industry and stimulate farmers to increase livestock numbers and achieve higher profits. However, it is important to note that this higher selling price is not beneficial to consumers. Therefore, it is advisable for the Shaanxi Provincial Government to establish market regulatory mechanisms, such as price support and reserve systems, to mitigate the effects of price volatility. Implementing these measures can help achieve the government's goal of boosting Shaanxi's livestock production.

However, this study's limitation is its oversight of the livestock industry's environmental and sustainability impact. Pursuing higher output through excessive resource exploitation may adversely affect Shaanxi's ecosystems and natural resources. Therefore, it is recommended that future studies focus on examining the environmental efficiency of Shaanxi's livestock industry. By quantifying the relationship between resource utilisation and environmental impacts, measures can be formulated to promote the sustainable development of Shaanxi's livestock industry.

ACKNOWLEDGEMENTS

REFERENCES

Aigner, D. J., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*(1), 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

AskCI Consulting. (2022, January 22). *China's meat imports analysis 2021*. AskCI. https://www.askci.com/news/data/maoyi/20220124/0843371733022.shtml

Banker, R. D., Charnes, A., & Cooper, W. E. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, *30*(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Cadell, C. (2020, November 15). China's sweeping COVID controls put frozen food importers on ice. *Reuters*. https://www.reuters.com/article/us-health-coronavirus-china-frozenfood-idUSKBN27T0YK

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). Springer.

Crop Farming Management Office of Shaanxi Province. (2022). *Interpretation of the development plan for fruit, vegetable, and tea industries.* Department of Agriculture and Rural Affairs of Shaanxi Province. http://nynct.shaanxi.gov.cn/www/bm7831/20221206/9811351.html

Dogan, N., Kaygisiz, F., & Altinel, A. (2018). Technical and economic efficiency of laying hen farms in Konya, Turkey. *Brazilian Journal of Poultry Science*, *20*(2), 263–272. https://doi.org/10.1590/1806-9061-2017-0649

Food and Agriculture Organization. (2021). *World food and agriculture statistical yearbook 2021*. https://www.fao.org/policy-support/tools-and-publications/resources-details/zh/c/1458575/

Food and Agriculture Organization. (2022a). *FAOSTAT-Crop and livestock products*. https://www.fao.org/faostat/en/#data/QCL

Food and Agriculture Organization. (2022b). Trade of agricultural commodities 2005–2021. https://openknowledge.fao.org/server/api/core/bitstreams/75eaf30a-b1b0-4e7b-8e21-d8970ff5833b/content#:~:text=The%20share%20of%20agricultural%20products,to%207.9%20percent%20in%202021.

Färe, R., & Lovell, C. a. K. (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, *19*(1), 150–162. https://doi.org/10.1016/0022-0531(78)90060-1

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical*

Society. Series a (General), *120*(3), 253–290. https://doi.org/10.2307/2343100

Fraser, I., & Cordina, D. (1999). An application of data envelopment analysis to irrigated dairy farms in Northern Victoria, Australia. *Agricultural Systems*, *59*(3), 267–282. https://doi.org/10.1016/s0308-521x(99)00009-8

Galluzzo, N. (2019). An analysis of technical efficiency in Icelandic dairy and sheep farms. *Studies in Agricultural Economics*, *121*(3), 144–150. https://doi.org/10.7896/j.1916

Greene, W. H. (1994). Econometric analysis. *Journal of the American Statistical Association*, *89*(428), 1567–1569. https://doi.org/10.2307/2291031

Hansson, H., & Öhlmér, B. (2008). The effect of operational managerial practices on economic, technical and allocative efficiency at Swedish dairy farms. *Livestock Science*, *118*(1–2), 34–43. https://doi.org/10.1016/j.livsci.2008.01.013

Heshmati, A., Kumbhakar, S. C., & Hjalmarsson, L. (1995). Efficiency of the Swedish pork industry: A farm level study using rotating panel data 1976–1988. *European Journal of Operational Research*, *80*(3), 519–533. https://doi.org/10.1016/0377-2217(94)00135-y

İkikat, T. E., Ağır, H., & Aydoğan, İ. (2020). Evaluating technical efficiency of hair goat farms in Turkey: The case of Mersin Province. *Tropical Animal Health and Production*, *52*(6), 3707–3712. https://doi.org/10.1007/s11250-020-02407-2

Jo, H., Nasrullah, M., Jiang, B., Li, X., & Bao, J. (2021). A survey of broiler farmers' perceptions of animal welfare and their technical efficiency: A case study in Northeast China. *Journal of Applied Animal Welfare Science*, *25*(3), 275–286. https://doi.org/10.1080/10888705.2021.1912605

Kuhn, L., Baležentis, T., Hou, L., & Wang, D. (2020). Technical and environmental efficiency of livestock farms in China: A slacks-based DEA approach. *China Economic Review*, *62*, Article 101213. https://doi.org/10.1016/j.chieco.2018.08.009

Lansink, A. O., & Reinhard, S. (2004). Investigating technical efficiency and potential technological change in Dutch pig farming. *Agricultural Systems*, *79*(3), 353–367. https://doi.org/10.1016/s0308-521x(03)00091-x

Liu, W., Shankar, S., & Li, L. (2021). Is specialization a strategy to improve farm efficiency in Northwest China? *Review of Development Economics*, *25*(3), 1695–1710. https://doi.org/10.1111/rode.12782

Luik-Lindsaar, H., Põldaru, R., & Roots, J. (2019). Estonian dairy farms' technical efficiency and factors predicting it. *Agronomy Research*, *17*(2), 593–607. https://doi.org/10.15159/ar.19.067

Meeusen, W., & Van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, *18*(2), 435–444. https://doi.org/10.2307/2525757

Mohd Idris, N. D., Chamhuri, S., & Talib, B. (2013). Determinants of technical efficiency on pineapple farming. *American Journal of Applied Sciences*, *10*(4), 426–432. https://doi.org/10.3844/ajassp.2013.426.432

Mugera, A. W., & Featherstone, A. M. (2008). Backyard hog production efficiency: Evidence from the Philippines. *Asian Economic Journal*. https://doi.org/10.1111/j.1467-8381.2008.00278.x

National Bureau of Statistics of China. (2010). *China statistical yearbook 2010*. https://www.stats.gov.cn/sj/ndsj/2010/indexeh.htm

National Bureau of Statistics of China. (2021). *China statistical yearbook 2021*. https://www.stats.gov.cn/sj/ndsj/2021/indexeh.htm

National Bureau of Statistics of China. (2023). *Statistical communiqué of the People's Republic of China on the 2022 national economic and social*

development. http://www.stats.gov.cn/english/PressRelease/202302/t20230227_1918979.html

Patton, D. (2022, February 11). China targets 15% increase in meat output by 2025. *Nasdaq.* https://www.zawya.com/en/business/china-targets-15-increase-in-meat-output-by-2025-hwsq83x2

Price Department of the National Development and Reform Commission (2019). *Compilation of national agricultural product cost and benefit data-2019.* China Statistics Press. https://www.stats.gov.cn/zs/tjwh/tjkw/tjzl/202302/t20230215_1907941.html

Reinhard, S., Lovell, C. A. K., & Thijssen, G. (1999). Econometric estimation of technical and environmental efficiency: An application to Dutch dairy farms. *American Journal of Agricultural Economics*, *81*(1), 44–60. https://doi.org/10.2307/1244449

Shaanxi Government. (2008). *Special plan for key areas of rural economic development during the 10th Five-Year plan for national economic and social development in Shaanxi province.* http://www.shaanxi.gov.cn/zfxxgk/fdzdgknr/ghxx/swgh/zxgh_2451/200801/t20080101_1661543.html

Shaanxi Government. (2022). *Modern fruit industry development plan for the 14th Five-Year plan in Shaanxi province.* http://www.shaanxi.gov.cn/xw/ldx/bm/202210/t20221027_2257223_wap.html

Shaanxi Provincial Department of Agriculture. (2018). *Modern agriculture development plan for the 13th Five-Year plan in Shaanxi province (2016-2020).* http://nynct.shaanxi.gov.cn/jgw/home/index/detail/catId/738/articleId/5200.html

Shaanxi Provincial Department of Agriculture. (2022). *Shaanxi Provincial Department of Agriculture and Rural Affairs on the issuance of the "14th Five-Year plan" for the development of animal husbandry and veterinary medicine in Shaanxi province.* http://nynct.shaanxi.gov.cn/www/snynctwj/20220428/9792947.html

Shaanxi Statistics Bureau. (2015). *Transformation of agricultural development mode and enhancement of product supply capacity.* http://tjj.shaanxi.gov.cn/sy/ztzl/zgcyksx/201510/t20151019_1615133.html

Sharma, K. R., Leung, P., & Zaleski, H. M. (1997). Productive efficiency of the swine industry in Hawaii: Stochastic frontier vs. Data envelopment analysis. *Journal of Productivity Analysis*, *8*(4), 447–459. https://doi.org/10.1023/a:1007744327504

Shi, Z. Z., Wang, M. L., & Hu, X. D. (2015). A study on beef market partial equilibrium model in China. *Journal of China Agricultural University, 20,* 278–290.

Siafakas, S., Tsiplakou, E., Kotsarinis, M., Tsiboukas, K., & Zervas, G. (2019). Identification of efficient dairy farms in Greece based on home grown feedstuffs, using the Data Envelopment Analysis method. *Livestock Science, 222,* 14–20. https://doi.org/10.1016/j.livsci.2019.02.008

Soh, N. C., Samsuddin, N. S., Ismail, M. R., & Habibullah, M. S. (2021). Technical efficiency of commercial stingless bee honey production in Peninsular Malaysia. *Pertanika Journal of Social Science and Humanities*, *29*(2). https://doi.org/10.47836/pjssh.29.2.03

Somwaru, A., Zhang, X., & Tuan, F. C. (2003). *China's hog production structure and efficiency* [Paper Presentation]. American Agricultural Economics Association Annual Meeting 2003, Montreal, Canada. https://doi.org/10.22004/ag.econ.22003

Terry, S. A., Basarab, J. A., Guan, L. L., & McAllister, T. A. (2021). Strategies to improve the efficiency of beef cattle production. *Canadian Journal of Animal Science*, *101*(1), 1–19. https://doi.org/10.1139/cjas-2020-0022

Tian, X., Sun, F., & Zhou, Y. (2015). Technical efficiency and its determinants in China's hog production. *Journal of Integrative Agriculture*, *14*(6), 1057–1068. https://doi.org/10.1016/s2095-3119(14)60989-8

Wang, G., Zhao, C., Shen, Y., & Yin, N. (2021). Estimation of cost efficiency of fattening pigs, sows, and piglets using SFA approach analysis: Evidence from China. *PLOS ONE*, *16*(12), Article e0261240. https://doi.org/10.1371/journal.pone.0261240

Wang, Y., Han, Y., Han, Y., & Li, W. (2021b). Does the grassland ecological compensation policy improve the herders' breeding technical efficiency in China? Based on the parallel mediation effect model. *PLOS ONE*, *16*(4), Article e0249990. https://doi.org/10.1371/journal.pone.0249990

Xiao, H., Wang, J., Oxley, L., & Ma, H. (2012). The evolution of hog production and potential sources for future growth in China. *Food Policy*, *37*(4), 366–377. https://doi.org/10.1016/j.foodpol.2012.02.002

Yan, Z., Wang, M., Li, X., & Jiang, H. (2023). Impact of African swine fever epidemic on the cost intensity of pork production in China. *Agriculture*, *13*(2), Article 497. https://doi.org/10.3390/agriculture13020497

Yang, C. C., Hsiao, C. K., & Yu, M. M. (2008). Technical efficiency and impact of environmental regulations in farrow-to-finish swine production in Taiwan. *Agricultural Economics*, *39*(1), 51–61. https://doi.org/10.1111/j.1574-0862.2008.00314.x

Zhang, Z., Zhang, Y., Li, F., Yang, H., Yuan, Y., & Yuan, X. (2017). Economic efficiency of small-scale tilapia farms in Guangxi, China. *Aquaculture Economics & Management*, *21*(2), 283–294. https://doi.org/10.1080/13657305.2016.1180644

Zhong, S., Li, J., Chen, X., & Wen, H. (2021). Research on the green total factor productivity of laying hens in China. *Journal of Cleaner Production*, *315*, Article 128150. https://doi.org/10.1016/j.jclepro.2021.128150

Zhou, Y., Zhang, X., Tian, X., Geng, X., Zhang, P., & Yan, B. (2015). Technical and environmental efficiency of hog production in China – A stochastic frontier production function analysis. *Journal of Integrative Agriculture*, *14*(6), 1069–1080. https://doi.org/10.1016/s2095-3119(14)60990-4

Zhu, J. (2009). *Quantitative models for performance evaluation and benchmarking: Data envelopment analysis with spreadsheets* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-85982-8

Gender Quota of Women Election Supervisors in The Implementation of Indonesian Election Law

Ricky Santoso Muharam[1,2]*, Ari Pradhanawati[1], Fitriyah[1] and Supratiwi[1]

[1]*Department of Political and Governance Studies, Faculty of Social and Political Sciences Universitas Diponegoro, Indonesia*
[2]*Sekolah Tinggi Pariwisata Ambarrukmo Yogyakarta, Indonesia*

ABSTRACT

Election supervisors are part of the election organizers in Indonesia, regulated by Law Number 7 of 2017 concerning general elections. The regulation on women's representation in the Election Supervisory Board (Bawaslu) is contained in Article 92, paragraph 11, as much as 30% of women's representation. This research uses a qualitative approach using archival studies on various elections and affirmative regulations. The findings in this study are the results of women's quotas spread across 514 regencies/cities. Male representatives numbered 1633 (85%) and women 274 (14%) out of 1907 commissioner seats contested. This study opens further discussion for the Indonesian government to amend the election law regarding women's representation in Bawaslu, as a regulatory sanction is required if the quota is not fulfilled.

Keywords: Elections, election supervisor, gender quota, Indonesia, women's representation

INTRODUCTION

The discussion on women's gender quotas in Indonesia has become a study for political scientists in Indonesia. As one of the countries using a democratic system, the issue of women's gender quotas in Indonesian politics has not had a significant impact. This problem still provides limited space for women to participate in the political system in Indonesia, which is still thick of its patriarchal culture. Concerning human rights and gender equality standards,

all institutions and political actors in the OSCE (Organization for Security and Co-operation in Europe) region have committed to respecting these standards, including countries that have committed to doing so. This act is conducted based on universal human rights instruments, which uphold equal rights for women and men, encouraging women's participation. The international principles governing the operation of democratic institutions also promote women's participation (Office for Democratic Institutions and Human Rights [ODIHR], 2014). The feminist movement encourages women to be directly involved in politics in order to break free from the culture of patriarchy that has been detrimental to women (Cole & Sabik, 2010). Including Indonesia, fulfilling women's representation in politics highlights the representation of Indonesian women in parliament, political parties, and regional head officials. Attention to women as election organizers in Indonesia still receives little attention from political scientists.

Regarding the gender gap in political empowerment, Indonesia is ranked 81st in 2023. Previously, Indonesia's position was ranked 90th in 2002 (World Economic Forum, 2023). As for the Asia and Pacific region, Indonesia is ranked ninth below New Zealand, the Philippines, Australia, Singapore, Laos, Vietnam, Thailand, and Mongolia (Muharam. et al., 2023). As a commitment to gender equality, the Indonesian government has made various rules regulating gender quotas, one of which has been handled in Law Number 8 of 2012 concerning the General Election of Members of the DPR, DPD, and DPRD, which governs the obligation of women's participation in general elections, which expressly regulates the right of women to participate in the political process regardless of their gender (Republik Indonesia, 2012). Moreover, many laws regulate women's involvement in the political process in Indonesia. In addition, the implementation of the zipper system in Indonesia has been in effect since the 2009 election (Prastiwi, 2018). Rules regarding women's representation have also been in Law Number 2 of 2011 concerning political parties in Article 2 Paragraph 5, which regulates the management of political parties at the central level. There should be 30% female representation (Republik Indonesia, 2011). The existence of regulations regarding women's quotas certainly requires a commitment to implementing rules that have been in force and have permanent legal force. This commitment requires solid moral values within the laws of women's political participation (Brisbane et al., 2023). This morality is responsible for strengthening women's political participation, especially in Indonesia's democratic system. Much research has been done on women's empowerment in politics and women's involvement in political participation, but there was less reviewed research on women's representation in electoral institutions. For instance, the role of women in the Election Supervisory Board (Bawaslu) has not been widely studied and considered in this regard.

So far, the issue of women's representation in Indonesia has only examined women's representation in parliament, in parties, and women's participation in politics (Muharam & Prasetyo, 2021). Although not much has been discussed, women's representation in organizing institutions is also essential. The involvement of women in election supervisory institutions will add a different color. Moreover, Law Number 7 of 2017 concerning General Elections has specifically regulated fulfillment of the women's quota in Bawaslu (Muharam et al., 2023). Affirmative policy regulations for women in the reform era have been introduced previously in Law Number 12 of 2003 concerning the General Election of Members of the People's Representative Council. DPRD and DPD Introduce Minimum Quota Limit for Women Politicians (Vasandani et al., 2022). One of the commitments to build gender equality is Indonesia's participation in activities (Joint Declaration on Trade and Women's Economic Empowerment) in 2017 held by Buenos Aires. This forum discusses the issue of gender equality. However, Indonesia does not yet have a gender equality and justice law that can strengthen Indonesian legal institutions in regulating gender issues (Kusumawardhana & Abbas, 2018). Of course, this regulation is vital in fighting for gender equality and justice, which is considered to have not been fully implemented. To realize democracy for the better, the Indonesian government has specifically regulated the concept of gender rules in Law Number 7 of 2017

concerning General Elections. Some articles have regulated women's representation in election-organizing institutions in Indonesia. Like the General Elections Commission regulated in Article 10, paragraph 7, the gender quota of the Election Supervisory Agency (Bawaslu) is regulated in Article 92, paragraph 11. However, the requirement for 30 percent female representation in the Election Organizing Agency (DKPP) is not regulated (Republik Indonesia, 2017). Certainly, the purpose of this rule is the development of a creative personality and a sense of togetherness with idealistic concepts. This concept contains the understanding of citizenship in the context of a women-friendly democracy based on the various experiences of both women and men (Lovenduski, 2019). The existence of this regulation actually answers the problem of gender inequality in election-organizing institutions, offering the possibility of affirmative action as a solution. Figure 1 is considered when observing how patriarchal culture organizes elections.

According to the data above, women's representation in the world of election organizers in KPU and Bawaslu still shows gender inequality in which the 30% quota for women has not been met. From the public votes of election organizer members, women's representation is only about 20% of election organizers. The need for efforts to encourage women's representation in various sectors of strategic public office is still on the common agenda (Pusat Kajian Politik Universitas Indonesia [Puskapol], 2021). The empirical inequality in women's representation in

	General Election Commission (KPU) Republic Indonesia	Election Supervisory Agency (Bawaslu) Republic Indonesia	General Election Commission (KPU) Province	Election Supervisory Agency (Bawaslu) Province	General Election Commission (KPU) Regency/City	Election Supervisory Agency (Bawaslu) Regency/City	Total
▮ Female percentage	14%	20%	21%	20%	17%	14%	44%
▮ Male percentage	86%	80%	79%	80%	83%	84%	56%
▮ Female	1	1	39	38	441	315	2119
▮ Male	6	4	146	150	2101	1599	2722

Figure 1. Comparison of the number of women's representation in the KPU and Bawaslu in the 2019 Election
Source: Authors' work

election organizing institutions shows the lack of women as members of the KPU and Bawaslu at various levels. Considering the ranks of Bawaslu commissioners, there is only one female representative out of the total number of Bawaslu commissioners, which is five people. Even in some regions, some still do not have female commissioners (Puskapol, 2021).

With further observation, there are obstacles to increasing women's representation in election administration due to the number of registrations for commissioner candidates, in which the number of male applicants is much larger than women. Meanwhile, from the selection process, there are serious problems related to the perspective of unequal representation of gender justice in the selection team and the

House of Representatives of the Republic of Indonesia. The push for technical regulations that are not strong enough to encourage the enforcement of gender justice is also believed to be one of the causes of inequality. This problem makes the position of women in political dynamics not yet able to run well. There is still a view that the existence of women has not maximally appeared in public because they are still shackled by the shadow of patriarchal culture still thick in the political system (Nurcahyo, 2016). Of course, this mentality and view must be eroded so that women's representation can be fulfilled appropriately, especially in the process of holding elections, where women's representation has not been given much attention. Women's issues are still focused on the legislative realm and political

parties. Therefore, this study focuses more on women's representation as election supervisors in Indonesia.

The purpose of this study is to provide an overview of the implementation of gender quotas in election law regulations in Indonesia and find out the extent of election law practice regarding women's representation in election supervisory institutions in Indonesia by looking at and considering 30% women's representation to serve as members of election supervisory commissioners in Indonesia.

This research contributes to understanding the representation of women as election supervisors in Indonesia in implementing the rule of law that applies in Indonesia. Gender quota is a solution to implementing a 30% quota as an election supervisor in Indonesia. This study will provide an overview of the application of legal rules on women's representation in Indonesia, especially as election organizers, which political science researchers in Indonesia have not widely studied.

LITERATURE REVIEW

Gender Quota Theory

The discourse on women's representation in the political system is an earnest study for political scientists. The perspective of representation in politics is considered one of the successes that need to exist in a democratic system. The understanding of women as individuals who only take care of domestic affairs and the influence of patriarchal culture in some countries make women considered unnecessary to play a role in determining policies and political decisions. According to Andrew Heywood's book *Political Theory: An Introduction*, representation suggests that an individual or group somehow represents a broader group of people. The relationship between two different entities is recognized by political representation (Heywood, 2004). Giving women the freedom to participate in politics as direct participants and compete for political office in modern democracies does allow women to earn their place in women's affirmative quotas.

In some countries with developed democracies, women are allowed to participate in the race for important government posts. Gender quotas shape women's political experiences and representations, and this will change normative perceptions so that female candidates can compete with male candidates (Deininger et al., 2015). Brian Turnbull explained that gender quotas did not allow female politicians to grow and develop in competitive gender competition. Suppose political parties continue to perceive women as a "risk" and female candidates competing outside the gender quota are unsuccessful in elections. In this case, the quota would be a top-and-bottom limit that limits women's participation (Turnbull, 2019). In other words, there is a gender quota as part of the formula for fulfilling representation in political participation. Klaus Deininger, Hari Krishnan Nagarajan, and Song Qing Jin even justified this statement by stating that organizational barriers to cultural norms, political discrimination, and policymakers' efforts to increase participation share

affected the political balance and long-term consequences (Deininger et al., 2015).

In addition, the influence of gender norms derived from religious and cultural traditions influences gender inequality (Boo, 2021). Of course, this can affect the position of women in seizing strategic positions in politics due to a paradigm attached to women. Hence, gender equality can be achieved since women have essential roles in economic and community development (Ang & Lai, 2023). Eventually, the gender-based quota, one of the solutions for women's empowerment used in more than a hundred countries, has become popular to increase the number of women elected in politics.

METHODS

This research was qualitative research, which is research using a natural background with the intention of interpreting phenomena that occur. It was carried out using various existing methods (Moleong, 2012). Nevertheless, John W. Creswell explained that qualitative research uses word-using methods to explore and understand the meaning processed by researchers derived from social problems (Creswell, 2014). Archival research studies use data records stored in physical and digital forms (Shaughnessy et al., 2012). The records are then analyzed to obtain conclusions about behaviors, attitudes and beliefs (Shaughnessy et al., 2012). The archival research methods in this research are based on various election regulations, such as General Election Commission Regulations, Election Supervisory Body Regulations, and statutory regulations regarding women's representation or affirmative action. The data processing process is in Figure 2.
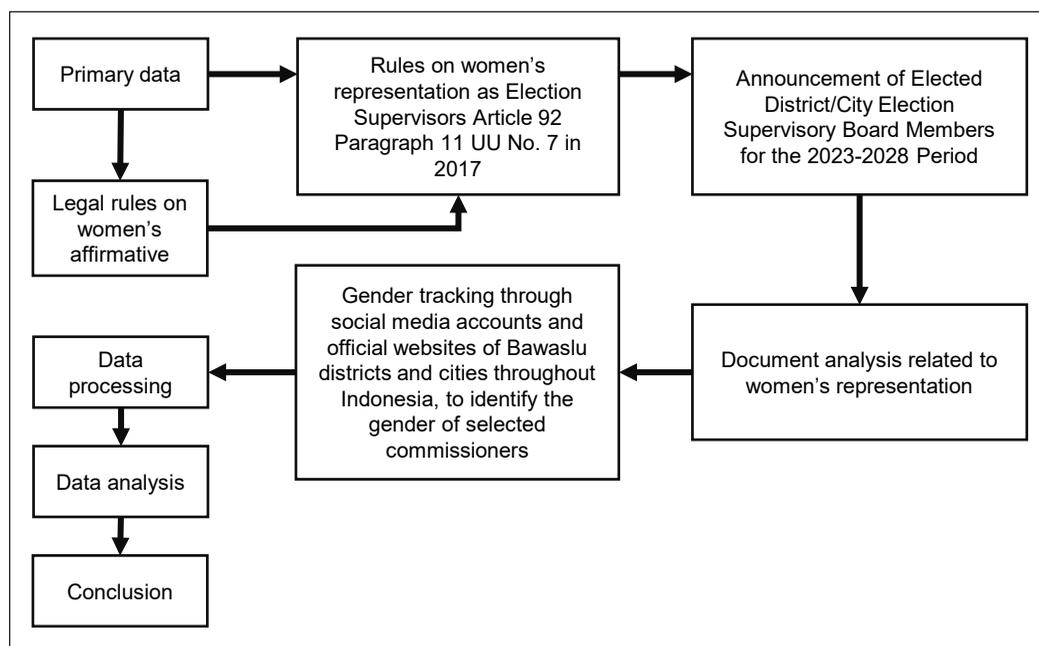
Figure 2. Research data processing (*Source*: Authors' work)

The theory used in this study is the gender quota theory, which questions the extent to which women's quotas can empower women, which is the goal. Gender quotas have actually proven effective in increasing women's representation in political positions such as in parliament in political parties as long as they are clearly enforceable and there are sanctions for those who do not comply with these rules (Childs & Krook, 2008; Dahlerup & Freidenvall, 2005). Primary data sources came from the general election stages in regencies and cities in Indonesia to the final stage of announcing candidates for Bawaslu members by making indirect observations from the decision results of the Bawaslu of the Republic of Indonesia.

Secondary data was obtained from various literature, such as journals, books, and articles discussing the recruitment process of prospective members of Bawaslu in regencies and cities of Indonesia. The quota of women's representation from 514 regencies/cities in Indonesia was then traced based on checking the data obtained and browsing social media accounts and websites owned by Bawaslu at regency and city levels. After the data was collected and cross-checked, the data obtained was then analyzed.

DISCUSSION

Affirmative Policy for Women in Indonesia

Gender quota enforcement will be able to run if the policies that have been regulated can be implemented properly and correctly. The importance of seriousness in carrying out practices in fulfilling formative quotas is not only a clear rule in regulations but also the need for mechanisms and applications in the field. Apart from regulations, of course, the success of fulfilling gender quotas requires the struggle of female politicians to be able to participate in the recruitment process for open political positions as a form of resistance and challenge to the hierarchy in the political system that has dominated (Phillips, 2019). This participation involves individual women being engaged in various activities (Hannah et al., 2021). Patriarchal cultural factors that do not compensate for access to affirmative action for women, such as quotas, contribute to women's underrepresentation in politics (Abu-Zayd et al., 2002). It hampered the opportunity for female politicians who have the ability or capacity to enter the political system. Female politicians struggle not only with a strong patriarchal culture but also with rules and regulations that are friendly, not gender-biased, and appropriately implemented, not just a matter of diction in paragraphs or statutory regulations. However, the commitment to implementing policies for affirmative women comes from applicable legal products. Regarding affirmative policies in Indonesia to increase women's representation, please see Table 1 as follows:

Based on Table 1 above, of course, the rules regarding women's affirmative quotas have been regulated in various laws and regulations, one of which is Law Number 7 of 2017 concerning general elections, which has held women's representation using a *zipper system*, rules for party

Table 1

Indonesia's affirmative policy in increasing women's representation

Regulation on Women's Affirmative Action	Rules in Law Number 7 of 2017
• Convention on Women's Political Rights (Law Number 68/1985) Women have the right to vote and be elected in elections.	• The policy of political parties is to place a minimum quota of 30% women in DCTs and political party officials.
• Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW) (Law Number 7/1964) Eliminating all forms of discrimination against women.	• A Zipper System does not mean alternating intervals. However, it means that one in three candidates is at least one woman.
• Law Number 2 of 2011 concerning Political Parties requires 30% representation of women as founders, administrators, and members of political parties.	• Placement of the permanent candidate list (DCT): For every 3 candidates, 1 female candidate was placed in order 1, 2, or 3. Not only the numbers 3 and 6 but also others.
• Law Number 7/2017 on women's representation in the management of central political parties and lists of candidates for councilors: Every three people are women.	• Sanctions against parties who violate the quota provisions above are regulated in KPU regulations.

Source: Asy'ari (2023)

management, legislature, and elections. This rule is an attempt to deal with discrimination when women enter the realm of masculine professions (Ahmad et al., 2019). The existence of women's affirmative policies as a basis for protecting women's rights, as well as minimizing negative perceptions of female leaders originating from women's gender roles, stems from the fact that these policies are not commensurate with society's expectations of leaders. When audiences see women as actual or potential residents of leadership roles, they can experience prejudice due to an inconsistency between the essential communal qualities they associate with women and the agent-dominant qualities they consider necessary to succeed as a leader. Therefore, people tend to have different beliefs about female and male leaders, as well as similar beliefs (Eagly & Karau, 2002). This statement is reinforced by (Primagita & Riantoputra, 2019). There is a convention that stipulates

that the role of women is only to take care of household affairs, where women have an essential role, making it difficult for women to develop their identity as leaders.

There is a need for ideas to help women meet their needs as well as enhance and contribute to gender-neutral career development (Diogo et al., 2021). The existence of this affirmative policy can undoubtedly provide benefits for prospective women who are taking the portion of the gender quota that has been prepared. However, it does not make women feel secure in winning public office without any effort made by them. Existing regulations in Indonesia certainly need to be used to provide opportunities to meet the target of women's representation in politics. Various affirmative policies for women in Indonesian regulations are listed in Table 2.

Table 2 shows that the affirmative policy of gender quota representation in the electoral system in Indonesia has provided

Table 2

Affirmative policy for women's representation in politics in Indonesia

General Election Law Number 12/2003	Law on Political Parties Number 2/2008	General Election Law Number 10/2008 and Law No. 8/2012	Law of Political Parties Number 2/2011	General Election Law Number 7/2017
Appeal for political parties participating in the election to nominate 30% of women as legislators	They are involving 30% of women as party founders.	The list of candidates contains 30% women.	Involving 30% of women as party founders	Affecting 30% of women as election organizers, both KPU, Bawaslu
	Political party management contains 30% of women	Every three names on the candidate list have at least one female candidate	The control of political parties includes 30% women	The management of political parties contains 30% women.
			Recruitment of legislative candidates involves 30% female representation.	Registration of candidates for legislative membership must involve 30% female representation.

Source: Asy'ari (2023)

space for women to be involved in political competition. The existence of apparent regulations for women in Indonesia can be used as an opportunity to fill strategic political positions and compete with male politicians.

Affirmative Quota Implementation in Election Supervisors

The Indonesian state explicitly regulates women's representation in politics, and as many as 30% of female representatives are involved in politics. Both the management of political parties, legislative candidates, and election organizers at the General Election Commission (KPU) and the Election Supervisory Agency (known as Bawaslu) at the central, provincial, and regency/city levels have been regulated in Law Number 7 of 2017 concerning General Elections, which

is still valid today. The implementation of women's representation that has been handled in the election law is significant, considering that Indonesia is one of the countries that implements the Pancasila democratic system and accommodates all levels of society regardless of differences due to religion, ethnicity, and gender, positioning the concept of equality in the political system in Indonesia. Implementing this rule of law is undoubtedly crucial for women to have the opportunity to fulfill women's representation in women's supervisory institutions in Indonesia. The very long selection process in the race for the commissioner seat as election supervisor begins with the registration process, administrative selection of written tests, psychological tests, medical tests, interviews, and due diligence. The process

that goes through each stage uses a knockout system where the selection team of the Election Supervisory Board in each province will be tasked with selecting candidates for election supervisory commissioners. When entering the interview and due diligence process, each required quota sends twice the required quota requirement. For example, if 5 seats are needed in one city or regency, ten candidates are included in the due diligence and interview. In the end, 5 people were selected, determined from the results of due diligence and interviews conducted at the center.

Overall, the seats contested in the election supervisory body amounted to 1907 seats from 514 regencies/cities in 38 provinces in Indonesia. The implementation of gender quotas in politics in Indonesia, as election organizers and election supervisors, has not been explicitly examined. The lack of attention to women's representation in Indonesian election supervisory institutions ultimately means that the position of women in contesting for political positions in election supervisory institutions has not met 30% of women's representation quota regarding the rules in Law Number 7 of 2017. According to the rule, the need for women's representation in election supervisory institutions in Indonesia has been clearly regulated. However, implementing the rule in the recruitment process of election supervisors has not run well and optimally. More details can be seen in Table 3.

Based on Figure 3 above, the low representation of women in election supervisory institutions is undoubtedly a terrible record of the implementation of women's representation in election supervisory institutions that have not complied with applicable laws and regulations. This finding shows that the factors causing inequality and low representation of women in election organizers are weak regulatory guarantees, limited access to recruitment information, incomprehensible understanding of the selection process mechanism, and a political environment that is gender insensitive, including the unequal perspective of gender justice of the selection team. Political support remains weak as cultural barriers hinder women's representation in election supervisory institutions. Other factors of low representation of women in election organizing institutions are (1) the low motivation and confidence level, inadequate access to information, and suboptimal support for women, (2) regulatory guarantees are still not supportive, (3) political support is still low, and (4) uneven gender perspectives (Puskapol, 2021). Women's representation in election supervisory institutions is not only an individual interest but also a form of expression in the policy consideration system and a symbol of the system to reflect women's identity (Alhuzail, 2021). The issue of women's representation in the administration of elections, apart from the patriarchal culture that is still thick in the political world, is the existing regulations that have not been implemented into the rules of democratic life, especially in fulfilling women's quotas in election administration.

Table 3

Women's representation as Members of Bawaslu Commissioners Regency and City throughout Indonesia for the 2023-2028 period

Province	City/ Regency	Male	Female	Percentage	Percentage	Total Gender
				M	F	
Aceh	23	66	10	87	13	76
Bali	9	26	8	76	24	34
Bangka Belitung	7	19	2	90	10	21
Banten	8	34	4	89	11	38
Bengkulu	10	28	2	93	7	30
Special Region of Yogyakarta	5	16	3	84	16	19
Special Capital Region of Jakarta	6	21	7	75	25	28
Gorontalo	6	16	2	89	11	18
Jambi	11	32	5	86	14	37
West Jawa	27	113	14	89	11	127
Central Java	35	140	27	84	16	167
East Java	38	153	23	87	13	176
South Kalimantan	13	40	5	89	11	45
Central Kalimantan	14	39	6	87	13	45
East Kalimantan	10	33	6	85	15	39
North Kalimantan	5	16	1	94	6	17
West Kalimantan	14	50	8	86	14	58
Riau Islands	7	18	5	78	22	23
Lampung	15	53	8	87	13	61
Maluku	11	30	3	91	9	33
North Maluku	10	24	5	83	17	29
West Nusa Tenggara	10	36	4	90	10	40
East Nusa Tenggara	22	54	14	79	21	68
Papua	9	21	6	78	22	27
West Papua	7	19	4	83	17	23
Southwest Papua	6	15	3	83	17	18
Highland Papua	8	30	2	94	6	32
South Papua	4	17	1	94	6	18
Central Papua	8	19	4	83	17	23
Riau	12	49	5	91	9	54
West Sulawesi	6	17	3	85	15	20
South Sulawesi	24	65	15	81	19	80
Central Sulawesi	13	36	7	84	16	43
Southeast Sulawesi	17	40	11	78	22	51
North Sulawesi	15	39	6	87	13	45
West Sumatera	19	55	8	87	13	63
South Sumatera	17	56	10	85	15	66
North Sumatera	33	98	17	85	15	115
Total	514	1633	274	86%	14%	1907
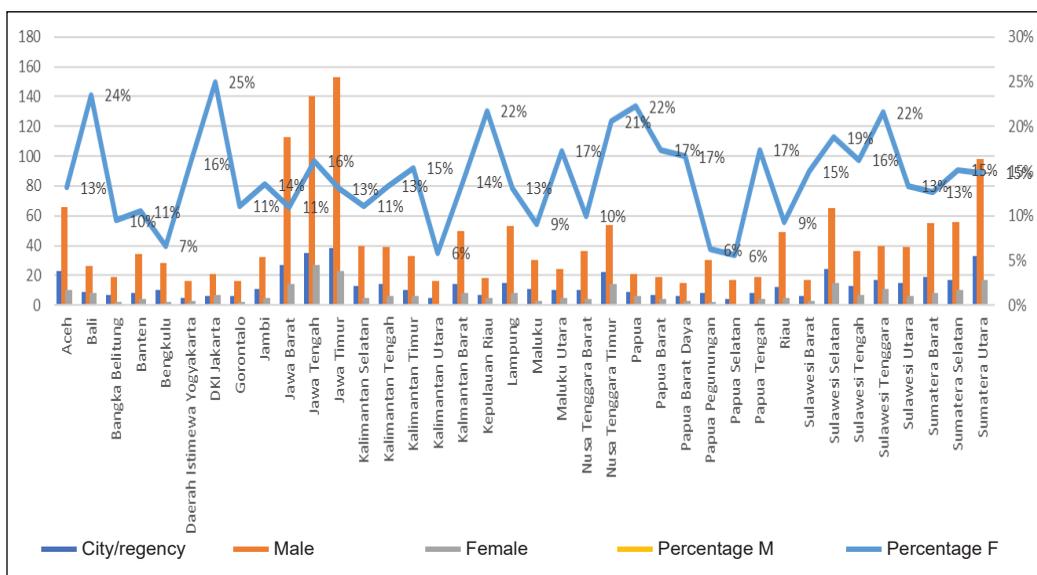
Figure 3. Women representatives as regency / City election supervisors in Indonesia during 2023–2028
Source: Authors' work

Meanwhile, other factors that hinder women's representation as election supervisors are a lack of understanding of gender in the political system in Indonesia, a lack of political networks owned by female candidates, and a risk burden for female election supervisors (Muharam et al., 2023a). The researchers' findings on women's representation as election supervisors, which did not comply with the rules on general elections that regulate the need for 30 percent of women's representation, indicate that there is still a negative stigma, such as gender bias, against women occupying very strategic political positions. Existing regulations were not applied in the selection process, and women's representation as election supervisors should be prioritized following applicable laws in Indonesia. There are obstacles to fulfilling women's quotas in fulfilling women's representation

as election supervisors due to several factors, including:

There is a Mistake in Understanding Gender in Politics

The perception of the political world that still views politics is the world of men (Muharam et al., 2023b; Muharam & Prasetyo, 2021; Warjiyati, 2016). There is still an understanding of patriarchy that men are the dominant party, and women are exploited (Kusumawardhana & Abbas, 2018). In addition, it discusses the issue of women's understanding of sex differences (sex) and the concept of gender (Warjiyati, 2016). This understanding will closely influence the understanding of gender differences and gender injustice with the structure of injustice in society. The existence of this understanding is that there are still many views that the political world tends to be considered a man's

world, causing stereotypes carried out by each individual (Zamroni, 2013). Even in making regulations on laws and regulations that regulate women's rights, especially in the representation of women 30%. Both as a legislature and party management member and as an election organizer such as the General Election Commission (KPU) and the General Election Supervisory Agency (Bawaslu). Not only that, the process of implementing human rights in exclusive laws and regulations regulates women's rights in fulfilling two fundamental democratic rights for women, namely the right to vote and the right to stand for elections.

The Burden of Risk Being an Election Supervisor

Being an election supervisor who certainly intersects a lot with various political interests is undoubtedly one of the factors in fulfilling women's representation as election supervisors. The risk of working as an election supervisor that requires women to adjust to various problems faced by election participants certainly requires courage for women to supervise, prevent and take action for perpetrators who commit violations in implementing in the field. Commitment for women who want to be involved as election supervisors must understand the workload and risks often faced, such as intimidation and harassment from election participants who are considered disturbed by the presence of election supervisors during campaigns and program socialization. With the workload and risks faced, female candidates involved

in election administration as election supervisors need to consider carefully the fulfillment of regulatory requirements stipulated in the election law and how the fulfillment of women's rights in politics can be achieved in accordance with regulations.

Limitations of Political Networks

One of the problems women face in competing for political office as commissioners of the election supervisory body is the limited political network women own. Few potential female candidates qualified in the race for seats as Bawaslu commissioners because they did not have political networks. This political network is considered important because it is undeniable that political relations for prospective election supervisors will make it easier for female candidates to be elected. The election of bawaslu commissioners is at the central level (Bawaslu Republik Indonesia). It is undeniable that this is one of the problems owned by women who lack political networks at the top level. These political networks can include right-wing political parties, religious organizations in Indonesia such as Nahdlatul Ulama (NU) and Muhammadiyah, and current stakeholder relations.

These three factors are one of the reasons why women's representation as election supervisors is less accommodated or not elected. Moreover, regulations on women's representation in the administration of elections, especially in the Election Supervisory Agency (Bawaslu), have not been adequately implemented and

practiced. The primary existence of this regulation seems to complement and abort the obligation of politicians to fulfill gender equality in politics. The selection process for occupying the commissioner position at the Election Supervisory Board was still dominated by men. Applicants from women could not be said to be significant. Of course, this impacts the results of selecting women's representatives in the Election Supervisory Agency.

CONCLUSION

Several regulations have been made to regulate the fulfillment of women's rights in politics. However, these regulations are only limited to rules, and there is no administrative punishment for disobeying these rules. For example, there is no obligation for election administrators to meet the quota of 30 percent women representation, as there is no sanction for those who do not fulfill the quota. Regulations regarding fulfilling women's political rights as supervisors have been regulated in Law Number 7 of 2017 concerning General Elections. However, in practice, several factors hinder the fulfillment of women's political rights as election supervisors. Some women choose not to exercise women's human rights as election supervisors because of the heavy burden of risk. An election supervisor's responsibilities and duties differ from those of other election organizers, such as in General Election Commissions (KPU), District Election Committees (PPK), Voting Committees (PPS), and Voting Organizing Groups (KPPS).

The findings of this study are various factors affecting the lack of women's electability in election supervisory institutions: (1) A selection process that uses a knockout system without considering the 30% female representation quota, (2) The political networks of female election supervisory candidates are very narrow and inflexible in building networks with decision-makers, (3) The patriarchal culture that is still dominant in some regions prioritizes men to be election supervisors. Women's representatives in some regencies/cities could not be found, (4) The selection team did not pay attention to women's gender representation. Meanwhile, at each stage of selection, there are women representatives. However, they were not selected in the final stage, and (5) Regulations on women's representation need to be changed, especially the rule that requires each election administrator to place 30% of women's representation seats. It is imperative to encourage women's participation in politics by creating an environment that supports and minimizes the burden of risk as an election supervisor. In addition, efforts are needed to increase awareness and understanding of the importance of women's representation in politics, especially in election supervisory institutions where not many parties understand women's representation as election organizers.

The implication of this study as study material for researchers in the field of gender politics is to understand that women's representation as election organizers is one of the studies that need to be developed.

Political research studies in gender studies focus more on affirmative quotas of women in parliament, in the management of political parties, and legislative candidates. However, the studies did not investigate much on election organizers, including election supervisors in Indonesia. The representation of women as commissioners at the Election Supervisory Board is one of the political positions that is very interesting to study. This study recommends that the Indonesian government make efforts to change the general election law, especially the rules on women's representation as election supervisors, in which the rules only stipulate up to the regency/city level. This rule differs from the representation of women as members of the General Elections Commission, where the authorities have reached the Voting Organizing Group (KPPS) level. In addition, rules regarding sanctions are necessary for the selection team or various parties involved in the selection process that do not fulfill 30% of women's representation as election supervisors.

ACKNOWLEDGEMENTS

REFERENCES

Abu-Zayd, G., Azizah, W., Ballington, J., Bylesjö, C., Dahlerup, D., Ginwala, F., Kadirgamar-Rajasingham, S., Karam, A., Lovenduski, J., Mar'iyah, C., Matland, R. E., Myakayaka-Manzini, M., Parawansa, K. I., Pintat, C., Rai, S., Reyes, S. L., Seda, F. S., Shvedova, N., & Skjeie, H. (2002). *Perempuan di parlemen bukan sekedar jumlah*. [Women in parliament are not just numbers]. International Institute for Democracy and Electoral Assistance. https://www.idea.int/sites/default/files/publications/perempuan-di-parlemen-bukan-sekedar-jumlah.pdf

Ahmad, A., Mahmood, Q. K., Saud, M., & Mas'udah, S. (2019). Perempuan dalam demokrasi: Partisipasi politik perempuan [Women in democracy: The political participation of women]. *Masyarakat, Kebudayaan, Dan Politik*, *32*(2), 114–122.

Alhuzail, N. A. (2021). The social representation of the Bedouin woman. *Women's Studies International Forum, 86,* Article 102474. https://doi.org/10.1016/j.wsif.2021.102474

Ang, C. W., & Lai, S. L. (2023). Women's empowerment in Malaysia and Indonesia: The autonomy of women in household decision-making. *Pertanika Journal of Social Sciences and Humanities*, *31*(2), 903–916. https://doi.org/10.47836/pjssh.31.2.22

Asy'ari, H. (2023). *Representasi perempuan dalam politik* [Representation of women in politics] [Paper presentation]. The Indonesia Women's Congress, Semarang, Indonesia.

Boo, H. S. (2021). Gender norms and gender inequality in unpaid domestic work among malay couples in Malaysia. *Pertanika Journal of Social Sciences and Humanities*, *29*(4), 2353–2369. https://doi.org/10.47836/pjssh.29.4.14

Brisbane, L., Hua, W., & Jamieson, T. (2023). Morality and the glass ceiling: How elite rhetoric reflects gendered strategies and perspectives. *Politics and Gender*, *19*(3), 806–840. https://doi.org/10.1017/S1743923X2200023X

Childs, S., & Krook, M. L. (2008). Critical mass theory and women's political representation. *Political Studies*, *56*(3), 725–736. https://doi.org/10.1111/j.1467-9248.2007.00712.x

Cole, E. R., & Sabik, N. J. (2010). Associations between femininity and women's political behavior during midlife. *Psychology of Women Quarterly*, *34*(4), 508–520. https://doi.org/10.1111/j.1471-6402.2010.01600.x

Creswell, J. W. (2014). *Research design qualitative, quantitative and mixed methods approaches*. SAGE Publications. https://www.pdfdrive.com/qualitative-quantitative-and-mixed-methods-approaches-e91943566.html

Dahlerup, D., & Freidenvall, L. (2005). Quotas as a "fast track" to equal representation for women. *International Feminist Journal of Politics*, *7*(1), 26–48. https://doi.org/10.1080/1461674042000324673

Deininger, K., Jin, S., Nagarajan, H. K., & Xia, F. (2015). Does female reservation affect long-term political outcomes? Evidence from rural India. *Journal of Development Studies*, *51*(1), 32–49. https://doi.org/10.1080/00220388.2014.947279

Diogo, S., Carvalho, T., & Breda, Z. (2021). Nomination vs. election: Do they influence women's access to institutional decision-making bodies? *Journal of Management and Governance*, *25*(3), 879–898. https://doi.org/10.1007/s10997-020-09538-6

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598. https://doi.org/10.1037/0033-295X.109.3.573

Hannah, C., Giroux, S., Krell, N., Lopus, S. E., McCann, L. E., Zimmer, A., Caylor, K. K., & Evans, T. P. (2021). Has the vision of a gender quota rule been realized for community-based water management committees in Kenya? *World Development*, *137*, Article 105154. https://doi.org/10.1016/j.worlddev.2020.105154

Heywood, A. (2004). *Political theory: An introductions*. Palgrave Macmillan. https://mgdc-charariSharief.com/elearn/Political Theory-1stSem.pdf

Kusumawardhana, I., & Abbas, R. J. (2018). Indonesia di persimpangan: Urgensi "undang-undang kesetaraan dan keadilan gender" di Indonesia pasca deklarasi bersama Buenos Aires pada tahun 2017 [Indonesia at a crossroads: The urgency of "gender equality and justice act" in Indonesia with Buenos Aire]. *Jurnal HAM*, *9*(2), Article 153. https://doi.org/10.30641/ham.2018.9.153-174

Lovenduski, J. (2019). Feminist reflections on representative democracy. *Political Quarterly*, *90*, 18–35. https://doi.org/10.1111/1467-923X.12563

Moleong, L. J. (2012). *Metodologi penelitian kualitatif* [Qualitative research methodology]. PT Remaja Rosdakarya.

Muharam., R. S., Pradhanawati, A., Fitriyah, & Supratiwi. (2023a). Fulfillment of women's political rights through mainstreaming gender quotas in Election Supervisory Board. *Otoritas*, *13*(3), 406–423. https://doi.org/https://doi.org/10.26618/ojip.v13i3.12691

Muharam, R. S., Pradhanawati, A., Fitriyah, Supratiwi, & Hidayat, A. (2023b). Women's empowerment in fulfilling the affirmative quota of election supervisors. *The 2023 International Conference of UNAIR Postgraduate School*, *3*(3), 11–19. https://doi.org/https://doi.org/10.31098/bmss.v3i3.685 Research

Muharam, R. S., & Prasetyo, D. (2021). Pemenuhan hak politik perempuan sebagai pengawas pemilihan kepala daerah serentak tahun 2020 [Womens political rights fulfilment to become supervisor the 2020 simultanous regional heads election]. *Jurnal HAM*, *12*(2), 273–284. https://doi.org/10.30641/ham.2021.12.273-284

Nurcahyo, A. (2016). Relevansi budaya patriarki dengan partisipasi politik dan keterwakilan perempuan di parlemen [The relevance of patriarchal culture to women's political participation and representation in parliament]. *Agastya: Jurnal Sejarah Dan Pembelajarannya*, *6*(01), 25-34. https://doi.org/10.25273/ajsp.v6i01.878

Office for Democratic Institutions and Human Rights. (2014). *Handbook on promoting women's participation in political parties*. https://www.osce.org/files/f/documents/e/f/120877.pdf

Phillips, A. (2019). Descriptive representation revisited. In R. Rohrschenider & J. Thomassen (Eds.), *The oxford handbook of political representation in liberal democracies* (pp. 1–23). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198825081.013.8

Prastiwi, J. H. (2018). Menakar kontribusi undang-undang pemilu tahun 2017 terhadap peningkatan keterwakilan perempuan [Measuring the contribution of the 2017 election law to increasing women's representation]. *Jurnal Wacana Politik*, *3*(1), 1–13. https://doi.org/10.24198/jwp.v3i1.15783

Primagita, P., & Riantoputra, C. D. (2019). Factors that influence woman–leader identity conflict among Indonesian women leaders. *Pertanika Journal of Social Sciences and Humanities*, *27*(3), 1769–1786.

Pusat Kajian Politik Universitas Indonesia. (2021). *Meningkatkan (lagi) keterwakilan perempuan dalam lembaga penyelenggara pemilu* [Increasing (again) women's representation in election organizing institutions]. https://puskapol.fisip.ui.ac.id/factsheet-meningkatkan-lagi-keterwakilan-perempuan-dalam-lembaga-penyelenggara-pemilu/

Republik Indonesia. (2011). *Undang-undang Republik Indonesia nomor 2 tahun 2011 tentang perubahan atas undang-undang nomor 2 tahun 2008 tentang partai politik* [Law of the republic of Indonesia number 2 of 2011 regarding amendments of law number 2 of 2008 about political party]. https://bphn.go.id/data/documents/11uu002.pdf

Republik Indonesia. (2012). *Undang-undang Republik Indonesia nomor 8 tahun 2012 tentang pemilihan umum anggota perwakilan rakyat, dewan perwakilan daerah, dan dewan perwakilan rakyat daerah* [Law of the Republic of Indonesia number 8 of 2012 concerning the general election of members of the people's representatives, regional representative councils, and regional people's representative councils]. https://peraturan.bpk.go.id/Details/37659/uu-no-8-tahun-2012

Republik Indonesia. (2017). *Undang-undang Republik Indonesia nomor 7 tahun 2017 tentang pemilihan umum* [Law of the Republic of Indonesia number 7 of 2017 concerning general elections]. https://www.mkri.id/public/content/pemilu/UU/UU No.7 Tahun 2017.pdf

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2012). *Research methods in psychology*. McGraw-Hill. https://perpus.univpancasila.ac.id/repository/EBUPT180048.pdf

Turnbull, B. (2019). Quotas as opportunities and obstacles: Revisiting gender quotas in India. *Politics and Gender*, *17*(2), 324–348. https://doi.org/10.1017/S1743923X19000722

Vasandani, M. R., Nugraha, D. P., & Susantijo, S. (2022). Affirmative action study on the political rights of women in the Indonesian constitution. *Constitutional Review*, *8*(1), 62–86. https://doi.org/10.31078/consrev813

Warjiyati, S. (2016). Partisipasi politik perempuan perspektif hukum Islam [Women's political participation: Islamic law perspective]. *Al-Daulah: Jurnal Hukum Dan Perundangan Islam*, *6*(1), 1–27. https://doi.org/10.15642/ad.2016.6.1.1-27

World Economic Forum. (2023). *Global gender report 2023*. https://giwps.georgetown.edu/resource/global-gender-gap-report-2023/

Zamroni, M. (2013). Perempuan dalam kajian komunikasi politik dan gender [Women in the study of political and gender communication]. *Jurnal Dakwah*, *14*(1), 103–132. https://doi.org/10.14421/jd.2013.14105

Case Study

The Development of College English Teachers' Pedagogical Content Knowledge in Moral Education Integrated College English Teaching in China

Lina Wang[1,2]*, Arshad Abd Samad[2] and Nur Surayyah Madhubala Abdullah[2]

[1]*School of Foreign Language Education, Zhuhai College of Science and Technology, 519000 Zhuhai, China*
[2]*School of Education, Taylor's University, 47500 Selangor, Malaysia*

ABSTRACT

In mainland China, college English (CE) teachers have been actively involved in integrating moral education into college English courses to promote a more rounded development of students. In the transition from traditional college English teaching to moral education integrated college English (MECE) teaching, college English teachers' pedagogical content knowledge (PCK) is subject to change. This study was designed as a qualitative case study to conduct a 3-year investigation to investigate the changes in PCK among college English teachers through the transition from CE to MECE. The trajectory of PCK development was collected from three teachers via content representation (CoRe), semi-structured interviews, and teaching materials. As suggested from the findings, in the process of transition from CE to MECE, the pathways that led to changes in PCK are specific objective-based study with reflection, participating in training sessions and workshops, interacting in effective online/face-to-face communities, as well as embracing the new relationship of co-learning with students. The findings underscored the need for comprehensive professional development programs that equip college English teachers with the skills to integrate moral education into their instruction. The research highlights the practical challenges that must be overcome by teacher training programs in being proactive, responsive, and adaptive to an evolving educational landscape, ensuring teachers are equipped not just to meet new pedagogical demands

E-mail addresses:
wanglina@sd.taylors.edu.my (Lina Wang)
arshad.abdsamad@taylors.edu.my (Arshad Abd Samad)
Surayyah.Abdullah@taylors.edu.my (Nur Surayyah Madhubala Abdullah)
* Corresponding author

but also to actively engage in the complex discourse on integrating moral education.

INTRODUCTION

Pedagogical content knowledge (PCK) is an important concept in educational research, and it has been widely studied since Shulman (1986) described it as a construct regarding teachers' ability to transmit their knowledge of a specific subject to the students. It was initially raised in the domain of teachers' professional development. According to Shulman (1986), teachers' PCK refers to the overlapped area of pedagogical knowledge and content knowledge. In the following decades, based on Shulman's (1986) PCK theory, research has emphasised the significance of context-specific approaches to PCK development, recognising the influence of teachers' prior knowledge, student needs, and institutional support (Loughran et al., 2012). In mathematics education, studies have explored the role of teacher education programs in the development of mathematics teachers' PCK, emphasising the significance of opportunities for reflection and collaboration (Lo, 2023). Much of the current literature on PCK still pays particular attention to the teachers in science (Mesci et al., 2020) and mathematics (Anney & Bulayi, 2020; Moh'd et al., 2021).

Research in English language teaching (ELT) has highlighted the complexity of the language learning process and the importance of PCK for effective teaching (Shulman, 1986). Theoretically, PCK is an integrated whole in which components cannot be separated. Based on Shulman's (1986) PCK theory, researchers have been focusing on the understanding of PCK components (Grossman, 1990; Li et al., 2021; Shariatifar et al., 2017; Tseng et al., 2022; Xu, 2015). Supported by the clarification of PCK components, teachers will pay more attention to understanding the language content, teaching strategies, and student's needs, which enable them to make informed instructional decisions and effectively engage learners in the language learning process. In the context of foreign language (FL) and second language (L2) teaching, research on PCK has been scarce in recent years (Han et al., 2021). Several studies have investigated the development of English language teachers' PCK (Elas et al., 2019; Han et al., 2021; Sadeghi et al., 2022). In foreign language education, research has highlighted the importance of integrating technology in teacher education programs to enhance language teachers' PCK (Rets et al., 2023; Tseng et al., 2022; Yeh et al., 2021). The existing literature on English language teachers' PCK revealed that PCK is subject-specific, domain-specific, and topic-specific (Großschedl et al., 2019; Li et al., 2021). In foreign language education, the context of teaching a specific language with a specific purpose is critical for teachers' PCK development. In the case of college English teachers, their PCK would encompass their understanding

of English language teaching and their knowledge of the best teaching practices and strategies for helping students improve their language skills.

Moral Education Integrated College English (MECE) in China

In mainland China, higher education institutions focus on the fundamental task of strengthening moral education for cultivating people. Moral education should be carried out in four aspects: ideal and faith education, patriotism education, socialist core values education, and excellent virtue education for the Chinese national. To put into practice, it is of great importance to integrate moral educational resources into all courses (Eryong & Li, 2021). An increasing number of higher education institutions have been encouraging teachers to integrate moral education into college English curriculum. College English teachers have been actively involved in the integration of moral education into college English courses by advocating the integration of moral education into English language teaching as a means of fostering students' character and values (Zhang, 2023). The integration of moral education into higher education courses is a topic of interest and debate. According to Li (2009), the roots of moral education in China can be traced back to Confucianism, which emphasised the importance of moral cultivation as a key element of education. In modern times, the Chinese government has increasingly focused on moral education to promote social stability and national development.

In China, the integration of moral education into college English courses is a relatively recent development that has gained attention in the last decade. It has been driven, in part, by the recognition that language learning is not simply a matter of acquiring linguistic competence but also involves the development of intercultural communicative competence and the promotion of cultural understanding and empathy (Byram, 2020). Guided by the fundamental goal of strengthening moral education for cultivating people, college English instruction has evolved beyond mere language acquisition, becoming a vital platform for imparting values and moral education. This shift acknowledges language learning as a cognitive process and an ethical endeavour. In that context, CE in China have become crucibles for ideal and faith education, patriotism education, socialist core values education and Chinese national excellent virtue education, where language learning intertwines with value transmission (Feng, 2019; Novawan et al., 2020).

Despite the clear directives from educational authorities, CE teachers face challenges in integrating moral education principles into their teaching practices. The research reveals a significant gap in the pedagogical content knowledge (PCK) required by teachers to effectively deliver such integrated lessons. The disparity between the educational goals set by policymakers and the practical knowledge of teachers in this new domain has motivated the present study, to explore and address the professional development needs of college

English teachers in China to successfully integrate moral education into their teaching repertoire.

This qualitative research, therefore, seeks to understand the research problem from the individual's perspective of the action they are taking or the skills they are developing when integrating moral education concepts into language teaching. It investigated the development of college English teachers' PCK by answering the research question: How does Chinese college English teachers' PCK change when they integrate moral education into college English teaching? What are the pathways that led to changes in college English teachers' PCK in integrating moral education into college English teaching?

MATERIALS AND METHODS

The Study

A qualitative case study approach was adopted, as this can assist the researcher in delving into the case by providing insights into the real problem through the examination of a particular situation (Stake, 1995).

Context and Participants

This study focused on a group of college English teachers who have experienced the transition of teaching from college English (CE) to moral education integrated college English (MECE). A 2-year longitudinal case study was carried out at X University in East China. In this university, CE teaching happens in the first two years,

after which the student takes the College English Test (CET), a task for college English learning. Since 2018, CE teaching has been required to integrate moral education, and all CE teachers have been integrating moral education into classroom activities. Considering the English language proficiency and the CET-taking needs of most students, teachers are encouraged to reconstruct their lessons in ways that are in line with the basic goal of fostering virtue. Such reconstruction provides the college English teachers with various opportunities to arrange specific activities to implement collaborative classroom teaching.

The participants of this study were a group of college English teachers who conducted moral education integrated college English (MECE) teaching from 2018 to 2022. The researchers formally invited college English teachers to participate in this research after receiving permission from the office of the College English Teaching Department. Nine CE teachers agreed to participate. The demographic information of those nine CE teachers was obtained. The influence of demographic factors on teachers' PCK development has been proven in previous studies (Shulman, 1986). It was also used as the criterion for purposeful sampling. Through purposeful sampling, three CE teachers were selected to participate in this qualitative case study. Table 1 provides the demographic information of those three participants. The three participants are given the pseudonyms Lilly, Sara, and Nina to guarantee anonymity. Along with the research plan introduction,

Table 1
The demographic information of the three selected participants

Participant *pseudonyms	Professional rank	Gender	Age	Educational background	Years of CE teaching	Years of MECE teaching
Lilly	Professor	Female	49	MA in English Language and Literature	20	4
Sara	Associate professor	Female	40	PhD candidate in Education	12	4
Nina	Lecturer	Female	30	PhD student in Applied Linguistics	8	2

Source: Authors' work

the participants voluntarily signed the consent forms before data collection so that the teachers understood what they were agreeing to do.

Data Collection

Data was collected in December 2018, December 2019, December 2020, and December 2021. During those years, two rounds of data collection were completed (2018–2019 is the first round, and 2020–2021 is the second round). The collected data from those two rounds were analysed to investigate whether changes happened and what happened in teachers' PCK during the moral education integrated college English (MECE) teaching compared with the prior college English (CE) teaching. The multiple sources of qualitative data, which consist of content representation (CoRe), interviews, and document analysis, were collected to implement triangulation. At first, the data reflecting the participants' PCK development along with the progress of MECE courses was collected to see if there were changes after the transition to MECE teaching. The timeline of data collection and

the data sources of the two collecting rounds were demonstrated (Figure 1).

The data collection lasted two years (approximately 12 months for each round), during which the researcher collected data four times (Figure 1). CoRe offers a holistic overview of particular concepts related to a content area. It is also helpful in accessing teachers' PCK by suggesting a powerful and useful representation of PCK responsive to practice (Loughran et al., 2012). By completing CoRe, the participants shared their understanding of PCK and developments. Teachers' PCK becomes evident by explicitly expressing the nature of their pedagogical reasoning and the related decision-making in teaching particular content (Loughran et al., 2012). It is also a very useful tool to explicitly focus teachers' attention on what is important in designing lessons (Hume & Berry, 2011). In each round, the participants were asked to complete the CoRe matrix each semester, which included a set of questions (Table 2).

Grossman (1990) proposed a model in which the Pedagogical Content Knowledge consists of four parts: conceptions of
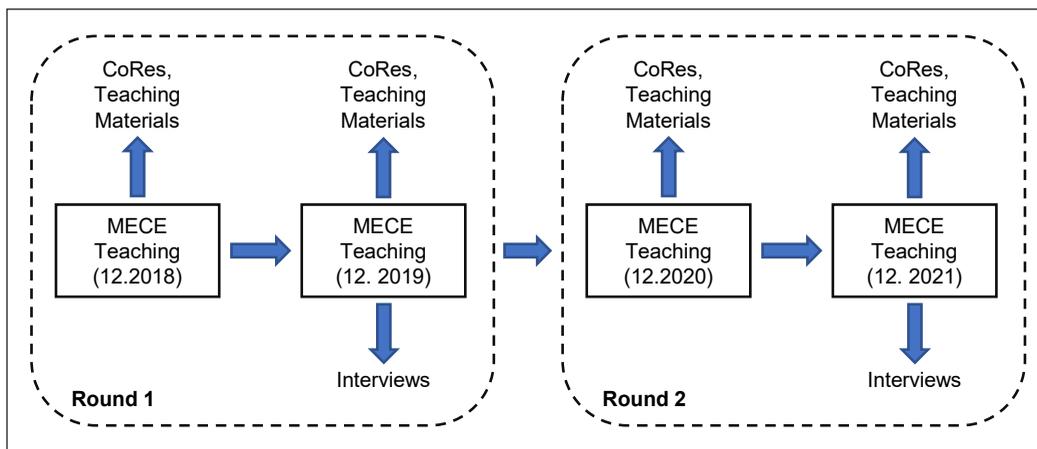
Figure 1. Data collection timeline and data sources
Source: Authors' work

Table 2
Content representation (CoRe) matrix

The year level for which this CoRe is designed	Content area		
	Big idea 1	**Big idea 2**	**Big idea 3**
(1) What do you intend the students to learn about this idea?			
(2) Why is it important for students to know this?			
(3) What else do you know about this idea (that you do not intend students to know yet)?			
(4) Are there any difficulties/limitations connected with teaching this idea?			
(5) What do you know about students' thinking which influences your teaching of this idea?			
(6) Are there any other factors that influence your teaching of this idea?			
(7) What teaching procedures are conducted, and what are the particular reasons for using these to engage with this idea?			
(8) What are the specific ways of ascertaining students' understanding or confusion around this idea?			

Source: Authors' work

purposes for teaching subject matter, knowledge of students' understanding (KOSU), Curricular knowledge (CK), as well as knowledge of instructional strategies (KOIS). Underpinned by Grossman's (1990) PCK model, the set of questions corresponds to the four specific components of PCK. For example, Question 1 looks at the conception of purposes for teaching MECE, Questions 2, 3 on curricular knowledge (CK), Questions 2, 3, and 5 on knowledge of students' understanding (KOSU), and Questions 4, 6, and 7 on knowledge of instructional strategies (KOIS). As for teaching materials,

the participants' syllabi and supplementary teaching materials were collected at the end of each semester. After two collections of CoRe and teaching materials, in-depth semi-structured interviews (SI) were conducted in each round.

Before conducting semi-structured interviews, the researcher prepared an interview guide with open-ended questions to encourage detailed responses (see Appendix I). The information-seeking questions were developed based on the specific components of PCK proposed by Grossman (1990) respectively. The participants clarified or confirmed their responses to the CoRe matrix in SI. Then, they responded to the questions concerning the changes after the transition to MECE teaching. All interviews were conducted in the participants' offices or classrooms, each lasting about one hour. With consent, the interviews were audio-recorded to capture the conversation for later transcription and analysis. The interview data were sent back to the participants for confirmation after being transcribed and translated by the researcher. Participation in both rounds of data collection was based on the written consent of the participants. Ethical approval was obtained from X University during the study and was strictly observed.

Data Analysis

Thematic analysis was undertaken to interpret the data and unveil respondents' thoughts and opinions through deconstruction. The analysis of data collected from CoRe responses and teaching materials

was conducted while the fieldwork was undertaken in parallel. The data collected from interviews were transcribed verbatim, and the informants confirmed the accuracy of the data. In this process, the researcher guaranteed the informants' anonymity to avoid potential negative influence during the data analysis. Four Core responses, four teaching material files, and two interview transcripts from each participant were collected. After a holistic reading of the CoRe responses, teaching materials, and interview transcripts, the data were coded in a way that aligned with the four PCK components proposed by Grossman (1990). Another data analysis strategy used was the comparative approach. Data from different sources in the two rounds were compared according to the PCK components to see if there were any changes, and this process was cyclical.

FINDINGS

This research collected only the changes in the transition to MECE teaching as the developments in teachers' PCK. The findings and discussion are presented case by case to portray the three teachers' PCK development over the two years. Table 3 demonstrates the participants' PCK during the two rounds of the research period.

Lilly, Sara, and Nina all went through a transition in integrating moral education into college English teaching. They all tried to adapt to the MECE teaching, modified their teaching goals, and reconstructed their lessons to align with the purpose of moral education.

Table 3

The development of the three teachers' PCK in moral education integrated CE teaching

	Conceptions of purposes for teaching MECE	Knowledge of students' understanding (KOSU)	Curricular knowledge (CK)	Knowledge of instructional strategies (KOIS)
Lilly	Focusing on language skills	Being active in what students know and what they need to learn.	Limited understanding of moral education content	Using similar strategies to CE
	Focusing on language skills and value guidance	Being active in what students know and what they need to learn.	Choosing MECE teaching materials critically	Using similar strategies to CE
Sara	Cultivating students' cultural confidence	Unaware of students' difficulties in learning MECE	Relying on the textbook	Trying some new strategies
	Cultivating students' cultural confidence	Trying to interact with students	Trying to be familiar with the MECE content	Using similar strategies to CE
Nina	Cultivating students' language skills to tell China's story	Unaware of students' interests	Conducting MECE teaching in a superficial way	Experimenting with different strategies
	Cultivating students' language skills to tell China's story	Failing to motivate students	Considering language and moral education as interwoven	Using similar strategies to CE

Source: Authors' work

Lilly

Lilly's conceptions of the purpose of teaching MECE changed a lot. She focused on language skills teaching in the first round and showed unawareness of the differences in the teaching aims. In the second round, she added value guidance to the teaching aims. As to the knowledge of students' understanding (KOSU), in the first round of teaching, she would use quizzes and tests to gauge students' understanding of a particular topic. In the second round, she conducted group discussions and then collected the students' feedback to get a sense of what students were struggling with and what they had grasped. From her responses in CoRe and her statement in the interview, her understanding of the content

and the structure of the moral education module were limited in the first round. When trying to adapt to MECE teaching, she realised that curricular knowledge (CK) could help teachers understand the teaching context and provide a framework for making instructional decisions. She managed to acquire the knowledge she needed from some sample lessons shared online and in pieces of training, for instance, the moral education subject matter being taught, the methods used to teach it, and the assessment strategies used to evaluate student learning. In the second year of MECE teaching, she realised that the purpose of teaching should be changed, so she added moral education activities in line with an additional teaching objective: value guidance. As for Knowledge

of instructional strategies (KOIS), Lilly kept using a product-oriented approach (POA) that emphasises the outcome of students' learning. She insisted that changes were not necessary for strategies because POA is good enough. She firmly believed that her role in classroom teaching would always be a participant.

Sara

Sara supported MECE teaching, and she developed an understanding of the purposes of teaching MECE before the transition happened. In the CoRe response, her big ideas of the content area were in line with cultivating students' cultural confidence. In her reflective journals, it can also be confirmed that she had changed her understanding of the purpose of teaching MECE. During the interview, she stated that moral education is such a big concept that she needed time to grasp it. In preparation for MECE lessons, she found that one of the teaching goals in the textbook was to build students' intercultural competency, especially cultural confidence. So, she focused on building the students' positive and self-assured attitude towards their own culture and a willingness to engage with and learn from other cultures. During the research, she developed the other three PCK components (KOSU, CK and KOIS). In the first round, Sara relied on the textbook due to her limited understanding of MECE content, and she tried some new strategies for teaching MECE.

In the second round, Sara tried to be more familiar with the MECE content and

brought critical thinking to the existing moral education section in the textbook. She did not change her understanding of the purpose of teaching MECE. While her KOSU has developed, she decided to go to the students and listen to the students' opinions on MECE learning. According to her statement in SI, she considered herself a participant in MECE teaching, not a teacher or instructor. As for CK, she found that she made a mistake when she thought of moral education and language teaching as two dissociated items, not as one aspect of a single unity. Therefore, she constructed the lessons and integrated moral education into activities based on unit-specific content. In the second round, her KOIS was developed, and she chose strategies similar to those used in college English teaching. She attributed her PCK development to her interaction with her colleagues.

Nina

Nina supported MECE teaching, and her conception of the purpose for conducting moral education in college English teaching is to cultivate students' language skills to tell China's story in the two rounds of teaching. She built that conception after participating in a teaching competition. In China, there are several competitions for college English teachers aimed at improving the quality of English language instruction in higher education institutions. Nina participated in the National English Teaching Contest for College English Teachers (NETC), held annually since 2001. The Ministry of Education and the State Language

Commission organised the contest. It aims to promote innovative teaching approaches and pedagogical practices among English teachers in colleges and universities across the country. The contest allowed English teachers to demonstrate their teaching abilities, share best practices, and learn from other teachers. During the preparation for the competition, Nina developed her conception of the purposes of MECE teaching. She believed that telling China's story in English is a good catalyst for promoting the student's English language acquisition and the ability to advocate Chinese culture to the world.

As for KOSU, Nina failed to motivate students and gave up collecting the students' feedback in the first round. However, she encouraged students to participate actively in learning in the second round. For CK, in her first round of MECE teaching, she added some videos and passages from the website to share the Chinese culture, and she told the students that there were specific activities for them to cherish the Chinese culture. She would highlight some keywords and sentences for the students to recite. From her statement in the SI, she admitted that her implementation of MECE was quite superficial in the first round. After the teaching contest, she considered language and moral education as interwoven. For KOIS, after analysing Nina's CoRe responses, it can be found that she has been experimenting with different strategies to integrate moral education into language teaching in the first round, such as flipped classrooms, projected-based learning, as well as communicative

language teaching. In the second round, she concluded that her instructional strategy was differentiated instruction, a similar strategy that she adopted in language teaching. She considered herself a facilitator rather than a teacher and encouraged students to participate actively in learning. While transitioning to MECE teaching, she designed her lessons based on her understanding of MECE and considered language and moral education interwoven.

DISCUSSION

Concerning the research question, it was found that all the respondents' CK changed, and they acquired knowledge of moral education from different resources by themselves, such as training sessions, peers, and sample lesson records. As for KOSI, Sara and Nina tried new instructional strategies and ended up with similar strategies to CE. A possible explanation might be that the teachers felt stressed when acquiring moral education knowledge. An innovation in instructional strategies makes them full of anxiety. Another important finding is that the conceptions of teaching purposes changed in two respondents (Sara and Nina) during the transition from CE to MECE lessons. Then, they maintained the same conception of purpose for teaching during the following two years of MECE teaching. Lilly only changed the conception of purpose for teaching in the second year after realising the transition was a need. The most unexpected finding is that all the respondents agreed that MECE teaching practices changed the relationship between students

and teachers. In KOSU, Lilly kept being active, Sara changed from unawareness of students' learning difficulties to trying to interact with students, and Nina failed to motivate students. Their KOSU results showed a difference; they all confirmed that co-learning is a new relationship when conducting moral education-related teaching and learning activities. In the process of transition from CE to MECE, the PCK growth pathways in college English teachers are suggested from the findings.

Specific Objective-based Study with Reflection

In MECE teaching, the CE teachers hold specific objectives when preparing MECE lessons: integrating moral education into college English lessons. With the various objectives, the CE teachers explore the resources of MECE and study by themselves as the first step in the transition from general to MECE teaching. The CE teachers who conduct self-study will focus on the moral education concepts that correlate with English lessons, teaching and learning activities and instructional pedagogies. Reflection is the intentional act of thinking through complex or unclear ideas in learning or tackling problems without clear solutions to derive meaning (Moon, 1999). With critical reflection during self-studying, CE teachers will develop their conceptions of purposes for teaching MECE, not just those limited to "cultural confidence cultivation" (Sara) or "value guidance" (Lilly). The CE teachers' CK will develop correspondingly during the classroom activities designing

and reflection process. Critical reflection provides opportunities for modelling reflection and seeking alternate rhetoric to improve practice (Ergas & Ragoonaden, 2020). With the designed MECE activities in the classroom, the CE teachers will develop KOSU by interacting with students and critically reflecting on the effectiveness of the designed MECE activities.

The objective-based self-study is like an evolutional procedure; it also provides chances to examine professional identity formation in CE teachers' practice and its impact on student learning (Ergas & Ragoonaden, 2020). Knowledge about self is important for transforming pedagogy and practice (Heilman,2003). In the process of studying and reflecting, the CE teachers will gradually develop their way of thinking and doing to adapt to the transition from general CE teaching to MECE teaching.

Participating in Training Sessions and Workshops

The participants' instructional strategies showed little change, potentially due to the stress associated with adopting new teaching methods. To ease the stress caused by instructional innovation, institutions can provide training sessions and workshops accordingly. Training sessions and workshops will provide resources such as instructional books, advanced instructional pedagogies, effective learning or teaching platforms or mobile applications, and experience-sharing opportunities. Continuous teacher training addresses the demands arising from societal shifts and

progress, as well as the complexities in grasping educational methodologies and curricula that challenge existing teaching approaches (Boumadan et al., 2020; Dyment & Downing, 2020; Rodrigues, 2020; Wong et al., 2023). The Ministry of Education (MOE) currently organises moral education integrated curriculum development training that is open to all university teachers. It is divided into specific modules based on different courses. University teachers can choose their courses and enrol in the training. The curriculum that consists of moral education integrated curriculum development training sessions can assist CE teachers in enhancing their understanding of moral education integrated CE courses in a short period.

Simultaneously, such training can invite distinguished educators to share cutting-edge theories and teaching methods in moral education integrated curriculum development. Training sessions and workshops can also invite experienced teachers with rich MECE teaching experience to share their insights. Furthermore, these training sessions are available for online viewing and can be replayed. It provides a sense of security for university teachers, as they can access these course resources for review anytime and anywhere. If the course modules of such training can be organised according to the Pedagogical Content Knowledge (PCK) framework, or if a separate module dedicated to enhancing PCK is established on the training website, it will result in a more targeted improvement of teachers' PCK.

Interacting in Effective Online/Face-to-face Communities

The findings demonstrate that all participants confirmed the effectiveness of interacting with peers in the transition from general CE teaching to integrating moral education into CE teaching. A potential reason for this response could be that teachers share their successes and challenges, providing mutual support and learning from each other's experiences, which is particularly valuable in navigating new educational contexts.

The training and workshops mentioned in the previous text are a rapid means to connect with the community. Besides that, the institutions could encourage collaboration between CE and moral education teachers. For example, they could gather CE teachers and moral education course teachers to construct MECE by sharing resources such as lesson plans, teaching materials, and assessments. It is a way to provide regular feedback and cooperation among teachers, highlighting areas where they can improve and offering guidance on how to make the changes, encouraging teachers to think critically about their teaching and consider ways to improve. The shared community discourse also allows entertaining a variety of viewpoints and utilises these as a source for reflection (Lin et al., 1999). In addition, creating a culture of continuous improvement and innovation within the institution can help to foster a mindset of openness to change and a willingness to develop PCK.

Embracing the New Relationship of Co-learning with Students

For both CE teachers and students, the content of moral education is entirely new in the early stages of MECE teaching and learning. The relationship between teachers and students tends to lean towards co-learning, mutual learning, and collaborative practice. It aligns with existing research, indicating a shift in teacher-student dynamics towards more cooperative models of education. This trend is evidenced by Harumi (2023), Polifiore (2021), and Wei (2023), who have identified a pivot towards co-learning environments, where teachers and students engage in joint educational endeavours. Complementary to this, Li et al. (2021) and Whitehouse et al. (2021) corroborate the efficacy of mutual learning, where shared experiences and knowledge contribute to the learning process. Additionally, the principle of collaborative practice, as detailed by Burns et al. (2014), Hansen et al. (2020), and Syahabuddin et al. (2020), further supports this study's observation of a more integrative and partnership-based approach in educational settings. Once CE teachers embrace this new relationship with students, the CE teachers will continuously receive feedback from students during the collaborative MECE learning process. Teachers will be active in what students know and need to learn, being aware of students' difficulties in learning (Lestari & Wahyudin, 2020). The CE teachers' KOSU will be enhanced when CE teachers manage to interact with the students as co-learners.

CONCLUSION

The findings support the idea that tracking teachers' PCK changes provides the development pathways of PCK. The study reveals that college English (CE) teachers' content knowledge evolves through independent learning from various sources, such as pieces of training and peer interactions, with a focus on integrating moral education into CE lessons. Notably, MECE practices transformed teacher-student relationships, fostering a co-learning environment despite varied teacher engagement and effectiveness levels. Professional development through training and workshops and participation in educational communities were instrumental in supporting teachers through this transition. By embracing new pedagogical content knowledge (PCK) and co-learning relationships, teachers are better equipped to respond to educational changes and contribute to students' learning experiences. The research identifies practical challenges that must be overcome by teacher training programs in being proactive, responsive, and adaptive to an evolving educational landscape, ensuring teachers are equipped not just to meet new pedagogical demands but also to actively engage in the complex discourse on integrating moral education.

Implications

This research emphasises that teachers' pedagogical content knowledge (PCK) is not static but evolves through diverse avenues, including independent study, peer collaboration, and formal training.

Such growth necessitates adaptive professional development frameworks within educational institutions, which should extend beyond mere knowledge acquisition to practical application within the classroom environment. The study also brings to light a transformative shift in teacher-student dynamics towards a co-learning model, advocating for educational ecosystems that promote student-centred learning experiences. This shift is critical in fostering an environment that values collaborative learning and mirrors modern educational paradigms that place the student at the heart of the learning process.

Limitations and Recommendations

This study adds to the limited research that tracks English teachers' PCK development when designing and implementing moral education practices in college English lessons. This study could be a starting point to develop a project with CE teachers to learn more about teachers' professional development. Meanwhile, the participants were from one university database associated with English language teaching, which shows the limitation of this study. Future studies that can recruit larger participant samples for foreign language lessons which integrate moral education concepts other than college English would help us better understand the factors that could influence PCK developmental trajectories.

ACKNOWLEDGEMENTS

REFERENCES

Anney, V. N., & Bulayi, M. (2020). Experienced mathematics teachers PCK in the use of learner-centred approaches in Tanzania's secondary schools. *Papers in Education and Development*, *37*(1), 94-122.

Boumadan, M., Soto-Varela, R., Ortiz-Padilla, M., & Poyatos-Dorado, C. (2020). What factors determine the value of an online teacher education experience from a teacher's perspective? *Sustainability, 12*(19), Article 8064. https://doi.org/10.3390/su12198064

Burns, M., Pierson, E., & Reddy, S. (2014). Working together: How teachers teach, and students learn in collaborative learning environments. *International Journal of Instruction, 7*(1), 17-32.

Byram, M. (2020). *Teaching and assessing intercultural communicative competence: Revisited.* Multilingual Matters. https://spada.uns.ac.id/pluginfile.php/253332/mod_resource/content/1/ICC%20Byram.pdf

Dyment, J. E., & Downing, J. J. (2019). Online initial teacher education: A systematic review of the literature. *Asia-Pacific Journal of Teacher Education*, *48*(3), 316–333. https://doi.org/10.1080/1359866X.2019.1631254

Elas, N., Majid, F., & Narasuman, S. (2019). Development of technological pedagogical content knowledge (TPACK) for English teachers: The validity and reliability. *International Journal of Emerging Technologies in Learning (IJET)*, *14*(20), 18-33. https://doi.org/10.3991/ijet.v14i20.11456

Ergas, O., & Ragoonaden, K. (2020). Two perspectives on teaching mindfulness in teacher education: A self-study of two selves. In O. Ergas & J. K. Ritter (Eds.), *Exploring self toward expanding teaching, teacher education and practitioner research* (pp. 179-196). Emerald Publishing Limited. https://doi.org/10.1108/S1479-368720200000034011

Eryong, X., & Li, J. (2020). What is the ultimate education task in China? Exploring "strengthen moral education for cultivating people" ("Li De Shu Ren"). *Educational Philosophy and Theory*, *53*(2), 128–139. https://doi.org/10.1080/001318 57.2020.1754539

Feng, W. D. (2019). Infusing moral education into English language teaching: An ontogenetic analysis of social values in EFL textbooks in Hong Kong. *Discourse: Studies in the Cultural Politics of Education, 40*(4), 458-473. https://doi.org/10.1080/01596306.2017.1356806

Großschedl, J., Welter, V., & Harms, U. (2019). A new instrument for measuring pre-service biology teachers' pedagogical content knowledge: The PCK-IBI. *Journal of Research in Science Teaching*, *56*(4), 402-439. https://doi.org/10.1002/tea.21482

Grossman, L. P. (1990). *The making of a teacher: Teacher knowledge and teacher education.* Teacher College Press. https://lccn.loc.gov/90038509

Han, J., Zhao, Y., Liu, M., & Zhang, J. (2021). The development of college English teachers' pedagogical content knowledge (PCK): From general English to English for academic purposes. *Asia Pacific Education Review, 22*(4), 609-621. https://doi.org/10.1007/s12564-021-09689-7

Hansen, J. H., Carrington, S., Jensen, C. R., Molbæk, M., & Secher Schmidt, M. C. (2020). The collaborative practice of inclusion and exclusion. *Nordic Journal of Studies in Educational Policy,* *6*(1), 47-57. https://doi.org/10.1080/20020317.2 020.1730112

Harumi, S. (2023). The role of learner-initiated questions as a pedagogical resource for co-learning: Development of teacher identity for leaderful classrooms. In S. Egitim & Y. Umemiya (Eds.), *Leaderful Classroom Pedagogy Through an Interdisciplinary Lens: Merging Theory with Practice* (pp. 103-117). Springer Nature. https://doi.org/10.1007/978-981-99-6655-4_7

Hume, A., & Berry, A. (2011). Constructing CoRes—a strategy for building PCK in pre-service science teacher education. *Research in science education*, *41*, 341-355. https://doi.org/10.1007/s11165-010-9168-3

Lestari, M., & Wahyudin, A. Y. (2020). Language learning strategies of undergraduate EFL students. *Journal of English Language Teaching and Learning, 1*(1), 25-30.

Li, S., Liu, L., & Jiang, A. L. (2021). Understanding the development of Chinese EFL student-teachers' pedagogical content knowledge. *Frontiers in Psychology, 12*, 1-16. https://doi.org/10.3389/fpsyg.2021.627728

Lin, X., Hmelo, C., Kinzer, C. K., & Secules, T. J. (1999). Designing technology to support reflection. Educational Technology, *Research and Development, 47*(3), 43–62. https://doi.org/10.1007/BF02299633

Lo, W. Y. (2023). The complexities of mathematical knowledge and beliefs within initial teacher education: An analysis of three cases. *Education Sciences, 13*(7), 697. https://doi.org/10.3390/educsci13070697

Loughran, J., Berry, A., & Mulhall, P. (2012). *Understanding and developing science teachers' pedagogical content knowledge* (Vol. 12). Springer Science & Business Media. https://rest.neptune-prod.its.unimelb.edu.au/server/

api/core/bitstreams/62da3fc4-2f5c-55f4-a882-7f7c038be0cf/content

Mesci, G., Schwartz, R. S., & Pleasants, B. A. S. (2020). Enabling factors of preservice science teachers' pedagogical content knowledge for nature of science and nature of scientific inquiry. *Science & Education, 29*(2), 263-297. https://doi.org/10.1007/s11191-019-00090-w

Moh'd, S. S., Uwamahoro, J., Joachim, N., & Orodho, J. A. (2021). Assessing the level of secondary mathematics teachers' pedagogical content knowledge. *Eurasia Journal of Mathematics, Science and Technology Education, 17*(6). https://doi.org/10.29333/ejmste/10883

Moon, J. A. (1999). *Reflection in learning and professional development theory and practice*. Routledge. https://doi.org/10.4324/9780203822296

Novawan, A., Aisyiyah, S., Miqawati, A. H., Wijayanti, F., & Indrastana, N. S. (2020). Exploring teachers' perspectives on morality in EFL Pedagogy. *Journal of ELT Research, 5*(1), 81-93.

Polifiore, M. (2021). Co-learning a language with your students—A classroom teacher's perspective. *Languages Victoria, 25*(2), 6-7.

Rets, I., Rienties, B., & Lewis, T. (2023). Transforming pre-service teacher education through virtual exchange: A mixed-methods analysis of perceived TPACK development. *Interactive Learning Environments, 31*(3), 1229-1241. https://doi.org/10.1080/10494820.2020.1826983

Rodrigues, A. L. (2020). Digital technologies integration in teacher education: The active teacher training model. *Journal of e-learning and knowledge society, 16*(3), 24-33. https://doi.org/10.20368/1971-8829/1135273

Sadeghi, S., Aliakbari, M., & Yasini, A. (2022). A model of EFL teachers' pedagogical content knowledge: A data-driven approach. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 41*(2), 205-243. https://doi.org/10.22099/tesl.2021.41332.3031

Shariatifar, S., Kiany, G., & Maftoon, P. (2017). High school EFL teachers' professional competencies: Content knowledge and pedagogical content knowledge. *Applied Research on English Language, 6*(4), 499-522. https://doi.org/10.22108/are.2018.107409.1191

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14. https://doi.org/10.3102/0013189X015002004

Stake, R. (1995). *The art of case study research*. Sage. https://doi.org/10.1007/978-3-030-85124-8_7

Syahabuddin, K., Fhonna, R., & Maghfirah, U. (2020). Teacher-student relationships: An influence on the English teaching-learning process. *Studies in English Language and Education, 7*(2), 393-406. https://doi.org/10.24815/siele.v7i2.16922

Tseng, J. J., Chai, C. S., Tan, L., & Park, M. (2022). A critical review of research on technological pedagogical and content knowledge (TPACK) in language teaching. *Computer Assisted Language Learning, 35*(4), 948-971. https://doi.org/10.1080/09588221.2020.1868531

Wei, L. (2023). Transformative pedagogy for inclusion and social justice through translanguaging, co-learning, and transpositioning. *Language Teaching, 57*(2). 1-12. https://doi.org/10.1017/S0261444823000186

Whitehouse, M., Rahm, H., Wozniak, S., Breunig, S., De Nardi, G., Dionne, F., Fujio, M., Graf, E. M., Matic, I., Mckenna, C. J., Steiner, F., & Sviķe, S. (2021). Developing shared languages: The fundamentals of mutual learning and problem solving in transdisciplinary collaboration. *AILA Review, 34*(1), 1-18. https://doi.org/10.1075/aila.00038.int

Wong, J. T., Bui, N. N., Fields, D. T., & Hughes, B. S. (2023). A learning experience design approach to online professional development for teaching science through the arts: Evaluation of teacher content knowledge, self-efficacy and STEAM perceptions. *Journal of Science Teacher Education, 34*(6), 593-623. https://doi.org/10.1080/1046560X.2022.2112552

Xu, W. (2015). Exploring ESL/EFL teachers' pedagogical content knowledge on reading strategy instruction. *English Language Teaching, 8*(11), 155-175. https://doi.org/10.5539/elt.v8n11p155

Yeh, Y. F., Chan, K. K. H., & Hsu, Y. S. (2021). Toward a framework that connects individual TPACK and collective TPACK: A systematic review of TPACK studies investigating teacher collaborative discourse in the learning by design process. *Computers & Education, 171*, Article 104238. https://doi.org/10.1016/j.compedu.2021.104238

Zhang, F. (2023). Inspiration of moral judgment evaluation methods on assessing curriculum ideological and political education for foreign language courses. *The Educational Review, USA, 7*(3), 344-349. https://doi.org/10.26855/er.2023.03.011

APPENDIX I

Interview Guide

Theme	Questions
Conceptions of purposes for teaching moral education integrated college English (MECE)	1. What is your attitude toward teaching freshmen MECE? 2. What knowledge or ability do you want your students to acquire when teaching MECE?
Curriculum knowledge (CK)	3. Do you have any new understanding of curricular knowledge? What are the sources of your understanding? 4. What kind of teaching material do you usually use?
Knowledge of instructional strategies (KOIS)	5. Have you developed any new strategies to facilitate the students' learning? How do you acquire them? 6. What is your role in classroom teaching?
Knowledge of students' understanding (KOSU)	7. What is the student's attitude toward learning MECE? 8. Are there any difficulties when the students learn MECE? What kind of difficulties?
Developing of PCK	9. Overall, what promotes or obstructs your PCK development? Please give some examples

The Quagmire of Internal Displacement: Reinstatement Challenges and Counter Strategies in District Swat, Pakistan

Muhammad Rafiq¹*, Muhammad Saeed² and Asan Ali Golam Hassan³

¹*College of Business Administration, Prince Mohammad Bin Fahd University KSA, Dhahran, Kingdom of Saudi Arabia*
²*Department of Global Studies, King Fahd University of Petroleum & Minerals (KFUPM) Dhahran, Kingdom of Saudi Arabia*
³*Azman Hashim International Business School, Universiti Teknologi, Malaysia,54100 Kuala, Lumpur, Malaysia*

ABSTRACT

In the period following the military operation against insurgents in the Swat district during 2009–10, the Pakistani government, with support from international organizations, initiated a comprehensive program aimed at the rehabilitation and reintegration of Internally Displaced Persons (IDPs). This research study utilizes the Delphi method to gather data through face-to-face interviews with officials involved in the rehabilitation efforts. The analysis revealed a range of challenges, such as security threats, infrastructural deficiencies, bureaucratic complexities, educational system revival, unmet donor commitments, financial constraints, weakened local governance, and communication breakdowns. Despite these obstacles, strategic interventions have ameliorated the IDPs' conditions, including enhanced security, temporary infrastructure, international partnerships, makeshift educational provisions, and a dedicated rehabilitation budget. Nevertheless, the full rehabilitation of IDPs remains an unachieved goal. This article suggests that successfully reintegrating affected communities requires innovative strategies tailored to Pakistan's unique socio-structural context, which could serve as a model for similar situations globally.

INTRODUCTION

Pakistan is susceptible to experiencing internal displacement caused by social and natural disasters (Sayed & González, 2014; Shah et al., 2018). In 2009-10,

a military operation against insurgents followed by a devastating flood forced around 2.3 million people to flee from Pakistan's Swat district (Ahmad, 2018; Din, 2010; Internal Displacement Monitoring Centre, 2012). This turbulence wreaked havoc on the local infrastructure and collapsed the socioeconomic conditions (Asian Development Bank, 2010). Reports illustrate that nearly 400 schools, 8500 houses, and several hotels and shops were partially or fully destroyed (Provincial Disaster Management Authority, 2019). The government's main challenge was reinstating the internally displaced persons (IDPs) who returned to their original place of residence after the completion of the military operation in 2010 (Sayeed & Shah, 2017). With the help of national and international NGOs and donor organizations, the government of Pakistan launched multifaceted plans for the rehabilitation and reinstatement of IDPs (Khan, 2011; Rafiq et al., 2021; Rafiq, Hassan, 2022). However, reinstating the IDPs and re-establishing normality in Swat was an unprecedented challenge due to the scale and speed of the repatriation (Ahmad, 2018; Zahid, 2012).

Several studies (Ahmad, 2018; Ahmad et al., 2017; Akhunzada et al., 2015; Azim et al., 2018; Din, 2010) have probed into the history and reasons behind the mass displacement of people from district Swat. Some studies have researched the socioeconomic conditions of IDP returnees, often relying on quantitative data and structured tools for data collection (Khan & De Nardi, 2024; Rafiq et al., 2021; Rafiq,

Hassan, 2022; Swaroop & DeLoach, 2015). A few studies have reported a grim situation of the rehabilitation and reinstatement of IDPs from district Swat. For example, bad governance, mismanagement, corruption, and nepotism impeded the rehabilitation and reinstatement of IDPs in Swat (Bangash, 2012; Elahi, 2015). Some studies have documented not only the lack of existence of a uniform policy at the state level to deal with the issue of Internal Displacement effectively but also have reported the no adherence to the United Nations Guiding Principles on Internal Displacement and Rehabilitation (Ahmad, 2018; Ahmad et al., 2017; Din, 2010). Research advocates that appropriate resettlement programs significantly improve subjective well-being (Randell, 2016) and that the happiness of people is associated with the resettlement scheme and associated social, cultural, and economic transformations (Boadi et al., 2022; Danquah et al., 2014; Herath et al., 2017; Xu et al., 2021).

The scholarly landscape notably lacks detailed examinations that illuminate the complexities and apply solutions in rehabilitating and reintegrating Internally Displaced Persons (IDPs) within the Swat district. This deficiency underscores an imperative to thoroughly explore the barriers and impacts stemming from rehabilitation efforts on the return of IDPs to Swat. This investigation is of paramount importance as it seeks to mend the existing schism by meticulously reviewing a wide range of relevant literature and weaving together various theoretical perspectives pertinent

to the successful rehabilitation of IDPs. Through the lens of firsthand accounts and insights from officials, this study casts a spotlight on the extensive array of challenges encountered. It delineates the strategies devised to overcome these hurdles. This innovative method augments our comprehension of the subject and lays a foundational stone for subsequent inquiries within this field. In this milieu, this study sets forth the following objectives.

1. To explore the political, administrative, and socioeconomic challenges of rehabilitation projects faced during the reinstatement of returnees in Swat, Pakistan.

2. To investigate the counterstrategies adopted by the authorities to cope with political, administrative, and socioeconomic challenges of rehabilitation projects faced during the reinstatement of returnees in Swat, Pakistan.

The scope of this study is to examine the efficacy of rehabilitation initiatives, delving into the insights, obstacles, and tactical responses associated with the rehabilitation process. The research methodically traverses through a critical review of existing literature and a robust theoretical framework, advancing toward the methodologies employed. Subsequently, the findings are revealed, and a thorough discussion is conducted before culminating in a concise conclusion. The study also thoughtfully acknowledges its own limitations, ensuring a balanced and reflective scholarly discourse.

Background of the Study

Swat, a former princely state with an independent government, merged with Pakistan in 1969 to improve the living conditions of its people. However, the merger ended up being a failure and had a negative impact on the local population (Khan & De Nardi, 2024). The English judicial system introduced did not align with the local population's customs and failed to outperform the traditional Jirga, a conflict resolution system based on Sharia Law (Lutfullah et al., 2022). Jirga was an effective forum for solving issues between the government and the militants. Still, due to differences, the Jirga has been manipulated by the state and the militants in Swat. Hence, the higher-level Jirga, such as the Swat Qaumi Jirga (Swat National Jargas), have become increasingly active since 2007, whereas lower-level (village) Jirga have become less common in practice in many parts of Swat (Azim et al., 2018). Consequently, civil and criminal cases that had previously been resolved in a matter of weeks began to drag on for years, causing high costs and stress for those involved (Bangash, 2012).

The genesis of militancy, exemplified by movements such as Tehreek Nifaz-e-Muhammadi (TNSM), started by Maulana Sufi Muhammad in 1989, and Tehreek-e-Taliban Pakistan (TTP), spearheaded by Maulvi Fazlullah in 2005, is deeply rooted in the soils of poverty and underdevelopment (Ahmad, 2018; Khan, 2020; Saeed, 2012). These conditions facilitated and accelerated the rise of such movements, serving as

fertile ground for their emergence and proliferation (Hashmi, 2016; Khalid & Helge, 2010; Peracha et al., 2012). In 2008, in an unexpected turn of events, the Tehrik-i-Taliban Pakistan (TTP) leader took command of the Swat district with a daring pledge: to return the region to its "golden era" under the Wali. This period was marked by the enactment of Sharia Law, offering prompt and cost-free justice, a stark contrast to the pre-merger era. This bold initiative aimed at dispelling the pervasive despair among the people of Swat. Mula Fazalullah, in a courageous move, openly defied the state's authority, securing considerable support from the public (Adnan, 2012). This support, however, swiftly transformed into widespread turmoil as the TTP embarked on a brutal campaign that included the

beheadings of adversaries, public lashings of both men and women, coerced marriages of women to TTP leaders, mutilation of corpses in public squares, prohibition of female education, and the destruction of law enforcement infrastructure (Khaliq, 2010). This political unrest made the region unsafe (Saeed & Griffin, 2019). This escalation of violence prompted the Pakistani government to launch a military operation against these militants, a move that resulted in the displacement of approximately 2.3 million individuals to safer areas within the country (Ahmad, 2018; Haider, 2009). See the map of district Swat below in Figure 1.

In the repercussion of the military operation's completion and the subsequent return of internally displaced persons (IDPs), the Swat region faced a devastating flood
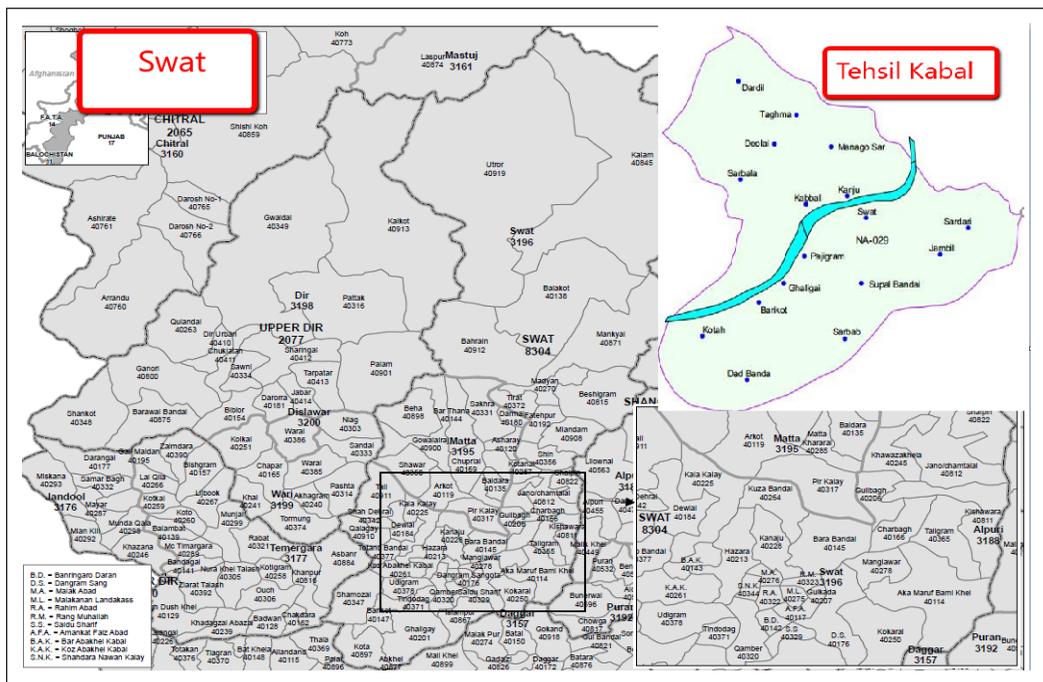
Figure 1. Map of Swat district
Source: Office for the Coordination of Humanitarian Affairs (2010)

(Internal Development Monitoring Centre, 2018). This calamity further exacerbated Swat's tribulations by displacing over a million individuals anew (Internal Development Monitoring Centre, 2018). These events inflicted considerable damage on the local infrastructure and precipitated a collapse in the socioeconomic conditions of the IDPs (Shaheen, 2011). Conversely, the repatriation process was remarkable in its scale and rapidity, presenting the Pakistani government with formidable challenges in facilitating the returnees' reintegration into swat and the resumption of their normal lives (Shaheen, 2011). Complicating matters, insurgent activities, including kidnappings for ransom and the execution of hundreds of government and international organization employees engaged in rehabilitation efforts, significantly shifted the administrative focus towards security concerns, detracting from the economic and social spheres (Din, 2010). This shift interrupted or decelerated various international organizations' rehabilitation projects (Elahi et al., 2015). Furthermore, the comprehensive restoration of infrastructure—encompassing homes, roads, hospitals, and schools—presented an overwhelming challenge for the government (Zahid, 2012), underscoring the multifaceted nature of the crisis and the intricate path towards recovery.

LITERATURE REVIEW

Various research (Aburas et al., 2018; Ahmad, 2018; Cantor et al., 2021; Ogunode et al., 2022; Rafiq, Hassan, 2022; Rafiq, Saeed, 2022) have been conducted on the rehabilitation of displaced people. The displacement is a growing concern as people lose their sources of livelihood and become dependent on aid provided by the government and aid agencies (Wistrand, 2023). A study by Kanishchev et al. (2022) found that a sudden influx of IDPs in a new host area can create economic and social issues, such as a decline in wages, pressure on schools, hospitals, and the job market, as well as an increase in crime rates (Khan et al., 2018; Shami, 1993; Ullah et al., 2017). Moreover, most IDPs experience stress and depression, making women, children, and the elderly particularly vulnerable (Morina et al., 2018; Murphy et al., 2018; Ulke et al., 2021).

Besides, the main contributory factors to the long-standing migration phenomena are economic consideration, natural disasters, socio-political issues, demographic growth and urbanization, conflicts, and family reunification (Urbański, 2022). Most migration studies have concentrated on push-pull factors for human migration (Ahmad et al., 2017; Anwar & Ahmad, 2017; Bangash, 2012; Edge et al., 2014). The push factors for displacement include political instability and conflict, whereas the pull factors are the rehabilitation projects that might improve their standard of living (Rafiq, Saeed, 2022). Rehabilitation efforts often aim to restore normalcy, including social lives, education, and economic systems (Elahi et al., 2015; Serghiou et al., 2016). In the studies by (Adekola et al., 2024) and (Akilova et al., 2022), it is argued that rebuilding social infrastructure, establishing shelters, and

fostering peace, political reconciliation, and improved security are crucial for encouraging the return of internally displaced persons (IDPs) to their original homes. However, completely restoring the socioeconomic conditions of IDPs requires more time and cost (Salgado-Gálvez, 2018). Unfortunately, rehabilitation projects often lack the necessary resources, such as funding (Lam et al., 2015). According to Bain and Parkinson (2010), rehabilitation projects need huge expenditures because they support the affected people by providing them with employment and business opportunities, ultimately contributing to the country's economy. In such scenarios, aid organizations and NGOs collaborate with the government to provide quick aid and reinstate affected communities (İçduygu, 2015).

Often, the displaced communities experience a decline in their standard of living. For example, it was documented that nearly half of the respondents were unsatisfied with resettlement development schemes for Malaysia's displaced Orang Asli community (Abdullah et al., 2016). However, appropriate rehabilitation programs produce advantageous social and economic benefits for a displaced population to improve subjective well-being (Randell, 2016). One of the most important components of the afflicted individuals' rehabilitation programs is the physical infrastructure, which significantly affects tourism (Bashir et al., 2022). After agriculture, tourism is the second largest source of revenue in Swat, severely impacted by the region's inadequate physical infrastructure and security threats (Sohail et al., 2023). However, collaboration between the government and local communities, acknowledging the importance of heritage tourism and reinforcing policies for cultural heritage conservation boosts the economy (Ma et al., 2023). In addition, the rehabilitation programs offer opportunities for enhancing the impacted community's hard and soft skills. Employment is improved when technical and practical business knowledge and hard and soft skills demonstrate the coexistence of current and traditional values (Chan et al., 2021). On the contrary, a study conducted by Kinyanjui et al. (2017) examined how rehabilitation processes affect the safety and norms of the residents. For example, the Syrian refugees fled from their country due to security concerns and faced another shelter problem (Molnar, 2017). Thus, the rehabilitation programs have positive and negative consequences depending on the migration's culture, norms, and nature. Positive impacts include increased socioeconomic status, while negative implications include norms, safety, and security challenges.

Research in Pakistan has also revealed positive and negative consequences of relocation and rehabilitation. Internal displacement caused by the construction of the Tarbela Dam in Pakistan, for example, has disturbed social life and the social networks of individuals in the affected areas (Abbasi-Shavazi & Sadeghi, 2015; Magsi & Torre, 2014). Similarly, Edge et al. (2014) also addressed the dynamic socioeconomic

constraints associated with the social and economic development of IDPs in Pakistan. However, the large displacement from Pakistan's Swat Valley presented numerous hurdles for the administration in providing essentials to IDPs and returnees (Akhunzada et al., 2015). Nevertheless, the IDP returnees' socioeconomic condition improved, so the government coped with the situation (Loescher, 2017). A similar finding was carried out by (Elahi, 2015) that the government has been taking considerable measures to promote the resettlement and rehabilitation of IDPs in Swat.

Contrarily, studies indicated that the government had paid little attention to fulfilling humanitarian issues associated with the Swat IDPs and returnees. For example, Khan and Wei (2016) urged that the government provide minimal monetary compensation, which was insufficient to rehabilitate and reinstate the affected people completely. Therefore, it is necessary to reconsider the implementation of rehabilitation programs to protect the rights of IDPs in the country. Adam (2015) criticized the service delivery by supporting and helping authorities to the IDPs in Pakistan. Hence, people cannot access the service at the right time to meet their needs. In addition, Adam (2015) suggested that policymakers and key stakeholders can take necessary actions to protect the rights of marginalized and deprived people in the country. The findings of these studies revealed that, somehow, the returnee's socioeconomic condition improved through rehabilitation,

but the administration has faced various challenges in the rehabilitation process.

In district Swat, the government of Pakistan initially focused on delivering fundamental amenities like food, shelter, healthcare, and infrastructure (Khan et al., 2017). The subsequent stage involved reinstating education, peace, and healthcare and reviving the economic and social sectors, as mentioned by Bhatti and Shah (2017). According to research, the government encountered various obstacles while carrying out rehabilitation programs (Abbasi-Shavazi & Sadeghi, 2015; Lam et al., 2015; Peterson et al., 2018). Therefore, this study explores the challenges and strategies adopted for reinstating returnees IDPs in Swat, Pakistan.

Theoretical Background

This study draws on Kukrit Pramoj's (1957-58) theory of good governance, emphasizing effective, transparent, and accountable political systems for managing change (Waters, 2022). In Swat, Pakistan, stakeholders encountered multiple hurdles during rehabilitation, such as inadequate infrastructure, approval processes, budget allocation, and local governance issues, as illustrated in Figure 2. Pramoj's theory posits that robust governance is essential for rehabilitation projects' well-being, security, and success. We argue that safeguarding the rights of internally displaced persons (IDPs) depends on the successful implementation of these projects through good governance. This research highlights the obstacles faced in managing and reintegrating IDPs,

Figure 2. Theoretical background of the study (*Source:* Authors' work)

using Pramoj's framework to explore how public and private institutions address these challenges.

MATERIALS AND METHODS

This research study employed an exploratory research method. Specifically, the Delphi technique, including interviews with officials involved in the planning and executing rehabilitation processes, was used for data generation. The Delphi method's systematic and unbiased approach aligns well with rehabilitation research's dynamic and interdisciplinary nature, making it effective for achieving consensus and generating valuable insights (Al-Qawasmi et al., 2021; Landeta, 2006). Further, the interview process creates an environment of trust where participants feel heard and valued (Mason et al., 2010). It goes beyond a simple Q&A session, fostering a genuine exchange of thoughts and perspectives

(Creswell & Clark, 2017). The study participants were accessed through the Provincial Development Management Authority (PDMA), Pakistan. However, respondents from the local government representatives, including the education and health departments, Tehsil Nazem, and Commissioner's office in Swat, were accessed through a local political leader of the Awami National Party (ANP) (refer to Appendix 2 for visuals).

A purposive sampling technique was employed to select appropriate respondents. The purposive sampling method is used when participants are chosen on purpose because they are thought to be the best fit for the study (Creswell & Clark, 2017). Since this study aimed to gain insights into the challenges faced in the rehabilitation process for internally displaced persons, fifteen expert interviewees were selected from government and non-governmental

organizations who were actively involved in the rehabilitation process (Appendix 1). Choosing an appropriate sample size in qualitative research involves the researcher's experience and judgment in evaluating the usefulness and quality of the collected information (Sandelowski, 1995). The Delphi technique's "information power" approach was utilized to ensure the sample was information-rich. According to the "information power" approach, the fewer participants are required, the more information the sample contains that is relevant to the study (Malterud et al., 2016). The interviewees initially discussed the challenges they faced in assisting IDP returnees in Swat with their rehabilitation. This discussion was complemented by open-ended questions, leading to further inquiries.

This article is based on the corresponding author's PhD thesis, for which the data was collected over six months from January to July 2019 in the district of Swat. This duration enabled a comprehensive examination of the subject, ensuring a robust dataset that highlights nuances and variations over time.

The study adhered to the ethical standards for social research. Participants provided informed consent and were assigned pseudonyms to ensure their confidentiality and anonymity. The corresponding author, a Ph.D. scholar, received approval from the Institutional Review Board (IRB) and a no-objection letter from the University of Technology Malaysia (ref UTM.K.55.01.03/13.11/1/4; UTM.K.55.01.03/13.11/2/4), confirming compliance with ethical guidelines prior to initiating data collection. No participants were placed in vulnerable situations, and all had the option to withdraw their participation or opinions at any time during the study.

Data Analysis

The data collected was analyzed using an inductive approach, following a 6-step framework for analysis of the collected data (Braun & Clarke, 2006), as shown in Figure 3.

First, the audio-recorded data was transcribed word-for-word into text using Microsoft Word to analyze textual data. Reviewing the transcripts allowed us to observe patterns and themes. Subsequently, we transferred all data to Microsoft Excel for sorting by font color and theme, both alphabetically and numerically. In the
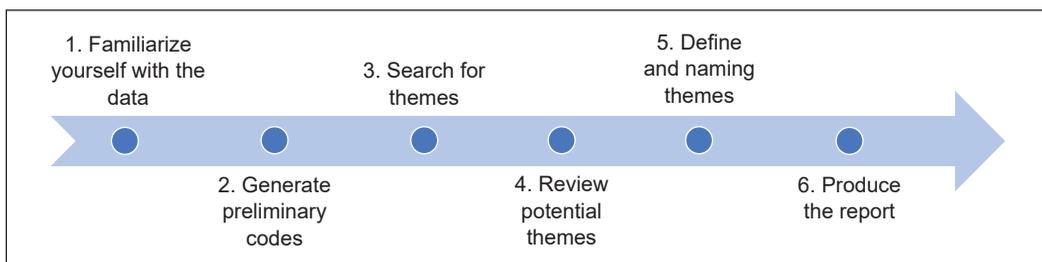
Figure 3. 6-step framework for analysis (*Source:* Author's work)

second step, we generated preliminary codes by assessing the themes and patterns in the transcripts. Rather than coding every word, we segmented the data into meaningful chunks, coding each part that reflected or revealed something significant about the research question. We then compared and adjusted these segments—the third phase involved organizing the codes into broader themes and subthemes guided by relevance. While codes might overlap, the determination of a theme remained flexible (Braun & Clarke, 2006). For example, codes related to an organization's challenges in rehabilitating returnees in Swat were amalgamated into a *field formation process* theme. In the fourth step, we refined these themes. We consolidated all relevant information under each theme using Microsoft Word's "cut and paste" feature to assess its validity and interrelations (Bree & Gallagher, 2016). This process helped us identify supportive data, combine related themes, and discard irrelevant details. The fifth step involved defining each theme and pinpointing its essence (Braun & Clarke, 2006). It included examining the relationships between main themes and subthemes, such as grouping sub-themes related to agricultural restoration and road/bridge reconstruction under the *infrastructural challenges* main theme. Finally, the sixth step focused on writing up the analysis. Here, we presented and interpreted the main themes, discussing the challenges in reinstating Internally Displaced Persons (IDPs) and the counterstrategies adopted.

RESULTS AND DISCUSSION

Challenge of Security and Mobility

The military operation in Swat District was designed to eradicate militants. However, these militants, familiar with the terrain, could seamlessly integrate as IDPs, thereby posing a security threat. This risk was a major concern for both locals and rehabilitation staff, who were potential targets. This theme's analysis corresponds with the study's first objective, which addresses political, administrative, and socioeconomic challenges. In this context, numerous interviewees voiced their concerns. For instance, Respondent R3 remarked:

> "*Due to the possibility of militants returning as IDPs, it was not safe to roam around the area freely. Furthermore, the public was reluctant to trust NGO employees, hindering them from providing rehabilitative programs without any restrictions*" (Respondent R3).

Despite government assurances of safety and security in the region, NGO workers and representatives from local communities continued to feel endangered. This finding aligns with the study by Reza (2010), who reported that in April and May of 2010, targeted assassinations claimed the lives of seven village defense committee members. This violence, underscored by a suicide bombing in Mingora that killed three people, signaled militant activity and heightened fears of an insurgency resurgence, significantly disrupting aid organizations' operations and mobility.

Especially for women workers, the volatile situation combined with cultural barriers deterred them from engaging in remote areas, thereby obstructing NGOs' ability to provide rehabilitation services. Another study by Muhammad et al. (2011) supports our findings, as their research noted that NGO workers encountered local distrust, security threats, and skepticism, impeding their movement and relief efforts due to perceptions of them as foreign agents.

In response, the government took steps to enhance security and facilitate ongoing rehabilitation projects, such as implementing stricter security measures, increasing checkpoints, and providing police escorts for personnel in high-risk zones. This successful strategy for addressing these challenges incorporates the good governance principles espoused by Kukrit Pramoj. However, these interventions also introduced complications. For example, Respondent R5 observed that enhanced security checks reduced threats and caused significant delays due to prolonged traffic queues. To mitigate these issues, the government issued special permits to NGOs and rehabilitation workers, alongside creating a dedicated vehicle lane at checkpoints for faster screening and smoother transportation for officials and aid workers.

Lack of Sufficient Physical Infrastructure

All interviewees identified the region's second most significant challenge for rehabilitation workers as the insufficient physical infrastructure, which impeded their mobility and service delivery. For instance, Respondent R5 stated:

> "*The difficulty of navigating the region due to significant damage to the physical infrastructure, worsened by the 2010 flood. The flood destroyed previously repaired facilities, further restricting access to remote mountain villages*" (Respondent R5).

Various studies complement our findings. One study has mentioned that during the military operation, Swat's infrastructure, encompassing roads and bridges, sustained substantial damage, impeding access to isolated regions (Bangash, 2012). This damage consequently hindered the timely rehabilitation of displaced individuals in these areas. Although the government and various aid organizations initiated numerous projects to rehabilitate the infrastructure, the catastrophic flood of 2010 negated much of their progress. The flood demolished numerous smaller bridges, three of which were significant ones in Gyeman, Shamozai, and Baghdairi ("Delay in construction of the bridge," 2013), severely restricting access to vital services and facilities. A report from the (United Nations, 2010) mentioned that the flood caused extensive damage, destroying approximately 170 bridges, 2,000 km of roads, and 158 government buildings, leading to major transportation delays. For example, traveling from Mingora to Kalam in a 4-wheel drive took around 13 hours to cover 93 km (Kiren, 2016). The disaster also devastated the agricultural infrastructure,

leaving irrigation channels, crops, and fields barren (Peracha et al., 2012). In an effort to mitigate these challenges, the government developed a strategy for establishing temporary infrastructure to facilitate essential relief services. This strategy aligns with Kukrit's approach to good governance. Respondent R6 has mentioned a detail of these initiatives:

"*Efforts were made to connect communities and principal regions swiftly by constructing temporary side roads and steel bridges. Eventually, we erected two major bridges, the Khwazakhela and Mingora bridges, and 15 roads spanning 44.07 km. We also refurbished and expanded 84 km of the road from Mingora to Fatehpur, rehabilitated 19 flood-affected irrigation channels, and founded an agricultural research center. Furthermore, we supported 801 small-scale infrastructure projects through the UNDP, such as culverts and local roads. However, the scarcity of roads and numerous security checkpoints escalated the cost of materials and hindered progres*s" (Respondent R6).

The temporary infrastructure reconstruction provided essential relief to displaced individuals, showcasing the government's smart strategy. Nonetheless, the efforts to reconstruct temporary roads and bridges led to increased costs and extended timelines. It was particularly evident in the case of the Ayoub Bridge, which, despite expectations to be completed by 2013, faced delays due to funding shortages. The rehabilitation of basic infrastructure and services spanned 3 to 4 years, with the government shouldering most expenses. This situation strained the government's budget, pushing the fiscal deficit to 6.3% of GDP in 2009/10, above the target of 5.1%. The total cost of reconstruction was estimated between Rs 662 billion (4.5% of 2009/10 GDP) and Rs 779 billion (5.3%). Challenges such as a lack of proper management plans, a scarcity of skilled labor, and insufficient heavy machinery further delayed the reconstruction efforts. Several respondents highlighted these issues. For example, respondents R1, R2, and R5 noted that the scarcity of labor, skilled personnel, and heavy machinery prolonged project timelines.

Security concerns and threats contributed to the labor and skills shortage, complicating reconstruction and rehabilitation efforts. Additionally, the local administration faced challenges in sourcing heavy machinery, often having to procure it from other provinces, which further delayed projects (Khan et al., 2017). The government and aid organizations offered higher wages to attract local labor and brought in skilled workers and heavy machinery from other regions to address these obstacles. Priority was given to restoring economic and social infrastructure, with 207 rehabilitation and reconstruction schemes implemented with UNDP support, facilitating access to

remote areas and aiding in the restoration of physical infrastructure (Mujtaba, 2013).

The government's initiative for temporary infrastructure reconstruction provided essential relief to displaced individuals. Despite its financial burdens and operational challenges, such as labor shortages and security concerns, this strategic approach to crisis management mirrors the principles of adaptive strategies and proactive management highlighted by Kukrit. These efforts in overcoming obstacles through higher wages and imported skilled labor underscore the application of good governance in crisis situations, aiming for efficient resource allocation and clarity in project execution.

Challenges in NOC Approval and Field Formation Process

The complex, time-intensive bureaucratic processes for securing No Objection Certificates (NOCs) have significantly impeded NGO operations. In the Swat region, for instance, these organizations could not function without the required NOC from pertinent authorities. Multiple respondents voiced their frustrations with the NOC approval mechanism, highlighting the daunting nature of acquiring NOCs for NGOs. Many respondents stated:

"*A stringent, multi-step approval sequence and tight security vetting as primary obstacles. Additionally, the army's necessity to grant final NOC approval further restricted many NGOs from engaging in rehabilitation initiatives for*

returnees, thus limiting their potential contribution" (Respondent R4, R11, R12, and R13).

One key concern the respondents raised was the rigorous NOC approval process for NGOs. According to them, NGOs would not be able to participate in rehabilitation without NOC approval. The process was complex and time-consuming, requiring approval from multiple government departments and security agencies. The government's main objective was to ensure that registered NGOs had no suspicious affiliations with banned organizations and adhered to the rule of law. However, aid organizations were allowed to collaborate with local registered NGOs to speed up the rehabilitation process. For instance, the government of Khyber Pakhtunkhwa updated the NOC process for aid organizations in 2018, allowing them to obtain their NOC within 35 days from proposal submission to approval. This one-step process would have given NGOs more time to rehabilitate and reinstate the returnees quickly while still ensuring the enforcement of strict policies against anti-state practices.

According to respondent R2, NGOs faced challenges in identifying potential beneficiaries due to the need for NOC approval and difficulties accessing isolated rural areas with limited employee mobility and damaged infrastructure. Consequently, the field formation process was time-consuming, causing delays in rehabilitation for returnees (Madeeha, 2017; Office for the Coordination of Humanitarian Affairs, 2013). To address this issue,

NGOs developed a structural process to identify beneficiaries at the village level, as mentioned by Respondent R7. He said:

"*Our approach to identifying poverty involved a three-tier system known as CRP (Community Resource Persons), which included the following sub-units: (1) CO (Community Organizer), responsible for 10-15 houses; (2) VO (Village Organizer), responsible for 5-10 villages; and (3) LSO (Local Supervisor Organizer), responsible for all villages. The affected individuals were assessed using the PSC (Poverty Score Care) form, which the CRP completed. Once the data was entered into the database and verified by the community, an EDP (Enterprise Development Plan) was prepared*" (Respondent R7).

Participatory Rural Appraisal (PRA) and Rapid Rural Appraisal (RRA) methodologies are frequently employed to gather information to address local community issues. RRA is particularly beneficial during a project's initial phases, where foundational information is necessary for devising a plan. Conversely, the PRA methodology emphasizes local community involvement and empowerment, a critical component for achieving project objectives (Townsley, 1996). This method fosters strong social cohesion, leading to rapid success through community efforts, as highlighted by Alifa and Nugroho (2019).

Social capital serves as the primary source of support for entrepreneurs, especially when formal institutions are inadequate (Khan & Sepulveda, 2022). Khan (2024) also suggests that a community-based rehabilitation approach that incorporates a local community holds the potential for effective reintegration. The effectiveness of this approach was exemplified in Swat, where the involvement of local community members was crucial in identifying vulnerable returnees. It supports Kukrit's approach, emphasizing how good governance facilitates public and private institutions in efficiently managing the affairs of internally displaced persons (IDPs) in Swat.

Restoration of the Education System

The Taliban displayed hostility towards Western education, especially female education, as they deemed it a catalyst for indecency and Westernization in the community. According to the Provincial Disaster Management Authority (2019), a total of 152 schools were bombed and destroyed, while 237 sustained partial damage. As a result, when the Internally Displaced Persons (IDPs) returned to their homes, they hesitated to send their children to schools that were damaged or destroyed. Respondent R13 shared this concern:

"*The extremists were against the current educational system… they thought it was promoting the Western agenda and was un-Islamic. Therefore, they destroyed schools, killed teachers, banned women's*

education, and wanted to replace it with the Islamic education system. After the return of the displaced people, enrollment in schools dropped because the parents were frightened and could not risk their kids' lives by sending them back to school" (Respondent R13).

Due to the fear of Taliban attacks and the absence of protective boundary walls, many parents were hesitant to send their kids to school. Approximately 55,000 students lost an academic year since most schools lacked fencing and were destroyed (Ullah et al., 2017). Therefore, the rehabilitation plan's priority was to restore education and renovate the schools. A respondent, R4, expressed the same sentiment.

"*After the unfortunate damage to schools, the projects promptly initiated the reconstruction process while also establishing new schools in rented buildings as per the required demand. Furthermore, they erected fences and boundary walls, provided school uniforms, and arranged recreational and vocational training sessions for both parents and children*" (Respondent R4).

Reiterating it, respondent R6 stated:

"*We focused on both rehabilitation and reconstruction and convinced the parents and children. We reconstructed 122 damaged schools in district Swat*" (Respondent R6).

Respondents R6 and R9 shared similar views about their education restoration initiatives in district Swat's post-conflict era. They said:

"*In urban areas, we have built a classroom capable of accommodating 40 students, a principal's office, a clerk's office, a stand for cycles, and a security hut for schools. Additionally, we have rented some buildings to establish community schools and constructed a two-room school where necessary. Our efforts also include providing these schools with staff, furniture, ICT, and solar systems*" (Respondent R6, R9).

Efforts were made by the government and project officials to restore education in post-conflict Swat. However, rebuilding the schools and bringing children back to school proved to be a challenging task. Community schools were established in rented buildings to address these challenges, with assurance given to parents of their children's safety. Negative propaganda by the Taliban concerning modern school education and the un-Islamic nature of female education was discouraged through local media and Friday masjids' sermons. Additionally, refreshments, entertainment, and sports were organized in schools to provide a sense of security and inclusion for traumatized parents and children. These initiatives helped restore confidence and trust in the education system. The resilience and capacity to recover from adversities

underscore the inevitability of challenges and the significance of responses to them (Carlson et al., 2012). It was particularly pertinent to the rehabilitation endeavors for IDP returnees in Swat, fostering resilience through infrastructure development and community inclusion in resolving their problems.

However, despite a decade of conflict, many schools remain unreconstructed. According to statistics from the Elementary and Secondary Education Department (2016), the government of Pakistan still needs to rebuild several schools in the Swat district. For instance, a report by "Pakistan: Plight of IDPs" (2010) indicated that out of the 94 destroyed and damaged girls' schools, only 79 had been reconstructed. Furthermore, while the government focused on rebuilding damaged schools, establishing new schools was put on hold, indicating that the education system in Swat is still not fully restored.

Unfulfilled Donors' Commitment and Budget Allocation Challenges

Despite making solid commitments to provide funds for rehabilitation, some donors failed to follow through, leaving the humanitarian efforts of those involved severely underfunded. Despite persistent requests for donor and project details, respondents were hesitant to share information. It became evident that certain donor organizations focused on short-term relief efforts rather than sustainable development, resulting in incomplete and underfunded compensation and rebuilding

processes, as reported by Reza (2010). The United Nations (2010) also found that only 24% of the $537.7 million intended for humanitarian operations was received. As a result, some humanitarian agencies were forced to shut down or cut back on their projects due to the lack of funding.

According to respondent A-1, some donors who wished to contribute to the rehabilitation process were not familiar with the local culture and requirements. They had unreasonable demands for unfeasible and uneconomical projects that were not well-suited to the local context (Farooq, 2017). The authorities and project team members must be free to carry out the projects without interference or undue influence from donors to ensure the smooth execution of rehabilitation initiatives (Levac et al., 2015). In Swat, for example, some donors were conducting quality checks on drinking water samples to ensure they were safe for human consumption (Khan et al., 2018). While this improves effectiveness, impatience with the slow pace of reforms and frustration can delay rehabilitation. It is crucial for authorities to maintain the integrity of rehabilitation projects and allow the project team to perform their responsibilities without pressure from donors (Guy et al., 2013). Maintaining a good relationship with them and involving them in strategic decision-making or changes to the project plan is essential to fulfilling the donor's commitment. The government was facing a significant challenge in deciding how to allocate funds. They had to choose between spending the developmental budget

on relief and rehabilitation or constructing new physical infrastructure. However, this decision posed a risk of neglecting a crucial aspect of reconstruction and reinstatement. Respondent R2 shared their thoughts on the matter.

> "*The government's primary goal was to ensure that people received prompt resolution and reinstatement. However, financing and fund diversion posed a significant risk. This dilemma necessitated obtaining approvals and simplifying audits*" (Respondent R2).

During disasters, the government focuses on relief and rehabilitation rather than new development initiatives because restoring the affected people requires significant funds. For example, instead of building new schools, the government diverted funds towards restoring education by opening community schools in rented buildings. Additionally, they prioritized temporary side roads and bridges over constructing new ones. To deal with post-displacement reconstruction, other countries such as China, Brazil, Canada, Columbia, and Japan made arrangements for additional investments (Cernea, 2008). However, the diversion of development funds to relief efforts required prior approval from the concerned authorities to avoid audit objections and complexities. This delay in the fund diversion process affected the overall rehabilitation of returnees and the region's new developmental

schemes. Rationalizing and reprioritizing the existing development budget could free up some fiscal space. Still, a higher fiscal deficit would have adverse fiscal and economic implications for the future (Asian Development Bank, 2010). In Swat's rehabilitation case, the government of Pakistan assigned a separate fund for future rehabilitation to avoid diverting the existing development budget in case of future disasters. This move on the part of the Pakistani government complements and endorses the good governance doctrine of Kukrit within the realm of rehabilitation in district Swat. Well-organized and planned initiatives ensured the judicious allocation of resources, clarity in the execution of projects, and citizens' active engagement in rehabilitation and reinstatement.

Disempowerment of the Local Government and Empowerment of the Security Agencies

The government's primary responsibility is to ensure the safety and well-being of internally displaced persons (IDPs) (Khan et al., 2018). Complying with this, the government of Pakistan, in collaboration with donor agencies, reinstates IDPs through multi-layered rehabilitation projects (Patil et al., 2017; Sáez et al., 2018). Due to the unstable security situation in the Swat district, security agencies assumed full control, previously vested in the local government. This shift in power posed significant challenges to the implementation of rehabilitation initiatives. Respondents R12 and R15 emphasized the impact of

this disempowerment on local officials. They said:

> "*The escalation of security threats by militants led to the Pakistan army being granted absolute authority over the region under the Civil Aid Action Regulation 2009, significantly reducing the local government's role in rehabilitation efforts and beyond. This unilateral decision prevented a potentially more effective collaborative approach between security agencies and the local government, which could have better served the Swat community. The agencies, inexperienced in community service and primarily versed in military operations, struggled to adopt a more nuanced approach to governance and community engagement*" (Respondent R12, R15).

Other study participants concurred that the local government's participation was instrumental in cost-effective rehabilitation. Azhar (2022) found that cost-effective disaster countermeasures include a holistic, integrative approach to governance that involves minimal bureaucracy and attempts to leverage local expertise. They argued for full engagement due to the government's intimate understanding of the region's culture, community, and dynamics. Such active involvement would have ensured the alignment of rehabilitation efforts with local needs and norms, optimizing project prioritization. Moreover, the local government's role in collecting information on internally displaced persons (IDPs) returning could prevent project redundancies and omissions. Their comprehensive knowledge of local traditions and ability to foresee risks significantly contribute to project success (Cruz & de Brito, 2015). In Swat, however, the local government faced operational constraints due to security concerns and the risk of information leakage to militants. Nonetheless, they implemented notable safety measures, including arming teachers, employing security guards for schools, and enhancing school infrastructures.

CONCLUSION

This study comprehensively examines the challenges and successes in rehabilitating and reintegrating Internally Displaced Persons (IDPs) in Swat, Pakistan, intertwining Kukrit Pramoj's good governance theory with resilience and conflict transformation frameworks. Utilizing the Delphi method and qualitative data analysis, numerous obstacles and facilitators to effective rehabilitation are identified. Key hurdles include security threats, bureaucratic delays, corruption, inadequate infrastructure, educational deficiencies, weakened local governance, financial limitations, and unmet donor commitments. On the flip side, effective strategies include improved security measures, infrastructure development, NGO collaboration, community involvement, and substantial international donor support, particularly from the USA and the UAE.

The research further details the positive impacts of these efforts, citing increased life

satisfaction, psychological well-being, and community participation among returnees, based on Rafiq and Hassan (2022). However, it acknowledges critiques, such as the insufficient financial compensation and lack of focus on humanitarian needs identified by Khan and Wei (2016), advocating for a more comprehensive rehabilitation approach. The study concludes by recommending regular evaluations of rehabilitation programs, incorporating community feedback and well-being indicators. It stresses the significance of collaborative efforts aligned with good governance and community engagement, involving community leaders, NGOs, and local governments. The insights garnered extend beyond Pakistan, offering valuable lessons for developing countries on sustainable development, peace, and disaster management strategies, enriching the global discourse on forced migration and rehabilitation.

Policy Implications and Recommendations

Despite significant displacement issues, Pakistan lacks a comprehensive policy for the rehabilitation of internally displaced persons (IDPs). However, it has implemented the Agency Standing Committee Framework for Durable Solutions for IDPs. This framework, anchored in international law, aims to address the rights and needs of IDPs without discrimination and ensure long-term solutions (Ali & Kandhro, 2015; Jaung et al., 2017; Mohsin, 2013; Sharma, 2015). Moreover, collaboration with the United Nations Development Program (UNDP) has been initiated to assist in rehabilitation and relocation efforts, aligning Pakistan with UN rehabilitation recommendations, specifically Principles 28-30 (Ali & Kandhro, 2015). Nonetheless, there are significant gaps in the implementation of these principles and domestic legislation concerning internal displacement (Ahmad et al., 2017).

This study on the rehabilitation of IDP returnees in Swat, Pakistan, reveals significant program challenges and proposes targeted policy measures to improve rehabilitation management. Key recommendations include:

1. Establishing a clear rehabilitation policy with defined roles for local governments, administrative bodies, and stakeholders.
2. Creating a dedicated fund to protect development budgets.
3. Involving local governments in projects to leverage their cultural insights.
4. Fostering local community involvement in project selection to reduce political bias and nepotism.
5. Setting up a monitoring team to ensure project integrity, track finances, and assess needs.
6. Streamlining NGOs' No Objection Certificate process to expedite rehabilitation efforts.
7. Implementing a centralized coordination system to avoid project overlap.
8. Emphasizing expert input in early planning to ensure efficient use of resources.

9. Increasing transparency and expert oversight in project planning and execution.

10. Ensuring timely and clear communication with donors for effective fund management and project adjustments.

These strategies aim to enhance the IDP rehabilitation framework in Pakistan through greater transparency, accountability, and community involvement, addressing existing challenges more effectively.

Limitations of the Study

This study's primary limitation lies in its narrow focus on the perspectives of returnees exclusively from Swat. Future studies should extend their scope to include returnees from additional regions such as Dir Lower, Bunner, Shangla, Bajaur, South Waziristan, and North Waziristan to better understand attitudes towards rehabilitation and reinstatement. Moreover, the findings predominantly reflect male viewpoints, potentially overlooking the unique perspectives of women on these issues. Future research should endeavor to incorporate women's perspectives. Additionally, expanding the investigation to cover diverse geographical areas and various socioeconomic strata within society could enrich the understanding of the subject matter.

ACKNOWLEDGEMENT

REFERENCES

Abbasi-Shavazi, M. J., & Sadeghi, R. (2015). Socio-cultural adaptation of second-generation Afghans in Iran. *International Migration, 53*(6), 89-110. https://doi.org/10.1111/imig.12148

Abdullah, J., Sayuti, N. M., Arshad, A. A. M., & Embong, M. R. (2016). Living conditions in orang asli resettlement project (PROSDET) of Pantos, Pahang, Malaysia. *Procedia-Social and Behavioral Sciences, 222*, 143-150. https://doi.org/10.1016/j.sbspro.2016.05.204

Aburas, R., Najeeb, A., Baageel, L., & Mackey, T. K. (2018). The Syrian conflict: A case study of the challenges and acute need for medical humanitarian operations for women and children internally displaced persons. *BMC medicine, 16*(1), 1-6. https://doi.org/10.1186/s12916-018-1041-7

Adam, I. F. (2015). The influence of maternal health education on the place of delivery in conflict settings of Darfur, Sudan. *Conflict and Health, 9*(1), 1-9. https://doi.org/10.1186/s13031-015-0057-2

Adekola, P. O., Cirella, G. T., & Brownell, G. (2024). Reintegration programs and the willingness of displaced persons to return home: Analyzing the role of social infrastructure in North-East Nigeria. *Journal of International Migration and Integration*, 25, 1351-1380. https://doi.org/10.1007/s12134-024-01128-x

Adnan, R. (2012). *How bad governance led to conflict: The Case of Swat, Pakistan.* Center for Public Policy and Governance. https://mpra.ub.uni-muenchen.de/35357/

Ahmad, S. (2018). Pakistan and international principles in conflict-induced internal displacement, 2009-2015. *Pakistan Horizon, 71*(4), 61-79.

Ahmad, S., Ghufran, N., & Usman, T. (2017). Conflict-induced internally displaced persons in pakistan and issues of their protection: A case of displaced Pakhtuns & the United Nations guiding

principles on internal displacement. *Frontier Women University Journal of Social Sciences,* (3), 67-78.

Akhunzada, W. A., Qadir, A., Maqsood, N., Rasool, G., & Rahman, R. U. (2015). Internally displaced persons (IDPS): Anxiety and depression. *The Professional Medical Journal, 22*(03), 337-342. https://doi.org/10.29309/TPMJ/2015.22.03.1352

Akilova, M., Borh, K. J., & Marzouk, H. A. (2022). Durable solutions: Return and reintegration of displaced populations and reconstruction in post-conflict societies. In N. J. Murakami & M. Akilova (Eds.), *Integrative social work practice with refugees, asylum seekers, and other forcibly displaced persons* (pp. 199-228). Springer. https://doi.org/10.1007/978-3-031-12600-0_9

Al-Qawasmi, J., Saeed, M., Asfour, O. S., & Aldosary, A. S. (2021). Assessing urban quality of life: Developing the criteria for Saudi cities. *Frontiers in Built Environment, 7*, Article 682391. https://doi.org/10.3389/fbuil.2021.682391

Ali, R., & Kandhro, S. H. (2015). National disaster management authority in Pakistan: Role of Pakistan army in disaster management. *Journal of Social and Administrative Sciences, 2*(1), 11-17.

Alifa, S., & Nugroho, F. (2019). The role of local community enterprise towards economic recovery of disaster-affected community in Indonesia. *Pertanika Journal of Social Sciences and Humanities, 27*(4), 2333-2349.

Anwar, Z., & Ahmad, W. (2017). Augmenting the conflict in Swat: Loopholes in legal governance. *Journal of Political Studies, 24*(2), 417 - 432.

Asian Development Bank. (2010). *Pakistan floods 2010: Preliminary damage and needs assessment.* https://www.adb.org/sites/default/files/linked-documents/44372-01-pak-oth-02.pdf

Azhar, A. (2022). Disaster management reforms in Pakistan. In A. Zahra, G. Bouckaert, M.

J. I. Zafar & N. Jabeen (Eds.), *Public sector reforms in Pakistan: Hierarchies, markets and networks* (pp. 329-348). Springer. https://doi.org/10.1007/978-3-030-96825-0_14

Azim, S. W., Mehmood, W., & Hussain, S. (2018). Swat conflict in retrospect: Violence and jarga among the Swat Pukhtuns in Pakistan. *Liberal Arts and Social Sciences International Journal (LASSIJ), 2*(1), 37-48. https://doi.org/10.47264/idea.lassij/2.1.5

Bain, A., & Parkinson, G. (2010). Resettlement and social rehabilitation: Are we supporting success? *Probation Journal, 57*(1), 63-74. https://doi.org/10.1177/0264550509354669

Bangash, S. (2012). Socio-economic conditions of post-conflict Swat: A critical appraisal. *A Journal of Peace and Development, 11*(2), 66-79.

Bashir, M. F., Ain, S. Q. U., Tariq, Y. B., & Iqbal, N. (2022). Impact of governance structure, infrastructure, and terrorism on tourism. *Pertanika Journal of Social Sciences & Humanities, 30*(4), 1607 - 1633. https://doi.org/https://doi.org/10.47836/pjssh.30.4.08

Bhatti, M. I., & Shah, A. A. (2017). Meaning in life and psychological well-being among flood victims in Pakistan: The moderating role of resilience. *Bahria Journal of Professional Psychology, 16*(2), 73 - 86.

Boadi, E. B., Chen, S., Amponsah, E. I., & Appiah, R. (2022). Antecedents of residential satisfaction in resettlement housing in Ellembelle: A PLS-SEM approach. *Sustainability, 14*(18), 1 - 13. https://doi.org/10.3390/su141811256

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Bree, R. T., & Gallagher, G. (2016). Using Microsoft Excel to code and thematically analyse qualitative data: A simple, cost-effective approach. *All*

Ireland Journal of Higher Education, 8(2), 2811 - 2819.

Cantor, D., Swartz, J., Roberts, B., Abbara, A., Ager, A., Bhutta, Z. A., Blanchet, K., Bunte, D. M., Chukwuorji, J. C., & Daoud, N. (2021). Understanding the health needs of internally displaced persons: A scoping review. *Journal of migration and health, 4,* Article 100071. https://doi.org/10.1016/j.jmh.2021.100071

Carlson, J., Haffenden, R., Bassett, G., Buehring, W., Collins III, M., Folga, S., Petit, F., Phillips, J., Verner, D., & Whitfield, R. (2012). *Resilience: Theory and application.* Argonne National Laboratory. https://publications.anl.gov/anlpubs/2012/02/72218.pdf

Cernea, M. M. (2008). Compensation and benefit sharing: Why resettlement policies and practices must be reformed. *Can compensation prevent impoverishment, 1*(1), 89-120. https://doi.org/10.1016/S1674-2370(15)30021-1

Chan, K. L. G., Hamid, B. A., & Selvadurai, S. (2021). Social inclusion and exclusion of women directors in Malaysia. *Pertanika Journal of Social Sciences & Humanities, 29*(4), 2483-2501. https://doi.org/10.47836/pjssh.29.4.22

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. SAGE Publications. https://books.google.com.sa/books?id=eTwmDwAAQBAJ

Cruz, C. O., & de Brito, J. (2015). Role of public administration in fostering urban housing rehabilitation. *Journal of Urban Planning and Development, 141*(4), Article 06014002. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000228

Danquah, J. A., Attippoe, J., & Ankrah, J. S. (2014). Assessment of residential satisfaction in the resettlement towns of the Keta basin in Ghana. *International Journal Civil Engineering, Construction and Estate Management, 2*(3), 26-45.

Delay in construction of bridge irks Swat people. (2013, June 19). *Dawn.* https://www.dawn.com/news/1019225/delay-in-construction-of-bridge-irks-swat-people

Din, N. U. (2010). *Internal displacement in Pakistan: Contemporary challenges.* https://dev.humanitarianlibrary.org/sites/default/files/2014/02/22.pdf

Edge, S., Newbold, K. B., & McKeary, M. (2014). Exploring socio-cultural factors that mediate, facilitate, & constrain the health and empowerment of refugee youth. *Social Science & Medicine, 117*, 34-41. https://doi.org/10.1016/j.socscimed.2014.07.025

Elahi, N. (2015). Militancy conflicts and displacement in Swat Valley of Pakistan: Analysis of transformation of social and cultural network. *International Journal of Humanities and Social Science, 5*(3), 226 - 236. http://hdl.handle.net/11250/2576231

Elahi, N., Nyborg, I. L., & Nawab, B. (2015). Participatory development practices: A critical analysis of gender empowerment and development in pre-and post-crises Swat, Pakistan. *Forum for Development Studies, 42*(2), 333-356. https://doi.org/10.1080/08039410.2015.1025828

Elementary and Secondary Education Department. (2016). *Elementary and secondary education department Khyber Pakhtunkhwa*. https://www.kpese.gov.pk/

Farooq, S. (2017). Rehabilitation of 2010 flood affectees in Pakistan: The role of Development Partners. *NUST Journal of Social Sciences and Humanities, 3*(2), 194-231. https://doi.org/10.51732/njssh.v3i2.23

Guy, A. J., Curnoe, D., & Banks, P. B. (2013). A survey of current mammal rehabilitation and release practices. *Biodiversity and Conservation, 22*(4), 825-837. https://doi.org/10.1007/s10531-013-0452-1

Haider, Z. (2009). How did Islamist militancy emerge in Pakistani paradise. *Reuters.* https://www.reuters.com/article/us-pakistan-swat-idUKTRE51F21620090216

Hashmi, A. S. (2016). Religious radicalisation in Pakistan: Defining a common narrative. *Strategic Studies, 36*(3), 58-77.

Herath, D., Lakshman, R. W., & Ekanayake, A. (2017). Urban resettlement in Colombo from a wellbeing perspective: Does development-forced resettlement lead to improved wellbeing? *Journal of Refugee Studies, 30*(4), 554-579. https://doi.org/10.1093/jrs/few043

İçduygu, A. (2015). Syrian refugees in Turkey: The long road ahead. *Reliefweb.* https://reliefweb.int/report/turkey/syrian-refugees-turkey-long-road-ahead-april-2015?gad_source=1&gclid=CjwK-CAjwoJa2BhBPEiwA0l0ImBqZDv6arCgOK5ANo6trNMjA3gBXTdjUuPRNb7KgDE1Uu6LU7WgTvRoCxU4QAvD_BwE

Internal Displacement Monitoring Centre. (2012). *Pakistan: Displacement caused by conflict and natural disasters, achievements and challenges*. https://reliefweb.int/report/pakistan/displacement-caused-conflict-and-natural-disasters-achievements-and-challenges

Internal Development Monitoring Centre. (2018). *Global report on internal 2018 displacement*. https://www.internal-displacement.org/global-report/grid2018/downloads/2018-GRID.pdf

Jaung, M., Jani, S., Banu, S., & Mackey, J. M. (2017). International emergency psychiatry challenges: Disaster medicine, war, human trafficking, displaced persons. *Psychiatric Clinics, 40*(3), 565-574. https://doi.org/10.1016/j.psc.2017.05.015

Kanishchev, V., Lyamin, S., & Zhukov, D. (2022). Demographic strategies of the Russian rural population: Intersubjective nonlinearity versus economic viability. *Pertanika Journal of Social Sciences & Humanities, 30*(2), 579 - 604. https://doi.org/10.47836/pjssh.30.2.09

Khalid, A., & Helge, L. (2010). *The main causes of the breakdown of governance and rise of militancy.* Norwegian Institute of International Affairs. https://www.files.ethz.ch/isn/121829/PB-06-10-Lurås%20and%20Aziz.pdf

Khaliq, F. (2010). Locals want bridge reconstructed before winter rains. *The Express Tribune*. https://tribune.com.pk/story/81434/locals-want-bridge-reconstructed-before-winter-rains

Khan, H., Khan, K., & Yousafzai, A. (2017). A sequential explana-tory mixed method study to identify human rights violations and asses the psychosocial health status of the internally displaced women of North-West Pakistan. *Journal of Community Medicine & Public Health Care, 4*, Article 27. https://doi.org/10.24966/CMPH-1978/100027

Khan, I. (2024). The resurgence of militant groups in Pakistan's borderlands: Analyzing shortcomings in rehabilitation and reintegration strategies. *Journal for Deradicalization* (38), 1-27.

Khan, K. S. (2010). Analysing local perceptions of post-conflict and post-floods livelihood interventions in Swat, Pakistan. *Development Policy Review, 32*(s2), 274-292. https://doi.org/10.1111/dpr.12421

Khan, K., Lu, Y., Saeed, M. A., Bilal, H., Sher, H., Khan, H., Ali, J., Wang, P., Uwizeyimana, H., & Baninla, Y. (2018). Prevalent fecal contamination in drinking water resources and potential health risks in Swat, Pakistan. *Journal of Environmental Sciences, 72*, 1-12. https://doi.org/10.1016/j.jes.2017.12.008

Khan, M. S., & De Nardi, S. (2024). The affectual-social ecology of cultural artefacts: Illegal markets and religious vandalism in Swat Valley, Pakistan. *Forum for Social Economics, 53*(1), 35-52. https://doi.org/10.1080/07360932.2022.2084440

Khan, M. S., & Sepulveda, L. (2022). Conflict, displacement, and economic revival: The case of

the internally displaced minority entrepreneurs in Pakistan. *Strategic Change, 31*(4), 461-477. https://doi.org/10.1002/jsc.2515

Khan, M. K., & Wei, L. (2016). When friends turned into enemies: The role of the national state vs. Tehrik-i-Taliban Pakistan (TTP) in the war against terrorism in Pakistan. *Korean Journal of Defense Analysis, 28*(4), 597-626.

Khan, Z. A. (2011). Military operations in FATA and PATA. *Strategic Studies, 31*, 129-146.

Kinyanjui, B., Umeasiegbu, V. I., & Bishop, M. L. (2017). Rehabilitation needs of refugees with disabilities in the United States: Implications for vocational rehabilitation counselors. *Journal of Applied Rehabilitation Counseling, 48*(2), 5-14. https://doi.org/10.1891/0047-2220.48.2.5

Kiren, A. C. (2016). *Economic calculation and strategies among resettled IDPs (SWAT) and current IDPS (Peshawar and Bannu), Pakistan.* International Growth Centre. https://www.theigc.org/publication/economic-calculation-and-strategies-among-resettled-idps-swat-and-current-idps-peshawar-and-bannu-pakistan/

Lam, E., McCarthy, A., & Brennan, M. (2015). Vaccine-preventable diseases in humanitarian emergencies among refugee and internally-displaced populations. *Human vaccines & immunotherapeutics, 11*(11), 2627-2636. https://doi.org/10.1080/21645515.2015.1096457

Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological forecasting and social change, 73*(5), 467-482. https://doi.org/10.1016/j.techfore.2005.09.002

Levac, D., Glegg, S. M., Camden, C., Rivard, L. M., & Missiuna, C. (2015). Best practice recommendations for the development, implementation, and evaluation of online knowledge translation resources in rehabilitation. *Physical therapy, 95*(4), 648-662. https://doi.org/10.2522/ptj.20130500

Loescher, G. (2017). UNHCR's origins and early history: Agency, influence, and power in global refugee policy. *Refuge: Canada's Journal on Refugees, 33*(1), 77-86. https://doi.org/10.25071/1920-7336.40450

Lutfullah, S., Rasheed, A. F., & Syed, N. B. B. (2022). English 1. judicial system of the Swat State in the Regime of Wali Swat and its appraisal in the light of Islamic law (Sharī'ah) (1917-1969). *Al Khadim Research Journal of Islamic Culture and Civilization, 3*(2), 89-100. https://doi.org/10.53575/arjicc.v3.02(22)e1.1-15

Ma, Y., Ariffin, N. F. M., Aziz, F. A., He, X., Liu, Y., & Feng, S. (2023). Potential of sense of place in cultural heritage conservation: A systematic review. *Pertanika Journal of Social Sciences & Humanities, 31*(4), 1465 - 1489. https://doi.org/10.47836/pjssh.31.4.07

Madeeha, A. (2017). NGOs: Pakistan government restrictions hamper campaigns to curb militancy. *Voice of America.* https://www.voanews.com/a/pakistan-ngos-red-tape/3733512.html

Magsi, H., & Torre, A. (2014). Proximity analysis of inefficient practices and socio-spatial negligence: Evidence, evaluations and recommendations drawn from the construction of Chotiari reservoir in Pakistan. *Land Use Policy, 36*, 567-576. https://doi.org/10.1016/j.landusepol.2013.10.009

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative health research, 26*(13), 1753-1760. https://doi.org/10.1177/1049732315617444

Mason, P., Augustyn, M., & Seakhoa-King, A. (2010). Exploratory study in tourism: Designing an initial, qualitative phase of sequenced, mixed methods research. *International Journal of Tourism Research, 12*(5), 432-448. https://doi.org/10.1002/jtr.763

Mohsin, Z. R. (2013). The crisis of internally displaced persons (IDPs) in the federally

administered tribal areas of Pakistan and their impact on Pashtun women. *Tigah: A Journal of Peace and Development, 3*(2), 92-117.

Molnar, P. (2017). Discretion to deport: Intersections between health and detention of Syrian refugees in Jordan. *Refuge: Canada's Journal on Refugees/ Refuge: Revue canadienne sur les réfugiés, 33*(2), 18-31. https://doi.org/10.7202/1043060ar

Morina, N., Akhtar, A., Barth, J., & Schnyder, U. (2018). Psychiatric disorders in refugees and internally displaced persons after forced displacement: A systematic review. *Frontiers in psychiatry, 9*, Article 433. https://doi.org/10.3389/fpsyt.2018.00433

Muhammad, I., Asadullah, K., Mitsuhiro, I., Muhammad, A., & Hassan, S. (2011). Identifying factors affecting agroforestry system in Swat, Pakistan. *African Journal of Agricultural Research, 6*(11), 2586-2593.

Mujtaba, M. R. (2013). *The UNDP development programme in Swat: Impact on peacebuilding and social cohesion.* https://info.undp.org/docs/pdc/Documents/PAK/Evaluation%20Report-%20CR%20Project-%202014.pdf

Murphy, A., Chikovani, I., Uchaneishvili, M., Makhashvili, N., & Roberts, B. (2018). Barriers to mental health care utilization among internally displaced persons in the republic of Georgia: a rapid appraisal study. *BMC health services research, 18*(1), 1-11. https://doi.org/10.1186/s12913-018-3113-y

Office for the Coordination of Humanitarian Affairs. (2013). *Humanitarian access in Pakistan.* https://www.unocha.org/publications/report/pakistan/humanitarian-access-pakistan

Ogunode, J., Chijindu, O. E., & Jegede, D. (2022). Provision of education services for internally displaced persons in IDPs camps in Nigeria: Challenges and way forward. *International Journal on Integrated Education, 5*(5), 14-22. https://doi.org/10.31149/ijie.v5i5.3011

Pakistan: Plight of IDPs. (2010, January 9). *Reliefweb.* https://reliefweb.int/report/pakistan/pakistan-plight-idps

Patil, V., Ghosh, R., & Kathuria, V. (2017). The role of access mechanisms in effective rehabilitation of displaced farmers due to development projects. *The Journal of Development Studies, 53*(4), 548-564. https://doi.org/10.1080/00220388.2016.1187725

Peracha, F. N., Khan, R. R., Ahmad, A., Khan, S. J., Hussein, S., & Choudry, H. R. (2012). Socio demographic variables in the vulnerable youth predisposed towards militancy (Swat, Pakistan). *Psychiatry, Psychology and Law, 19*(3), 439-447. https://doi.org/10.1080/13218719.2011.598635

Peterson, R. B., Kapiyo, R. A., Campbell, E. M., & Nyabua, P. O. (2018). Gully rehabilitation trusts: Fighting soil erosion through community participation in western Kenya. *Journal of rural studies, 58*, 67-81. https://doi.org/10.1016/j.jrurstud.2017.12.028

Provincial Disaster Management Authority. (2019). *A disaster resilient Khyber Pakhtunkhwa.* https://www.pdma.gov.pk/

Rafiq, M., Hassan, A. A. G., & Saeed, M. (2021). The reinstatement of returnees in District Swat, Pakistan: An evaluative study of the rehabilitation initiatives. *Social Sciences, 10*(12), 476. https://doi.org/10.3390/socsci10120476

Rafiq, M., Hassan, A. A. G., & Saeed, M. (2022). Perceptions of returnees concerning their rehabilitation and reinstatement in Swat District, Pakistan: An evaluative study. *Journal of Population and Social Studies [JPSS], 30*, 354-376. https://doi.org/10.25133/JPSSv302022.021

Rafiq, M., Saeed, M., & Israr, M. (2022). A qualitative study of the factors driving the temporary migration of workers from the Khyber-Pakhtunkhwa, Pakistan to Saudi Arabia. *Journal of Humanities, Social and Management*

Sciences (JHSMS), 3(1), 169-186. https://doi.org/10.47264/idea.jhsms/3.1.12

Randell, H. (2016). The short-term impacts of development-induced displacement on wealth and subjective well-being in the Brazilian Amazon. *World development, 87*, 385-400. https://doi.org/10.1016/j.worlddev.2016.07.005

Reza, J. (2010). Paradise regained: Swat one year on. *Chritical Threats*. https://www.criticalthreats.org/analysis/paradise-regained-swat-one-year-on

Saeed, M. (2012). *Pakhtun men's perceptions of the conditions promoting domestic violence in their culture* [Unpublished doctoral dissertation]. University of York. https://core.ac.uk/download/pdf/5225494.pdf

Saeed, M., & Griffin, G. (2019). Researching a sensitive topic in an unstable environment: Fieldwork dilemmas in the Khyber-Pakhtunkhwa province of Pakistan. *Qualitative Research Journal, 19*(3), 248-258. https://doi.org/10.1108/QRJ-02-2019-0013

Sáez, P. V., Astorqui, J. S. C., del Río Merino, M., Moyano, M. d. P. M., & Sánchez, A. R. (2018). Estimation of construction and demolition waste in building energy efficiency retrofitting works of the vertical envelope. *Journal of Cleaner Production, 172*, 2978-2985. https://doi.org/10.1016/j.jclepro.2017.11.113

Salgado-Gálvez, M. A. (2018). Estimating the lost economic production caused by internal displacement because of disasters. *International Journal of Disaster Risk Science, 9*(4), 496-506. https://doi.org/10.1007/s13753-018-0190-9

Sandelowski, M. (1995). Sample size in qualitative research. *Research in nursing & health, 18*(2), 179-183. https://doi.org/10.1002/nur.4770180211

Sayed, S. A., & González, P. A. (2014). Flood disaster profile of Pakistan: A review. *Science Journal of Public Health, 2*(3), 144-149. https://doi.org/10.11648/j.sjph.20140203.11

Sayeed, S., & Shah, R. (2017). *Displacement, repatriation and rehabilitation: Stories of dispossession from Pakistan's frontier'*. https://www.swp-berlin.org/publications/products/arbeitspapiere/Sayeed_and_Shah_2017_Internal_Displacement_Pakistan.pdf

Serghiou, M. A., Niszczak, J., Parry, I., Li-Tsang, C., Van den Kerckhove, E., Smailes, S., & Edgar, D. (2016). One world one burn rehabilitation standard. *Burns, 42*(5), 1047-1058. https://doi.org/0.1016/j.burns.2016.04.002

Shah, A. A., Ye, J., Abid, M., Khan, J., & Amir, S. M. (2018). Flood hazards: Household vulnerability and resilience in disaster-prone districts of Khyber Pakhtunkhwa province, Pakistan. *Natural hazards, 93*, 147-165. https://doi.org/10.1007/s11069-018-3293-0

Shaheen, A. (2011). The emergence of conflict in swat valley and the post-conflict management issues. *Journal of the Pakistan historical society, 59*(3), 19-59.

Shami, S. (1993). The social implications of population displacement and resettlement: An overview with a focus on the Arab Middle East. *International Migration Review, 27*(1), 4-33. https://doi.org/10.1177/019791839302700101

Sharma, B. B. (2015). Revisiting the United Nations' 1951 convention relating to the status of refugees: A critical analysis of the international refugee law. *Social Development Issues, 37*(2), Article 80.

Sohail, M., Muhammad, S., Mehmood, K., Anees, S., Rabbi, F., Tayyab, M., Hussain, K., Hayat, M., & Khan, U. (2023). Tourism, threat, and opportunities for the forest resources: A case study of Gabin Jabaa, District Swat, Khyber-Pakhtunkhwa, Pakistan. *International Journal of Forest Sciences, 3*(3), 194-203. https://doi.org/10.5281/zenodo.10442867

Swaroop, S. R., & DeLoach, C. D. (2015). Voices of trauma and resilience: Cultural and gender distinctive responses to war and displacement in Pakistan. *Psychology and developing societies, 27*(1), 1-30. https://doi.org/10.1177/0971333614564743

Townsley, P. (1996). *Rapid rural appraisal, participatory rural appraisal and aquaculture.* Food and Agriculture Organization. https://www.fao.org/4/w2352e/W2352E00.htm

Ulke, C., Gfesser, T., Fleischer, T., Altweck, L., Hahm, S., Mühlan, H., Heller, A., Beutel, M. E., Schmidt, S., & Grabe, H. J. (2021). Later-life depressive symptoms and anxiety attacks in displaced and nondisplaced populations. *Journal of affective disorders reports, 3*, Article 100061. https://doi.org/10.1016/j.jadr.2020.100061

Ullah, A., Khan, K., & Mahmood, H. (2017). Understanding the nexus between conflict, displacement, and schooling: A case study of IDPs in Pakistan. *Asian Journal of Peacebuilding, 5*(1), 155-168. https://doi.org/10.18588/201705.00a010

United Nations. (2010). *Crisis persists in Pakistan's Swat Valley despite return of many displaced – UN.* https://news.un.org/en/story/2010/04/337062

Urbański, M. (2022). Comparing push and pull factors affecting migration. *Economies, 10*(1), Article 21. https://doi.org/10.3390/economies10010021

Waters, T. (2022). MR Kukrit Pramoj's theory of good governance and political change: The dialectics of Farang Sakdina. *Humanities and Social Sciences Communications, 9*(1), 1-11. https://doi.org/10.1057/s41599-022-01158-9

Wistrand, J. S. (2023). A development approach to a protracted IDP Situation: Lessons from sAzerbaijan. *Journal on Migration and Human Security, 11*(1), 23-40. https://doi.org/10.1177/23315024231158558

Xu, J., Shi, G., Li, B., Fischer, T. B., Zhang, R., Yan, D., Jiang, J., Yang, Q., & Sun, Z. (2021). Skills' sets and shared benefits: Perceptions of resettled people from the Yangtze-Huai river diversion project in China. *Impact assessment and project appraisal, 39*(5), 429-438. https://doi.org/10.1080/14615517.2020.1848242

Zahid, A. (2012). Pakistan: Challenges of conflict-induced displacement. *Peace Insight.* https://www.peaceinsight.org/en/articles/pakistan-the-challenges-of-conflict-induced-displacement-of-people/?location=pakistan&theme=development

APPENDICES

Appendix 1: Table of organizations involved in the rehabilitation process

S.N0	Code Name	Pseudonym Name	Organization Name
1	R1	Liaqat	PDMA (Provincial Disaster Management Authority)/Relief Dep
2	R2	Nawaz	PDMA(Provincial Disaster Management Authority)
3	R3	Ali	PDMA(Provincial Disaster Management Authority)
4	R4	Ikram	Directorate of Education /UNICEF
5	R5	Ahmad	UNICEF
6	R6	Shah	UNDP
7	R7	Lal	EPS (Environmental Protection Society)a
8	R8	Akbar	PARSA (Provincial Reconstruction Rehabilitation and Settlement Authority)
9	R9	Ibrahim	Lasona (the organization name)
10	R10	Noman	Hujra Foundation (the organization name)
11	R11	Mahmood	Tehsil Kabal Nazim
12	R12	Speen	Assistance Commissioner
13	R13	Tariq	ASDO (Assistant District Education Officer)
14	R14	Yousaf	In Charge Principal Medical Officer BHU Kabal
15	R15	Adnan	Political Leader

Appendix 2: Field trip visuals

REFEREES FOR THE PERTANIKA
JOURNAL OF SOCIAL SCIENCES & HUMANITIES

Vol. 32 (3) Sep. 2024

The Editorial Board of the Pertanika Social Sciences & Humanities wishes to thank the following:

Irniza Rasdi
(UPM, Malaysia)

Jihad Mohammad
(UCSI, Malaysia)

John Robert Rilveria
(UPD, Philippines)

Juliana Abdul Wahab
(USM, Malaysia)

Juraifa Jais
(UNITEN, Malaysia)

Kelly Wong Kai Seng
(UPM, Malaysia)

Kim Ling @ Geraldine Chan
(UKM, Malaysia)

Kris Green
(St. John Fisher University, USA)

Ku Hasnita Ku Samsu
(UPM, Malaysia)

Latifat Kehinde Olatinwo
(KWASU, Nigeria)

Lim Seong Pek
(UNISEL, Malaysia)

M. Zainora Asmawi
(IIUM, Malaysia)

Made Hery Santosa
(UNDIKSHA, Indonesia)

Madhubala Bava Harji
(SEGi University, Malaysia)

Magdalene Ang Chooi Hwa
(USM, Malaysia)

Mansour Amini
(USM, Malaysia)

Manuel Ángel Fernández Gámez
(UMA, Spain)

Marcelo Coertjens
(UFDPar, Brazil)

Marhana Mohamed Anuar
(UMT, Malaysia)

Melissa Moralli
(UNIBO, Italy)

Mohamad Fadhli Rashid
(UTM, Malaysia)

Mohamad Termizi Borhan
(UPSI, Malaysia)

Mohd Asyraf Mansor
(USM, Malaysia)

Mohd Faiq Abd Aziz
(UPM, Malaysia)

Mohd Faizal Kasmani
(USIM, Malaysia)

Mohd Nihra Haruzuan
Mohamad Said
(UTM, Malaysia)

Mohd Zairul Mohd Noor
(UPM, Malaysia)

Mohd. Mahadee Ismail
(UPM, Malaysia)

Mou Leong Tan
(USM, Malaysia)

Mozhgan Ghassemiazghandi
(USM, Malaysia)

Mua'azam Mohamad
(UUM, Malaysia)

Muafi
(UII, Indonesia)

Muhammad Abrar ul haq
(UTB, Bahrain)

Muhammad Ayub Jan
(UOP, Pakistan)

Muhammad Danial Azman
(UM, Malaysia)

Muhammad Fuad Othman
(UUM, Malaysia)

Muhd Khaizer Omar
(UPM, Malaysia)

Musaev Talaibek
(UM, Malaysia)

Nellie Ismail
(UPM, Malaysia)

Ng Boon Kwee
(UM, Malaysia)

Nor Malina Malek
(USM, Malaysia)

Norliza Kushairi
(UUM, Malaysia)

Norzalina Zainudin
(UPM, Malaysia)

Nuha Loling Othman
(UNIMAS, Malaysia)

Nurahimah Mohd Yusoff
(UUM, Malaysia)

Nurul Liyana Mohd Kamil
(UM, Malaysia)

Ong Chuan Huat
(SEGi University, Malaysia)

Priyanka Tripathi
(IIT Patna, India)

Pui Yee Chong
(UNITEN, Malaysia)

Raja Muhammad Ishtiaq Khan
(MU, Saudi Arabia)

Rajwani Md Zain
(UUM, Malaysia)

Rani Ann Balaraman
(USM, Malaysia)

Reza Kafipour
(SUMS, Iran)

Rita Wong Mee Mee
(UPNM, Malaysia)

Rizal Mohd Razman
(UM, Malaysia)

S. Kanageswari Suppiah
Shanmugam
(UUM, Malaysia)

Saima Afzal
(BZU, Pakistan)

Samsudin Suhaili
(UPSI, Malaysia)

Sanmugam Annamalah
(SEGi University, Malaysia)

Shahanum Md Shah
(UiTM, Malaysia)

Content