ISSN 2708-7581 (Online)
ISSN-L 2708-7573

JLTM

Journal of Learning Theory and Methodology

Scientific journal

August 2026
Volume 7
Number 2

JLTM
LLC OVS

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Journal of Learning Theory and Methodology
Abbreviated key-title: J. learn. theory methodol. (Online)

Scientific journal
Three issues per year. Established in 2020
https://www.ltmjournal.com.E-mail: editor-in-chief@ltmjournal.com

Abstracting and Indexing:

Web of Science Research Commons (analytical visibility); CrossRef; ROAD (Directory of Open Access scholarly Resources); DOAJ (Directory of Open Access Journals); ERIH PLUS; ICI Journals Master List / ICI World of Journals; Google Scholar; Open Ukrainian Citation Index (OUCI); Scilit (A database of scientific & scholarly literature); WorldCat

DOI: https://doi.org/10.17309/jltm.2026.7.2

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Журнал теорії та методології навчання
Abbreviated key-title: J. learn. theory methodol. (Online)

Науковий журнал

Три випуски на рік. Заснований у 2020 році

https://www.ltmjournal.com.E-mail: editor-in-chief@ltmjournal.com

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Journal of Learning Theory and Methodology
Scientific journal
August 2026, Vol. 7, Num. 2

Contents

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Журнал теорії та методології навчання
Науковий журнал
Серпень 2026, Том 7, Номер 2

Зміст

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання
ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Motor Function as a Conceptual Framework for Interpreting Fundamental Movement Skills

Oleg Khudolii[1ABCD], Olha Ivashchenko[1ABCD] and Mykola Khudolii[1ABCD]

[1]Kharkiv State Academy of Physical Culture

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.02

Abstract

Purpose. To substantiate motor function as a functional system underlying the development of human motor activity, to determine the place of fundamental movement skills within its structure, and to develop a conceptual model of motor function development.

Material and Methods. The study was theoretical and methodological in nature and was conducted as a conceptual narrative review. The methodological foundation was based on conceptual analysis and knowledge synthesis. The materials included classical and contemporary scholarly works addressing motor function, motor control, motor development, motor learning, and fundamental movement skills. Historical-logical analysis, content analysis of scientific sources, comparative analysis, systems analysis, and conceptual modeling were employed.

Results. Motor function was conceptualized as an integrative functional system formed through the interaction of motor abilities, motor control mechanisms, sensory support systems, and motor experience. Motor control was identified as the central component of motor function, with its development progressing from the accurate reproduction of movement parameters to the differentiation of spatial, temporal, and force characteristics of movement actions. Fundamental movement skills were interpreted not as the foundation of motor development but as a form of motor function development and a mechanism for the accumulation of motor experience. A conceptual model of motor function development and a functional classification of fundamental movement skills were proposed, comprising stabilization, spatial orientation, temporal organization of movement, motor control, and adaptation. Programming was substantiated as the mechanism for the purposeful development of motor function.

Conclusions. Motor function represents the system-forming construct of motor development, integrating motor abilities, motor control mechanisms, sensory support systems, and motor experience in the process of solving motor tasks. Fundamental movement skills should be interpreted as a form of motor function development, whereas programming serves as the mechanism for its purposeful formation. The proposed framework expands the explanatory potential of contemporary Fundamental Movement Skills, Motor Competence, and Physical Literacy approaches.

Keywords: motor function, fundamental movement skills, motor development, motor control, motor experience, programming, motor learning.

Introduction

Research on human motor development constitutes one of the fundamental domains of motor learning theory, physical education, and sports training. A prominent position in contemporary research is occupied by the concept of Fundamental Movement Skills (FMS), according to which basic movement actions are considered the foundation for subsequent motor development, the formation of motor competence, and lifelong engagement in physical activity (Gallahue et al., 2012; Stodden et al., 2008; Lubans et al., 2010; Hulteen et al., 2018).

Within the FMS framework, a substantial body of scientific evidence has accumulated regarding the structure of fundamental movement skills, the patterns of their development, assessment methods, and their relationships with physical fitness, physical activity, and health outcomes (Lubans et al., 2010; Logan et al., 2012; Barnett et al., 2016; Khudolii et al., 2025a). The most widely accepted classifications divide fundamental movement skills into locomotor, object-control (manipulative), and stability skills (Gallahue et al., 2012; Barnett et al., 2016; Khudolii et al., 2025b). This approach provides a framework for describing basic forms of motor activity and identifying their role within the process of motor development.

Corresponding Author: Olha Ivashchenko,
e-mail: ivashchenko.olha21@gmail.com

However, contemporary classifications of fundamental movement skills primarily reflect the external structure of movement actions and their place within developmental progression (Gallahue et al., 2012; Barnett et al., 2016; Hulteen et al., 2018). Considerably less attention has been devoted to the mechanisms underlying skill acquisition, variability, and effective application under changing environmental conditions. Consequently, there is a need for a theoretical framework that would allow motor development to be examined as an integrated process and fundamental movement skills to be interpreted as outcomes of a complex functional system.

The theoretical foundations for such an approach can be found in the works of Bernstein devoted to the problems of motor control and movement construction (Bernstein, 1967; Bernshtein, 1966). Further development of the concept of motor function is associated with the studies of Farfel (1975), Shlemin (1973), Liubomirskii (1974), Vilchkovskii (1983, 1993), and other scholars who considered motor function as a complex integrative formation that ensures movement acquisition, motor control, and adaptation to the conditions of activity. At the same time, research on fundamental movement skills and research on motor function have evolved largely independently, limiting the possibilities for a comprehensive explanation of the motor development process (Stodden et al., 2008; Barnett et al., 2016; Hulteen et al., 2018).

An analysis of the scientific literature reveals a contradiction between the high level of development of fundamental movement skill classifications and the insufficient understanding of the mechanisms underlying their formation. Contemporary concepts provide convincing explanations regarding which movement skills emerge through development and learning (Gallahue et al., 2012; Stodden et al., 2008; Barnett et al., 2016; Hulteen et al., 2018), yet they offer less insight into how motor abilities, sensory support systems, motor control mechanisms, and motor experience are integrated during skill formation.

Addressing this problem requires consideration of motor function as the system-forming construct of motor development. Within this framework, fundamental movement skills may be interpreted both as outcomes of motor function development and as mechanisms through which motor experience is accumulated. Such an approach makes it possible to move beyond the description of individual forms of motor activity toward the analysis of the principles governing the formation, development, and functioning of the human motor system.

The purpose of this study was to substantiate motor function as a functional system underlying the development of human motor activity, to determine the place of fundamental movement skills within its structure, and to develop a conceptual model of motor function development.

Materials and Methods

This study has a theoretical and methodological character and was conducted in the form of a conceptual narrative review aimed at examining the relationships among the concepts of motor function, motor abilities, motor experience, and fundamental movement skills within the context of contemporary theories of motor learning and motor development.

The methodological foundation of the study was based on the principles of conceptual analysis and knowledge synthesis proposed by Khudolii, Ivashchenko, and Khudolii (2026a, 2026b). According to this approach, the preparation of a scientific text is viewed as a sequential process of knowledge synthesis that includes orientation within the scientific field, problem conceptualization, evidence analysis, synthesis, and interpretation. Within the framework of the present study, the central meta-problem was to determine the place of fundamental movement skills within the structure of human motor development and to substantiate motor function as the system-forming concept for their interpretation.

The materials for the study consisted of classical and contemporary scholarly works addressing motor function, motor control, motor development, motor learning, and fundamental movement skills. The analysis incorporated the works of M.O. Bernshtein, V.S. Farfel, A.M. Shlemin, L.Ye. Lyubomirskiy, E.S. Vilchkovskyi, and Yu.V. Verkhoshanskyi, as well as contemporary research in the fields of Fundamental Movement Skills (FMS), Motor Competence, Physical Literacy, and Motor Learning.

The following methods were employed:

- *Historical-logical analysis* to examine the evolution of scientific views on motor function and to determine its place within the system of human motor development;
- *Content analysis of scientific sources* to identify key concepts, categories, and theoretical approaches to the interpretation of motor function, motor abilities, and fundamental movement skills;
- *Comparative analysis* to compare contemporary Fundamental Movement Skills concepts with classical perspectives on motor function and motor development;
- *Systems analysis* to investigate the structure of motor function, identify relationships among its components, and substantiate their role in skill formation;
- *Conceptual modeling* to develop a theoretical model of motor function and determine the place of fundamental movement skills within its developmental process.

The study was conducted in four stages.

During the first stage, the evolution of scientific views on motor function and fundamental movement skills was analyzed, and the principal directions in the development of theoretical approaches to their interpretation were identified.

During the second stage, the problem was conceptualized through the examination of motor abilities, motor control, motor experience, and fundamental movement skills, and their place within the system of human motor development was determined.

During the third stage, the relationships among motor abilities, motor control mechanisms, sensory support systems, vestibular stability and sensitivity, motor experience, and fundamental movement skills were examined.

During the fourth stage, the obtained findings were synthesized and interpreted, a conceptual model of motor function was developed, and an approach to interpreting fundamental movement skills as a form of motor function development was substantiated.

The methodological basis for conceptual modeling was the proposition that motor function represents an integrative functional system formed through the solution of motor tasks as a result of the interaction among motor abilities, motor control mechanisms, sensory support systems, and motor experience. Within this framework, fundamental movement skills are considered both an outcome of motor function development and a mechanism for the accumulation of motor experience that supports the further refinement of the human motor system.

Results

Evolution of the Concept of Motor Function

An analysis of the scientific literature indicates that the concept of motor function emerged at the intersection of physiology, psychology, pedagogy, and motor learning theory (Bernshtein, 1966; Farfel, 1975; Shlemin, 1973).

In physiological research, motor function was primarily viewed as the result of the activity of bodily systems responsible for movement execution. V.S. Farfel argued that human motor abilities are supported by several physiological systems, including the motor system itself, the motor control system operating in conjunction with systems for information perception and processing, the energy supply system for muscular activity, the system of functional regulation and integration, and the system of higher mental functions. According to Farfel, each of these systems is morphologically and functionally embedded within the human genetic apparatus and contributes to the characteristics of human motor abilities (Farfel, 1975).

A further development of the concept of motor function is associated with the work of M.O. Bernshtein, who viewed movement as a process of constructing and reconstructing motor actions based on mechanisms of motor control, sensory corrections, and accumulated motor experience. Within this framework, the central object of analysis is not an isolated movement but rather the process of solving a motor task, which requires continuous interaction between motor control mechanisms and sensory regulation of movement activity (Bernshtein, 1966; Bernstein, 1967).

In pedagogical research, motor function gradually came to be understood as a complex integrative construct whose development depends not only on morphofunctional prerequisites but also on learning conditions. Particular importance should be attributed to the works of A.M. Shlemin, L.Ye. Lyubomirskiy, and E.S. Vilchkovskyi, in which the development of motor function is associated with the formation of movement control capabilities, the improvement of coordination mechanisms, and the accumulation of motor experience (Shlemin, 1973; Lyubomirskiy, 1974; Vilchkovskyi, 1983, 1993).

Our own investigations of the structure of motor function in preschool children demonstrated that its state is characterized by indicators of motor control, movement coordination, vestibular stability, flexibility, speed, strength and speed-strength fitness, as well as performance in tasks requiring movement speed and accuracy (Khudolii & Kasian, 2011). At the same time, the pivotal component of the motor function structure was found to be represented by motor control characteristics, particularly the accuracy and speed of movement execution. These characteristics integrate the manifestations of individual motor abilities during the process of solving motor tasks and determine the effectiveness of motor skill acquisition.

Taken together, these findings suggest a gradual evolution in the understanding of motor function—from a physiological characteristic of movement performance to an integrative functional system that combines motor abilities, motor control mechanisms, sensory support, and motor experience in the process of solving motor tasks.

Motor Function as a Functional System

Contemporary research on human motor development predominantly examines individual components of motor activity, including motor abilities, coordination abilities, motor skills, fundamental movement skills, and motor competence (Gallahue et al., 2012; Stodden et al., 2008; Robinson et al., 2015; Barnett et al., 2016; Hulteen et al., 2018). Although this approach is appropriate for analyzing specific aspects of motor activity, it does not fully explain the mechanisms underlying the formation of integrated motor behavior.

V.S. Farfel argued that human motor abilities are supported by several physiological systems, including the motor system itself, the motor control system operating in conjunction with systems responsible for information perception and processing, the energy supply system for muscular work, the system of regulation and integration of functional processes, and the system of higher mental functions. Each of these systems is morphologically and functionally embedded within the human genetic apparatus and determines the characteristics of motor abilities (Farfel, 1975). However, motor abilities cannot be regarded as isolated entities because their practical significance becomes evident only in the process of solving a specific motor task (Farfel, 1975; Bernshtein, 1966).

From the perspective of functional systems theory, the development of the human motor domain should be viewed not as the parallel development of separate abilities or skills but as the formation of an integrated functional system directed toward achieving a useful result (Anokhin, 1975). Within this framework, the development of motor abilities, the refinement of sensory support systems, vestibular stability and sensitivity, and the accumulation of motor experience are not independent goals but components of motor function development (Anokhin, 1975; Bernshtein, 1966).

The interpretation of motor function as an integrative construct is consistent with the fundamental principles of P.K. Anokhin's functional systems theory, according to which every functional system is organized around the achievement of a useful result. Within such a framework, individual morphological and functional components of the organism are not considered in isolation but are integrated into a system whose activity is directed toward producing a specific adaptive effect (Anokhin, 1975).

From this perspective, motor function cannot be reduced either to a set of motor abilities or to a collection of motor skills. Motor abilities, sensory support systems, motor control mechanisms, and motor experience represent components of a functional system that become integrated during the process of solving a motor task (Anokhin, 1975; Bernshtein, 1966; Farfel, 1975). The system-forming factor is not any individual component but the useful result of motor activity itself.

In the context of motor development, this result is expressed as an individual's ability to solve motor tasks effectively under changing conditions of activity. The requirement to achieve this result determines the nature of the interaction among motor abilities, motor control mechanisms, sensory support systems, and accumulated motor experience.

Experimental studies of motor function development have demonstrated that improvements in individual motor abilities do not automatically lead to improvements in motor skill quality (Khudolii & Kasian, 2011). In particular, the maximal development of strength abilities is not necessarily accompanied by improvements in motor control (Khudolii, 2011). Rather, an optimal level of motor ability development creates the prerequisites for the refinement of motor control mechanisms, which in turn support the development of the capacity to differentiate the spatial, temporal, and force characteristics of movement actions. Such differentiation creates the conditions necessary for the formation of variable motor skills adapted to changing activity conditions (Khudolii & Kasian, 2011; Khudolii, 2011).

These findings make it possible to interpret motor function development as a process of progressive integration of its components. At the initial stages, functional prerequisites are established through a certain level of motor ability development, sensory support, and vestibular stability. Subsequently, motor control mechanisms emerge and develop, enabling the accurate reproduction and differentiation of spatial, temporal, and force characteristics of movement actions. The integration of these components ultimately results in the formation of variable motor skills that can be effectively applied under changing activity conditions (Khudolii & Kasian, 2011).

Motor experience occupies a particularly important place within the structure of motor function. Unlike traditional approaches, in which motor experience is viewed as a collection of acquired skills, the functional-systems perspective considers it a mechanism for accumulating and integrating the outcomes of motor activity (Bernshtein, 1966; Bernstein, 1967). Motor experience preserves and refines successful ways of solving motor tasks, thereby providing a foundation for the further development of motor function.

Of particular importance is the principle of feedback, which occupies a central place in functional systems theory. The results of movement execution are accumulated in the form of motor experience and subsequently used to improve motor control mechanisms. Consequently, motor experience functions not only as an outcome of motor function but also as a factor driving its further development (Anokhin, 1975; Bernshtein, 1966; Bernstein, 1967).

Thus, motor function development may be viewed as a continuous process of integration and reorganization of a functional system in response to new activity demands. Within this process, motor abilities create the functional prerequisites for development, motor control mechanisms regulate movement activity, sensory support systems provide the informational basis for movement control, and motor experience serves as a mechanism of feedback and accumulation of activity outcomes.

The result of motor function is not merely the acquisition of individual motor skills but, more importantly, the ability to use them effectively for solving motor tasks under changing conditions of activity. This ability represents the primary criterion of motor function development and reflects the degree of integration of its components into a coherent functional system.

Within the proposed framework, motor function can be operationalized through a system of indicators characterizing motor control and its sensory support. These indicators include the accuracy of reproducing spatial, temporal, and force characteristics of movements, the ability to differentiate these characteristics, measures of vestibular stability and sensitivity, and performance indicators in motor tasks that reflect the level of accumulated motor experience.

Therefore, motor function is a functional system that develops through learning and development as a result of the integration of motor abilities, motor control mechanisms, sensory support systems, and motor experience. The outcome of this system is the effective solution of motor tasks and the formation of variable motor skills adapted to changing activity conditions. Such an understanding of motor function provides the theoretical foundation for interpreting fundamental movement skills as a form of motor function development.

Fundamental Movement Skills: Contemporary Approaches to Classification

The concept of Fundamental Movement Skills (FMS) has emerged as one of the leading frameworks for studying children's motor development. It is based on the assumption that a set of basic movement actions serves as the foundation for the subsequent acquisition of more complex forms of motor activity, physical exercises, and sport-specific skills (Gallahue et al., 2012; Stodden et al., 2008; Hulteen et al., 2018).

In contemporary scientific literature, fundamental movement skills are generally defined as basic movement patterns acquired through learning and development that enable the effective performance of a wide range of motor tasks. Unlike reflexive or naturally occurring forms of movement behavior, fundamental movement skills are acquired in nature and develop under the influence of pedagogical and environmental factors (Gallahue et al., 2012; Logan et al., 2012; Barnett et al., 2016).

The most widely accepted classification divides fundamental movement skills into three major categories (Gallahue et al., 2012; Barnett et al., 2016):

- Locomotor skills;
- Object-control skills;
- Stability skills.

Locomotor skills include walking, running, jumping, hopping, galloping, and other forms of body displacement through space. Object-control skills encompass various forms of interaction with external objects, including throwing, catching, striking, and dribbling. Stability skills involve maintaining balance, body control, turning movements, and other actions associated with controlling body position and posture (Gallahue et al., 2012).

The further development of the FMS concept has expanded understanding of their role in the formation of physical literacy, motor competence, and lifelong engagement in physical activity. Within this perspective, fundamental movement skills are viewed not only as individual movement actions but also as a foundation for subsequent motor development (Stodden et al., 2008; Lubans et al., 2010; Robinson et al., 2015; Barnett et al., 2016; Hulteen et al., 2018).

At the same time, most contemporary classifications of FMS are based primarily on the observable manifestations of motor activity and the manner in which movement actions are performed. While this approach provides an effective framework for organizing and describing diverse forms of movement, it offers a more limited understanding of the mechanisms underlying their formation and their place within the broader structure of motor development.

Consequently, the contemporary concept of fundamental movement skills provides a robust framework for describing and classifying basic forms of motor activity. However, it leaves unresolved the question of how fundamental movement skills are related to the structure of motor function and through which mechanisms they emerge during the processes of development and learning.

Fundamental Movement Skills as a Form of Motor Function Development

Traditional approaches to the interpretation of Fundamental Movement Skills (FMS) are based on the assumption that they constitute the foundation of subsequent motor development. Within this framework, locomotor, object-control, and stability skills are viewed as the basis for the acquisition of more complex movement actions, sport-specific skills, and motor competence (Gallahue et al., 2012; Stodden et al., 2008; Barnett et al., 2016; Robinson et al., 2015).

However, an analysis of motor function theory allows this issue to be considered from a different perspective. If motor function is understood as a functional system formed through the integration of motor abilities, motor control mechanisms, sensory support systems, and motor experience (Anokhin, 1975; Bernshtein, 1966; Farfel, 1975), then fundamental movement skills cannot be regarded as its primary element. Rather, they represent both an outcome of motor function development and a means for its further refinement.

The development of motor function involves the progressive formation of motor control mechanisms. Motor abilities create the functional prerequisites for this process but do not, by themselves, determine the quality of motor performance (Farfel, 1975). Of decisive importance is the capacity for motor control, which is expressed through the differentiation of the spatial, temporal, and force characteristics of movement actions (Bernshtein, 1966; Khudolii & Kasian, 2011).

The formation of fundamental movement skills results from the integration of these mechanisms during the process of solving motor tasks. Within this perspective, a skill is not merely an isolated movement action but a specific mode of realizing motor function under particular activity conditions (Bernshtein, 1966; Bernstein, 1967).

Of particular importance is the fact that an acquired motor skill becomes incorporated into motor experience and subsequently serves as a basis for solving new motor tasks. Consequently, fundamental movement skills are simultaneously outcomes of motor function development and elements of motor experience that support its further refinement (Bernshtein, 1966; Bernstein, 1967).

From the perspective of the functional-systems approach, motor function development represents the formation of a functional system integrating motor abilities, motor control mechanisms, sensory support systems, and motor experience (Anokhin, 1975). Motor abilities create the functional prerequisites for motor function development but do not directly determine the quality of motor skills. What is critical is the integration of motor abilities into the motor control system, enabling the differentiation of spatial, temporal, and force characteristics of movement actions. Such differentiation constitutes the mechanism through which the organism's functional capacities are transformed into acquired motor skills (Khudolii & Kasian, 2011; Khudolii, 2011). The outcome of motor function development is the formation of variable motor skills that are adapted to changing activity conditions and capable of ensuring the efficient solution of motor tasks.

Accordingly, fundamental movement skills should be interpreted not as the initial foundation of motor development but as a form of motor function development. From this perspective, FMS reflect the level of development of motor control mechanisms, the degree of integration of motor abilities into motor activity, and the volume of accumulated motor experience.

Classification of Fundamental Movement Skills as Elements of Motor Function Development

Traditional classifications of fundamental movement skills are based primarily on the observable characteristics of motor activity and the manner in which movement actions are performed. The most widely adopted approach categorizes fundamental movement skills into locomotor skills, object-control skills, and stability skills (Gallahue et al., 2012; Barnett et al., 2016). While this classification provides an effective framework for describing diverse forms of motor activity, it does not explain their place within the process of motor function development.

From the perspective of the functional-systems approach, fundamental movement skills should be considered forms of manifestation and development of motor function. Within this framework, the classification of skills should reflect not only the external structure of movement actions but also the functional mechanisms that underlie their formation and application (Anokhin, 1975; Bernshtein, 1966).

Motor function development is accompanied by the integration of motor abilities, sensory support systems, motor control mechanisms, and motor experience into a unified functional system (Farfel, 1975; Anokhin, 1975). Throughout this developmental process, fundamental movement skills represent specific forms through which motor function is expressed at different levels of organization.

A functional analysis makes it possible to identify five principal directions of motor function development that simultaneously define the functional structure of fundamental movement skills: stabilization, spatial orientation, temporal organization of movement, motor control, and adaptation.

Stabilization

Stabilization ensures the maintenance of body position and balance under both static and dynamic conditions of activity. Fundamental movement skills associated with this direction involve the development of vestibular stability, postural control, and the maintenance of equilibrium during movement execution (Vilchkovskyi, 1983; Khudolii & Kasian, 2011).

Spatial Orientation

Spatial orientation enables the determination of body position and the position of body segments in space, the control of movement direction and trajectory, and orientation relative to surrounding objects. Within motor function development, this direction is expressed through skills related to locomotion, changes in movement direction, and spatial control of movement actions (Bernshtein, 1966).

Temporal Organization of Movement

Temporal organization refers to the ability to coordinate movements in time and includes the regulation of movement tempo, rhythm, and sequence. It creates the prerequisites for the development of more advanced forms of motor control (Bernshtein, 1966; Liubomirskii, 1974).

Motor Control

Motor control occupies a central position within the structure of motor function. At the first level of its development, individuals acquire the ability to accurately reproduce the spatial, temporal, and force parameters of movement actions. At the second level, motor control is manifested in the ability to differentiate these spatial, temporal, and force characteristics according to the requirements of a specific motor task. Such differentiation enables the transition from the reproduction of predefined parameters to their purposeful modification and regulation (Bernshtein, 1966; Khudolii & Kasian, 2011; Khudolii, 2011).

Adaptation

Adaptation represents the highest level of motor function expression and characterizes the ability to apply acquired motor skills under changing activity conditions. At this level, movement parameters are reorganized in accordance with environmental demands, enabling the effective solution of novel motor tasks (Bernstein, 1967; Hulteen et al., 2018).

Thus, fundamental movement skills may be interpreted as forms of motor function development reflecting the progressive complexity of mechanisms related to stabilization, spatial orientation, temporal organization, motor control, and adaptation. Within this framework, fundamental movement skills are viewed not as the initial foundation of motor development but as outcomes of the integration of motor function components and, simultaneously, as mechanisms for its further refinement.

The proposed classification is based not on the observable characteristics of movement actions but on the functional mechanisms that support their execution. The identified

Table 1. Levels of Motor Function Development and Forms of Fundamental Movement Skills Expression

Level of Motor Function Development	Stabilization	Spatial Orientation	Temporal Organization	Motor Control	Adaptation
Sensory Foundation	Maintenance of balance	Body orientation in space	Response to temporal cues	Basic movement control	Orientation to environmental changes
Parametric Control	Stability during movement	Control of movement direction and trajectory	Regulation of tempo and rhythm	Spatial, temporal, and force accuracy of movements	Anticipation of changing task conditions
Fundamental Movement Skills	Stable execution of movement actions	Controlled movement through space	Temporal coordination of movements	Differentiation of spatial, temporal, and force characteristics of movement	Variable execution of movement actions
Combined Motor Programs	Maintenance of stability during complex movement actions	Spatial integration of multiple movement actions	Temporal coordination of movement sequences	Control of complex movement actions	Modification of movement actions according to task conditions
Adaptive Motor Programs	Stability under uncertain conditions	Orientation within dynamic environments	Variable temporal organization of movements	Flexible regulation of movement parameters	Solving novel motor tasks under changing environmental conditions

Note. The rows represent successive levels of motor function development, whereas the columns represent functional directions of development. Progression across levels reflects increasing integration of motor abilities, sensory support systems, motor control mechanisms, and motor experience within a functional system.

directions reflect the progressive development of the motor system—from maintaining stability and orientation in space to the regulation of movement and adaptation to changing activity conditions. Each subsequent direction builds upon and incorporates the preceding one as part of a more complex level of functional organization. Consequently, stabilization, spatial orientation, temporal organization of movement, motor control, and adaptation should be understood not as separate groups of movement actions but as functional levels in the development of motor function (Table 1).

The proposed model reflects the hierarchical nature of motor function development. Each successive level does not replace the preceding one but incorporates it into a more complex functional system. Progression from the sensory foundation to adaptive motor programs is accompanied by increasing complexity of motor control mechanisms and the continuous expansion of motor experience (Anokhin, 1975; Bernshtein, 1966).

Motor control occupies a central position within the structure of motor function. At the level of parametric control, it is manifested in the ability to accurately reproduce the spatial, temporal, and force parameters of movement. At the level of fundamental movement skills, motor control is characterized by the ability to differentiate these characteristics according to the requirements of a specific motor task. Such differentiation of spatial, temporal, and force characteristics of movement provides the foundation for the development of variable motor skills capable of effective application under diverse activity conditions (Bernshtein, 1966; Khudolii & Kasian, 2011; Khudolii, 2011).

Within the proposed model, fundamental movement skills occupy an intermediate position between parametric movement control and combined motor programs. This position allows them to be interpreted not as the initial foundation of motor development but as an outcome of the integration of motor function components and, simultaneously, as a mechanism for accumulating motor experience necessary for the further development of motor function.

Consequently, fundamental movement skills represent a transitional level at which the functional capacities of the motor system are transformed into organized, variable, and task-oriented forms of movement behavior.

Programming as a Mechanism for Motor Function Development

Contemporary research on Fundamental Movement Skills (FMS) has focused primarily on the description of movement structures, age-related characteristics of development, and the assessment of skill proficiency (Gallahue et al., 2012; Barnett et al., 2016; Hulteen et al., 2018). In contrast, the mechanisms underlying the purposeful development of motor function remain insufficiently explored.

From the perspective of the functional-systems approach, motor function development is not a spontaneous process of accumulating motor experience. Rather, motor function develops under the influence of systematically organized pedagogical interventions that facilitate the integration of motor abilities, motor control mechanisms, sensory support systems, and motor experience into a unified functional system (Anokhin, 1975).

Within the processes of learning and development, pedagogical influences should be directed not toward the isolated development of individual motor abilities or the acquisition of separate motor skills but toward creating conditions for their integration within the structure of motor function. Consequently, the effectiveness of motor function development is determined not by the magnitude of individual indicators of strength, speed, or endurance but by the degree to which these capacities are incorporated into motor control mechanisms (Bernshtein, 1966; Farfel, 1975).

Experimental evidence indicates that improvements in individual motor abilities alone do not ensure the development of motor function and do not automatically lead to improvements in motor skill quality. Motor abilities create only the functional prerequisites for motor function development. Their integration into the motor control system is the decisive factor. In particular, research has demonstrated that an optimal level of motor ability development creates more favorable conditions for the refinement of motor control mechanisms than an exclusive orientation toward maximizing individual physical qualities. The development of motor control, in turn, enables the transition from the accurate reproduction of spatial, temporal, and force parameters of movement to their differentiation according to the requirements of a specific motor task. Such differentiation constitutes a necessary condition for the formation of variable motor skills adapted to changing activity conditions. These findings support the view that effective skill formation depends not on the isolated development of motor abilities but on their integration into the motor control system and on the structure of learning tasks employed during instruction (Khudolii & Kasian, 2011; Khudolii, 2011; Khudolii, Iermakov, & Bartik, 2020).

Thus, motor function development is a regulated process governed by the interactions among motor abilities, motor control, and motor experience. Within this framework, programming serves as the mechanism through which motor function can be purposefully developed.

Programming is understood as the scientifically grounded determination of the content, volume, intensity, and sequence of pedagogical interventions aimed at producing anticipated changes in the structure of motor function (Verkhoshanskii, 1988; Ivashchenko, 2020). Its primary purpose is not merely to develop individual components of motor activity but, above all, to create conditions for their integration into a coherent functional system.

This understanding of programming is consistent with contemporary pedagogical concepts of motor learning organization. Within the Spectrum of Teaching Styles framework, the learning process is viewed as a system of pedagogical decisions that determine the nature of teacher–learner interaction and facilitate the achievement of intended learning outcomes (Mosston & Ashworth, 2008). In the context of motor function development, this implies that the selection of movement organization strategies, feedback procedures, levels of learner autonomy, repetition schedules, and sequences of learning tasks becomes an instrument for the purposeful development of motor control mechanisms, the accumulation of motor experience, and the advancement of motor function as a whole.

The outcome of programming is the formation of variable motor skills that are adapted to activity conditions and capable of ensuring the effective solution of motor tasks in changing environments. In this sense, programming functions both as a mechanism for directing motor function development and as a mechanism for the accumulation of motor experience.

Conceptual Model of Motor Function Development

The results of the present analysis support the interpretation of motor function development as a process of forming a functional system directed toward solving motor tasks under changing activity conditions (Anokhin, 1975; Bernshtein, 1966). Within the proposed framework, the central object of development is neither an individual motor ability nor an individual motor skill, but motor function itself as an integrative outcome of the interaction among motor abilities, motor control mechanisms, sensory support systems, and motor experience.

Motor abilities create the functional prerequisites for motor function development (Farfel, 1975). Through the processes of learning and development, these abilities become integrated into the motor control system, thereby supporting the formation of mechanisms responsible for the regulation of motor activity. A central role is played by motor control, which ensures the coordination of the spatial, temporal, and force characteristics of movement actions in accordance with the requirements of a specific motor task (Bernshtein, 1966; Bernstein, 1967).

Two interrelated levels may be distinguished in the development of motor control. The first level is characterized by the ability to accurately reproduce the spatial, temporal, and force parameters of movement actions. The second level is associated with the development of the ability to differentiate these spatial, temporal, and force characteristics according to activity conditions and task requirements. It is this process of differentiation that enables the transition from simple movement reproduction to purposeful movement regulation (Bernshtein, 1966; Khudolii & Kasian, 2011; Khudolii, 2011).

As motor function develops, fundamental movement skills emerge as forms through which motor function is expressed and further developed. The accumulation and refinement of fundamental movement skills contribute to the development of motor experience, which subsequently becomes an important factor in the continued advancement of motor function (Bernstein, 1967).

The ultimate outcome of motor function development is the formation of variable motor skills that are adapted to changing activity conditions and capable of ensuring the efficient solution of motor tasks. Accordingly, the ability to effectively apply acquired motor skills in variable environments represents the primary criterion of motor function development.

Based on the proposed framework, the process of motor function development may be represented as the following sequence:

Motor abilities → integration into the motor control system → movement accuracy → differentiation of spatial, temporal, and force characteristics of movement → fundamental movement skills → motor experience → variable motor skills adapted to activity conditions.

Programming functions as the mechanism through which this developmental process is regulated and directed, ensuring the purposeful formation of motor function at all stages of its development (Shlemin, 1973; Verkhoshanskii, 1988; Khudolii, 2011; Ivashchenko, 2020).

Consequently, motor function development can be understood as a process of integrating motor abilities, sensory support systems, motor control mechanisms, and motor experience into a functional system whose outcome is the formation of both fundamental and variable motor skills adapted to changing activity conditions. Within the proposed conceptual framework, fundamental movement skills should be interpreted not as the initial foundation of motor development but as a form of motor function development and a mechanism for the accumulation of motor experience.

Discussion

The scientific novelty of the proposed approach lies in interpreting fundamental movement skills not as the initial foundation of motor development but as a form of motor function development. Unlike FMS-based frameworks, which focus primarily on the classification and assessment of basic movement skills, Motor Competence approaches, which emphasize the level of motor proficiency, and Physical Literacy frameworks, which focus on lifelong engagement in physical activity, the present approach positions motor function as the system-forming construct that integrates motor abilities, motor control mechanisms, sensory support systems, and motor experience in the process of solving motor tasks.

The proposed concept of motor function does not contradict contemporary approaches to the study of motor development; rather, it extends their explanatory potential. Within the Fundamental Movement Skills framework, the primary object of analysis is the development and assessment of basic movement skills. The concept of Motor Competence focuses on the degree of proficiency in performing motor tasks. Physical Literacy emphasizes readiness for lifelong physical activity, encompassing not only motor competence but also motivational, cognitive, and behavioral dimensions. In contrast, the present framework identifies motor function as the central object of investigation and conceptualizes it as an integrative functional system responsible for the formation and application of motor skills during motor task performance.

The results of the present analysis confirm that the concept of Fundamental Movement Skills represents one of the most influential approaches to understanding children's motor development. Its contribution lies in identifying basic movement actions that support the acquisition of more complex forms of motor activity, physical exercises, and sport-specific skills. Consequently, FMS have been widely regarded as the foundation of motor competence, physical literacy, and lifelong participation in physical activity (Stodden et al., 2008; Lubans et al., 2010; Robinson et al., 2015; Barnett et al., 2016; Hulteen et al., 2018).

At the same time, most contemporary FMS models describe the forms of motor activity rather than the mechanisms through which they emerge. Traditional approaches primarily focus on movement classification, age-related developmental characteristics, and skill assessment (Gallahue et al., 2012; Barnett et al., 2016). In contrast, the question of how motor abilities, motor control mechanisms, sensory support systems, and motor experience are integrated during skill formation remains insufficiently addressed.

The proposed framework interprets fundamental movement skills not as the starting point of motor development but as a form of motor function development. This perspective shifts attention from the description of observable movement patterns to the analysis of the mechanisms underlying their formation. Within this framework, motor function is conceptualized as a functional system (Anokhin, 1975), whereas fundamental movement skills are viewed simultaneously as outcomes of the integration of its components and as mechanisms for accumulating motor experience.

The relationship among motor function, fundamental movement skills, and motor experience is therefore cyclical in nature. Motor function enables the formation of fundamental movement skills; acquired skills are accumulated within the structure of motor experience; and motor experience subsequently becomes a factor contributing to the further development of motor function and the emergence of more complex forms of motor activity.

Particular importance within the proposed framework is assigned to motor control as the central mechanism responsible for integrating the components of motor function. The present findings are consistent with theoretical and experimental evidence indicating that improvements in individual motor abilities do not automatically result in improved motor performance or higher-quality motor skills (Farfel, 1975; Khudolii & Kasian, 2011; Khudolii, Iermakov, & Bartik, 2020). Experimental studies have shown that the effectiveness of motor learning depends not only on functional preparedness but also on the organization of the learning process and its capacity to facilitate the development of motor control mechanisms (Khudolii, Iermakov, & Bartik, 2020). Within the proposed model, motor control is regarded as the system-forming component of motor function that integrates motor abilities, sensory support systems, and motor experience during motor task performance. In this context, the ability to differentiate the spatial, temporal, and force characteristics of movement serves both as an indicator of motor control development and as a prerequisite for the formation of variable motor skills adapted to changing activity conditions (Bernshtein, 1966; Khudolii, 2011).

An important feature of the proposed framework is the possibility of operationalizing motor function as an object of empirical investigation. Within this conceptualization, the state of motor function may be assessed through indicators reflecting motor control and its sensory support. Such indicators include the accuracy of reproducing spatial, temporal, and force characteristics of movement, the ability to differentiate these characteristics, measures of vestibular stability and sensitivity, and performance indicators that reflect the level of accumulated motor experience.

Additional support for the proposed framework is provided by studies examining the influence of theoretical learning models on the effectiveness of skill acquisition. These studies indicate that the theory of functional systems and the theory of movement construction make substantial contributions to the organization of effective learning processes. The effectiveness of learning depends not only on the content of motor tasks but also on the theoretical framework guiding the organization of instruction. Such findings suggest that motor skill acquisition is not simply the result of accumulating motor experience or developing individual motor abilities. Rather, it depends on the purposeful management of the integration of motor abilities, motor control mechanisms, sensory support systems, and motor experience. This interpretation is consistent with the present framework, according to which motor function development is a regulated process achieved through the programming of pedagogical interventions aimed at forming an integrated functional system whose outcome is the development of variable motor skills adapted to changing activity conditions (Khudolii, Iermakov, & Bartik, 2020).

The proposed framework also offers a new interpretation of the role of motor abilities in motor development. Traditional models frequently regard motor abilities as the direct basis for motor skill acquisition. The present findings suggest that motor abilities create the functional prerequisites for motor function development but acquire practical significance only through their integration into motor control systems and motor activity itself (Farfel, 1975; Bernshtein, 1966).

Another important implication concerns the role of programming in the learning process. If fundamental movement skills are interpreted as a form of motor function development, programming should be directed not toward the isolated development of individual abilities or skills but toward creating conditions for the integration of motor function components and the formation of variable motor skills adapted to activity conditions (Shlemin, 1973; Verkhoshanskii, 1988; Khudolii, 2011; Ivashchenko, 2020).

In summary, the concept of motor function expands the possibilities for interpreting fundamental movement skills by placing them within the broader context of motor development as an integrated functional process. Future research should focus on the empirical validation of the proposed model, the refinement of the structure of motor function, and the development of evidence-based technologies for its purposeful formation across different stages of ontogenesis.

Conclusions

The analysis of the evolution of scientific views on motor function demonstrated a transition from understanding motor function as a set of physiological mechanisms supporting movement to interpreting it as an integrative functional system responsible for the effective solution of motor tasks throughout learning and development.

Motor function is formed through the integration of motor abilities, motor control mechanisms, sensory support systems, and motor experience. Motor abilities create the functional prerequisites for motor function development but acquire practical significance only through their integration into motor activity.

Motor control occupies a central position within the structure of motor function. Its development proceeds from the accurate reproduction of spatial, temporal, and force parameters of movement to the differentiation of these characteristics according to task requirements and activity conditions.

Fundamental movement skills should be interpreted as a form of motor function development rather than as the initial foundation of motor development. Within the proposed framework, fundamental movement skills are both outcomes of the integration of motor function components and mechanisms for the accumulation of motor experience.

A conceptual model of motor function development was developed in which motor abilities create the prerequisites for the formation of motor control mechanisms, motor control enables the differentiation of movement characteristics, and this process ultimately leads to the formation of variable motor skills adapted to changing activity conditions.

Programming functions as the mechanism through which motor function is purposefully developed. By ensuring the integration of motor abilities, motor control mechanisms, sensory support systems, and motor experience, programming facilitates the formation of variable motor skills capable of supporting the effective solution of motor tasks in changing environments.

Ethics Approval

Ethical approval was not required for this study because it is a theoretical and conceptual analysis based on published scientific literature and does not involve human participants, animals, or the collection of personal data.

Informed Consent

Not applicable. This study did not involve human participants or the collection of personal data.

Data Availability Statement

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

AI Transparency Statement

The authors used AI-assisted tools to support language editing, conceptual organization, and manuscript preparation. All intellectual content, theoretical interpretations, conclusions, and final manuscript revisions were developed and approved by the authors, who take full responsibility for the content of this publication.

Acknowledgements

Not applicable.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Gallahue, D.L., Ozmun, J.C., & Goodway, J.D. (2012). *Understanding motor development: Infants, children, adolescents, adults* (7th ed.). McGraw-Hill.

Stodden, D.F., Goodway, J.D., Langendorfer, S.J., Roberton, M.A., Rudisill, M.E., Garcia, C., & Garcia, L.E. (2008). A developmental perspective on the role of motor skill competence in physical activity. *Quest, 60*(2), 290-306. https://doi.org/10.1080/00336297.2008.10483582

Lubans, D.R., Morgan, P.J., Cliff, D.P., Barnett, L.M., & Okely, A.D. (2010). Fundamental movement skills in children and adolescents: Review of associated health benefits. *Sports Medicine, 40*(12), 1019-1035. https://doi.org/10.2165/11536850-000000000-00000

Hulteen, R.M., Morgan, P.J., Barnett, L.M., Stodden, D.F., & Lubans, D.R. (2018). Development of foundational movement skills: A conceptual model for physical activity across the lifespan. *Sports Medicine, 48*(7), 1533-1540. https://doi.org/10.1007/s40279-018-0892-6

Logan, S.W., Robinson, L.E., Wilson, A.E., & Lucas, W.A. (2012). Getting the fundamentals of movement: A meta-analysis of the effectiveness of motor skill interventions in children. *Child: Care, Health and Development, 38*(3), 305-315. https://doi.org/10.1111/j.1365-2214.2011.01307.x

Barnett, L.M., Stodden, D.F., Cohen, K.E., Smith, J.J., Lubans, D.R., Lenoir, M., Iivonen, S., Miller, A.D., Laukkanen, A., Dudley, D., Lander, N.J., Brown, H., & Morgan, P.J. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education, 35*(3), 219-225. https://doi.org/10.1123/jtpe.2014-0209

Khudolii, M., Ivashchenko, O., & Khudolii, O. (2025a). Fundamental movement skills research from 2000 to 2025: A bibliometric analysis with pedagogical interpretation. *Journal of Learning Theory and Methodology, 6*(3), 138-145. https://doi.org/10.17309/jltm.2025.6.3.05

Khudolii, M., Ivashchenko, O., & Khudolii, O. (2025b). Fundamental movement skills development in children aged 5-12: Theory, evidence, and pedagogical models. *Physical Education Theory and Methodology, 25*(6), 1456-1467. https://doi.org/10.17309/tmfv.2025.6.16

Bernstein, N.A. (1967). *The co-ordination and regulation of movements.* Oxford: Pergamon Press.

Bernshtein, N.A. (1966). *Ocherki po fiziologii dvizhenii i fiziologii aktivnosti* [Essays on the physiology of movements and physiology of activity]. Moskva: Meditsina.

Farfel, V.S. (1975). *Upravlenie dvizheniiami v sporte* [Movement control in sport]. Moskva: Fizkultura i sport.

Shlemin, A.M. (1973). *Iunyi gimnast* [Young gymnast]. Moskva: Fizkultura i sport.

Liubomirskii, L.E. (1974). *Upravlenie dvizheniiami u detei i podrostkov* [Movement control in children and adolescents]. Moskva: Pedagogika.

Vilchkovskii, E.S. (1983). *Razvitie dvigatelnoi funktsii u detei* [Development of motor function in children]. Kiev: Radianska shkola.

Vilchkovskii, E.S. (1993). *Teoriia i metodika fizicheskogo vospitaniia detei doshkolnogo vozrasta* [Theory and methodology of physical education of preschool children]. Kiev: Radianska shkola.

Khudolii, O., Ivashchenko, O., & Khudolii, M. (2026a). From Knowledge Mapping to Evidence Synthesis: A Critical Path Model for Constructing Coherent Scientific Texts. *Physical Education Theory and Methodology, 26*(3), 411-418. https://doi.org/10.17309/tmfv.2026.3.01

Khudolii, O., Ivashchenko, O., & Khudolii, M. (2026b). The Meta-Problem in Scientific Research in Physical Education and Sport: Architecture and Pathways to Knowledge. *Physical Education Theory and Methodology, 26*(4), 646-654. https://doi.org/10.17309/tmfv.2026.4.01

Khudolii, O.M., & Kasian, A.V. (2011). Zakonomirnosti formuvannia rukhovoi funktsii u ditei doshkilnoho viku [Regularities of motor function formation in preschool children]. *Physical Education Theory and Methodology,* 12, 35-43.

Robinson, L.E., Stodden, D.F., Barnett, L.M., Lopes, V.P., Logan, S.W., Rodrigues, L.P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine, 45*(9), 1273-1284. https://doi.org/10.1007/s40279-015-0351-6

Anokhin, P.K. (1975). *Ocherki po fiziologii funktsionalnykh sistem* [Essays on the physiology of functional systems]. Moskva: Meditsina.

Khudolii, O.M. (2011). *Teoretyko-metodychni osnovy systemy pidhotovky yunykh himnastiv* [Theoretical and methodological foundations of the system of young gymnasts' training] (Doctoral dissertation).

Khudolii, O., Iermakov, S., & Bartik, P. (2020). Didactics: Methodological Basis of Motor Learning in Children and Adolescents. *Journal of Learning Theory and Methodology, 1*(1), 5-13. https://doi.org/10.17309/jltm.2020.1.01

Verkhoshanskii, Iu.V. (1988). *Osnovy spetsialnoi fizicheskoi podgotovki sportsmenov* [Fundamentals of special physical preparation of athletes]. Moskva: Fizkultura i sport.

Ivashchenko, O. (2020). Research Program: Modeling of Motor Abilities Development and Teaching of Schoolchildren. *Physical Education Theory and Methodology, 20*(1), 32–41. https://doi.org/10.17309/tmfv.2020.1.05

Mosston, M., & Ashworth, S. (2008). *Teaching physical education.* Spectrum Institute for Teaching and Learning.

Рухова функція як концептуальна основа інтерпретації фундаментальних рухових навичок

Олег Худолій[1ABCD], Ольга Іващенко[1ABCD], Микола Худолій[1ABCD]

[1]Харківська державна академія фізичної культури

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 11 с., 1 табл., 24 джерел.

Мета. Обґрунтувати рухову функцію як функціональну систему розвитку рухової діяльності людини, визначити місце фундаментальних рухових навичок у її структурі та розробити концептуальну модель розвитку рухової функції.

Матеріал і методи. Дослідження має теоретико-методологічний характер і виконане у форматі концептуального наративного огляду. Методологічною основою стали положення концептуального аналізу та синтезу знань. Матеріалами дослідження були класичні та сучасні наукові праці, присвячені руховій функції, управлінню рухами, руховому розвитку, руховому навчанню та фундаментальним руховим навичкам. Використано методи історико-логічного аналізу, контент-аналізу наукових джерел, порівняльного аналізу, системного аналізу та концептуального моделювання.

Результати. Показано, що рухова функція є інтегративною функціональною системою, яка формується в процесі взаємодії рухових здібностей, механізмів управління рухами, сенсорного забезпечення та рухового досвіду. Обґрунтовано, що центральне місце в структурі рухової функції займає управління рухами, розвиток якого пов'язаний із переходом від точного відтворення до диференціювання просторових, часових і силових характеристик рухових дій. Запропоновано розглядати фундаментальні рухові навички не як вихідну основу рухового розвитку, а як форму розвитку рухової функції та механізм накопичення рухового досвіду. Розроблено концептуальну модель розвитку рухової функції та функціональну класифікацію фундаментальних рухових навичок, що включає стабілізацію, просторову орієнтацію, часову організацію рухів, управління рухами та адаптацію. Обґрунтовано програмування як механізм цілеспрямованого формування рухової функції.

Висновки. Рухова функція є системоутворюючим концептом рухового розвитку, який інтегрує рухові здібності, механізми управління рухами, сенсорне забезпечення та руховий досвід у процесі розв'язання рухових завдань. Фундаментальні рухові навички доцільно розглядати як форму розвитку рухової функції, а програмування — як механізм її цілеспрямованого формування. Запропонований підхід розширює пояснювальні можливості сучасних концепцій Fundamental Movement Skills, Motor Competence та Physical Literacy.

Ключові слова: рухова функція, фундаментальні рухові навички, руховий розвиток, управління рухами, руховий досвід, програмування, моторне навчання.

Information about the Authors:

Khudolii Oleg: khudolii.oleg@gmail.com; https://orcid.org/0000-0002-5605-9939; Department of Olympic and Professional Sports, Kharkiv State Academy of Physical Culture, Klochkivska St, 99, Kharkiv, 61022, Ukraine.

Ivashchenko Olha: ivashchenko.olha21@gmail.com; https://orcid.org/0000-0002-2708-5636; Department of Theory and Methods of Physical Education, Kharkiv State Academy of Physical Culture, Klochkivska St, 99, Kharkiv, 61022, Ukraine.

Khudolii Mykola: khudolii.mykola88@gmail.com; https://orcid.org/0009-0002-1482-6498; Kharkiv State Academy of Physical Culture, Klochkivska St, 99, Kharkiv, 61022, Ukraine.

Cite this article as: Khudolii, O., Ivashchenko, O., & Khudolii, M. (2026). Motor Function as a Conceptual Framework for Interpreting Fundamental Movement Skills. Journal of Learning Theory and Methodology, 7(2), 62-72. https://doi.org/10.17309/jltm.2026.7.2.02

Received: 25.05.2026. Accepted: 15.06.2026. Published: 30.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Conceptual Models in Learning Theory

The "Release Illusion": A Conceptual Model of Motor Skill Formation in Archery

Oleksandr Kalynichenko[1ABCD]

[1]Stepan Gzhytskyi National University of Veterinary Medicine and Biotechnologies Lviv

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.01

Abstract

Background. In archery, the final phase of the shot (release) is traditionally considered a controllable action. However, high-level performance indicates that attempts to consciously control this phase reduce stability and are associated with disruptions such as target panic. This creates a contradiction between the conventional interpretation of release as a controlled action and the actual mechanisms underlying effective performance.

Objectives. To develop a conceptual model of motor skill formation in archery that explains the emergence of release as an automated event, defines the conditions for stable shot execution, and identifies the mechanisms underlying skill disruption.

Materials and Methods. The study has a conceptual and theoretical character and is based on the analysis and synthesis of scientific and methodological literature in sport theory, motor learning, and movement psychology, as well as on the generalization of practical coaching experience in archery. A systems approach to movement analysis and conceptual modeling methods were applied.

Results. The shooting action is interpreted as an integrated functional system in which performance stability depends on the coordination of subsystems rather than the accuracy of individual elements. Two subsystems are identified: the aiming system and the execution system, whose functional separation is essential for stability. Release is not considered a discrete controlled action but emerges as an automated event when the execution system reaches a critical functional state. Skill disruption is explained as a transition from a program-based to a reactive mode of action due to conscious interference in the final phase. Follow-through is defined as a stabilizing phase that ensures the formation of sensorimotor feedback.

Conclusions. Shot stability is ensured by preserving the integrity of the motor program and maintaining the functional separation between aiming and execution systems. Release serves as an indicator of a correctly organized action rather than an object of control. Skill disruption is associated with a shift to a reactive mode of organization. The proposed model generalizes the principles of the Ukrainian school of archery and provides a framework for application in other complex coordination activities.

Keywords: archery, motor learning, motor skill, release, sensorimotor control, target panic.

Introduction

In archery, shot performance is traditionally associated with the accuracy of individual technical elements, among which the final phase—the release—occupies a central position. In coaching practice, release is often treated as an action subject to direct control and expected to be executed at a specific moment (Ertan, 2009; Kalinichenko, 2020). At the same time, empirical observations indicate that attempts to consciously control this phase lead to decreased performance stability and the emergence of specific disruptions, including phenomena such as target panic (Hill et al., 2010; Nieuwenhuys & Oudejans, 2012).

This creates a contradiction between the conventional interpretation of release as a controllable action and the practice of high-level performance, where stability is achieved in the absence of direct intervention in the final phase (Masters, 1992; Beilock & Carr, 2001). Within existing approaches, this contradiction remains insufficiently explained, as analysis is typically focused on individual elements of technique rather than on the mechanisms underlying the organization of motor action as an integrated system (Kalinichenko, 1994, 1995).

Contemporary perspectives on motor activity conceptualize it as a hierarchically organized process in which performance outcomes are determined by the interaction of multiple levels of regulation (Bernstein, 1967; Wolpert et al., 2011). In this context, particular importance is attributed

JLTM
LLC OVS
Corresponding Author: Oleksandr Kalynichenko,
e-mail: kalinarch@gmail.com

to the distinction between orientational and execution components of action, as well as to the role of automated mechanisms in ensuring stable performance (Galperin, 1969; Schmidt & Lee, 2019). However, these theoretical positions have not been sufficiently specified for the context of archery and do not adequately explain the mechanism of release and its disruption (Kalinichenko, 2008, 2009).

This necessitates a reconsideration of the structure of motor action in archery from a systems perspective, allowing the shot to be interpreted as the result of interaction between functionally distinct subsystems. Such an approach provides a framework for explaining both performance stability and the mechanisms underlying its disruption within a unified model (Anokhin, 1974; Shadmehr & Krakauer, 2008). At the same time, a research tradition has been established in archery that consistently develops a systems-based interpretation of motor action and is represented in international specialized publications (Kalinichenko, 1994, 1995, 2008, 2011, 2020; Kalynichenko, 2019, 2020). The practical relevance of this line of work is confirmed by its application in training athletes across different levels of expertise, from beginners to elite performers.

Within existing approaches, release is predominantly interpreted as an action subject to control or optimization within a specific technical element. The present approach revises this interpretation by treating release not as an action, but as an automated event that emerges when the execution system reaches a critical functional state. This perspective allows for a redefinition of the mechanisms underlying motor skill stability and explains skill disruption not as an execution error, but as a shift in the principle of action organization.

The scientific novelty of this study lies in the development of a conceptual model that explains the formation and disruption of motor skills through the principle of functional separation between the aiming and execution systems, and through the interpretation of release as an automated event. Within this framework, performance stability is understood as the preservation of the motor program, while disruption is interpreted as a transition to a reactive mode of organization.

The aim of the study is to develop a conceptual model of motor skill formation in archery that explains the emergence of release as an automated event, defines the conditions for stable shot execution, and identifies the mechanisms underlying disruption of the motor program.

Materials and Methods

The study has a conceptual and theoretical character and is aimed at developing a generalized model of motor action organization in archery. The work is based on the analysis and synthesis of scientific and methodological publications in sport theory and methodology, movement psychology, and motor learning, as well as on the generalization of practical coaching experience (Schmidt & Lee, 2019; Wolpert et al., 2011). The source base includes scientific articles, instructional and methodological materials, coaching guidelines, and the results of long-term practice in training athletes of different qualification levels (Kalinichenko, 1994, 1995, 1999, 2008, 2011, 2020; Kalynichenko, 2019, 2020).

The selection of materials was carried out based on their substantive relevance to the problem of motor skill formation in complex coordination activities. The analysis included sources that allow motor action to be interpreted as an integrated functional system (Anokhin, 1974; Bernstein, 1967; Kalinichenko, 2008, 2011), as well as studies addressing the mechanisms of movement stabilization and disruption (Masters, 1992; Beilock & Carr, 2001; Kalinichenko, 2009, 2013). Sources limited to the description of individual technical elements without analysis of their functional role were not considered decisive for model construction.

The search and selection of sources were conducted using open scientific databases and libraries with keywords related to motor learning, movement coordination, action control, and archery technique. Additional sources reflecting the practical experience of the Ukrainian school of archery, represented in international specialized publications, were also included (Kalinichenko, 1994, 1995, 1999, 2020; Kalynichenko, 2019, 2020).

The methodological foundation of the study is the systems approach to movement analysis, according to which action is considered as the result of interaction among interrelated subsystems (Anokhin, 1974; Kalinichenko, 2011). The theoretical framework is based on the concepts of hierarchical organization of motor actions and stage-based skill formation (Bernstein, 1967; Galperin, 1969), which make it possible to analyze the relationship between orientational and execution components of action at different stages of its realization (Kalinichenko, 2009, 2012).

The main methods of the study were analysis and synthesis of scientific literature, comparative analysis of approaches to the interpretation of shooting technique, and conceptual modeling. Modeling was applied to construct a generalized scheme of motor action organization, allowing the mechanisms of stable performance and disruption to be explained (Shadmehr & Krakauer, 2008; Kalinichenko, 2011, 2013).

The limitations of the study are determined by its conceptual nature and its focus on generalizing functional principles of motor action organization. The study does not aim to provide quantitative analysis or systematic comparison of the effectiveness of different training methods. The primary focus is on the development of a theoretical model and its interpretation within a unified conceptual framework.

The generalization of the obtained positions was carried out through their integration into a conceptual model that reflects the interaction between the aiming and execution systems, the conditions for the emergence of release as an automated event, and the mechanisms of motor skill stabilization.

Results

The results are presented in the form of a conceptual model describing the organization of motor action in archery through the interaction of functional subsystems and the mechanisms of their coordination.

1. Conceptual Model of Motor Skill Formation in Archery

Motor skill formation in archery can be understood as the process of organizing an integrated functional system "archer–bow," in which performance outcome is determined not by the accuracy of individual elements, but by the coordination of their interaction and the preservation of the integrity of the motor program (Kalinichenko, 2008, 2011).

Within this system, the shot is not a sequence of discrete actions executed under direct control, but the result of an unfolding organized process in which individual components perform different functions and are governed by different regulatory principles (Kalinichenko, 1994, 1995). Disruption of this organization leads not to local errors, but to a change in the mechanism of action realization.

The model distinguishes two relatively autonomous subsystems—the aiming system and the execution system—whose interaction determines both shot stability and the conditions under which disruption occurs (Kalinichenko, 2009, 2011). These subsystems are not interchangeable and perform different functions: one establishes the conditions for action, while the other ensures its realization.

Thus, motor skill formation in archery is not based on refining individual movements, but on establishing a mode of system organization in which the motor program can be realized without distortion, and its final phase emerges as a natural consequence of this organization (Kalinichenko, 2011).

2. Aiming System

The aiming system is responsible for forming the spatial and dynamic conditions under which the motor program of the shot can be realized (Kalinichenko, 2009). Its function is not to execute the action, but to bring the "archer–bow" system into a state that allows stable completion of the movement without additional triggering influence.

Aiming includes visual alignment, stabilization of posture, control of oscillations, and reproducibility of reference points. However, the critical factor is not the list of components, but their coordinated organization. Functionally, the aiming system establishes an acceptable range of deviations within which further action remains stable under internal and external perturbations (Kalinichenko, 2009, 2010).

Importantly, the aiming system should not perform a triggering function. The use of visual signals as a basis for initiating the shot alters its role: from a system that establishes conditions for action to a source of a trigger, leading to interference in the execution process and disruption of its temporal structure (Kalinichenko, 2013).

Thus, the effectiveness of the aiming system is determined not by instantaneous precision of target alignment, but by its ability to create a stable system state within which the motor program can be executed without distortion and without transition to a reactive mode of regulation.

3. Execution System

The execution system ensures realization of the shot as a continuous process in which the motor program unfolds without transition to a discrete triggering act (Kalinichenko, 2011). Its function is not to initiate the action, but to maintain a system state in which movement completion occurs as a natural consequence of prior organization.

A key element of the execution system is the "expansion" phase, which performs not only a biomechanical but also a regulatory function. It ensures continuity of the action, prevents its interruption at a critical point, and eliminates the need for conscious initiation of the final phase (Kalinichenko, 2013). In this state, muscular activity is organized in such a way that the system gradually approaches a threshold condition at which further удержание becomes unstable.

Within this logic, release is not initiated by a discrete command, but emerges as a consequence of the execution system reaching a critical functional state (Kalinichenko, 2020). Any attempt to deliberately influence the timing of movement completion disrupts process continuity, alters the distribution of muscular effort, and shifts the system from a program-based to a reactive mode of operation.

Thus, the effectiveness of the execution system is determined not by the ability to control the final phase, but by the ability to maintain continuity and integrity of the motor program until its completion, thereby creating conditions for the emergence of release as an automated event.

4. Principle of Functional Separation

From the functional characteristics of the aiming and execution systems follows the principle of their separation: the formation of action conditions and the realization of action are governed by different regulatory mechanisms and cannot be effectively performed within a single conscious act (Kalinichenko, 2011).

The aiming system operates in an orientational mode and ensures stabilization of spatial parameters, whereas the execution system functions in a mode of continuous unfolding of the motor program. Attempts to combine these modes in the final phase of action lead to a regulatory conflict: visual control characteristic of aiming interferes with execution, which requires continuity and absence of a triggering act (Kalinichenko, 2009, 2013).

As a result, the system shifts from a program-organized mode to a reactive one, where the timing of the shot is determined not by the internal logic of the action, but by an external signal. This is accompanied by disruption of temporal structure, altered distribution of muscular effort, and loss of movement integrity.

Thus, functional separation of the aiming and execution systems is a necessary condition for motor skill stability.

5. Release as an Automated Event

Within the proposed model, release is interpreted not as a response to situational afferentation and not as a purposeful action, but as an automated event emerging during the realization of a formed motor program when the execution system reaches a critical state (Kalinichenko, 2013, 2020).

The automated nature of release does not imply uncontrolled triggering of a dominant motor stereotype. Rather, it reflects the system's ability to maintain integrity of the selected action until its completion. In this case, the final phase is not initiated by a separate signal but unfolds as a natural consequence of prior movement organization (Kalinichenko, 2011, 2013).

In terms of stage-based action formation, the orientational basis ensures the creation of conditions for action during preparation and aiming stages. However, after reaching an automated level, it should not directly interfere in the final phase. Reintroduction of conscious control at the moment of release leads to disruption of temporal and coordination structure or to a regression toward less adequate action patterns (Kalinichenko, 2009, 2013).

Thus, release functions not as an object of control, but as an indicator that the motor program is preserved and realized without distortion (Kalinichenko, 2020).

6. Mechanism of Disruption

Disruption of motor skill arises not as a random execution error, but as a закономірний consequence of a change in the principle of action organization (Kalinichenko, 2009, 2013).

Under normal conditions, shot execution occurs within an automated motor program with preserved functional separation between aiming and execution systems (Kalinichenko, 2011). A critical factor in disruption is the return of conscious control to the final phase, where it does not perform an organizing function but alters the mechanism of movement realization.

In this case, an artificial triggering connection is formed, in which visual alignment with the target begins to serve as a trigger for shot initiation (Kalinichenko, 2013). Transition to this reactive mode leads to disruption of temporal structure, disorganization of muscular coordination, and loss of movement continuity.

The consolidation of such patterns has a cumulative character: repeated execution under reactive conditions strengthens maladaptive coordination structures and forms stable patterns of disrupted performance. In practice, this manifests as phenomena such as target panic, which should be understood not as isolated psychological reactions, but as systemic consequences of deformation of the motor program (Kalinichenko, 2013).

7. Follow-through

Follow-through is not an external technical element or a formal completion of the shot, but a functionally necessary phase that ensures preservation of the integrity of the motor action after string release. Its absence or reduction leads to premature cessation of muscular activity, altering the structure of the final phase and retrospectively affecting shot quality (Kalinichenko, 2011, 2013).

Functionally, follow-through serves as a stabilizing continuation of the action, allowing the system to maintain the organization established in previous phases. This prevents abrupt transition from active state to relaxation, which is critical for preserving temporal and coordination structure (Kalinichenko, 2011).

At the same time, follow-through performs a sensory function. It is during this phase that proprioceptive and kinesthetic information about the executed action is integrated, forming the basis for performance evaluation. Unlike external results, follow-through provides access to internal parameters that determine reproducibility and stability (Kalinichenko, 2011, 2012).

Thus, follow-through acts as a mechanism of feedback formation, supporting refinement of the motor program rather than correction of isolated errors.

8. Generalization

The results allow motor skill formation in archery to be interpreted as a process of organizing an integrated functional system in which performance accuracy is an emergent effect of coordinated interaction between aiming and execution subsystems and preservation of motor program integrity (Kalinichenko, 2008, 2011).

Performance effectiveness is determined not by the intensity of conscious control in the final phase, but by its functional appropriateness: the orientational basis organizes conditions at preparatory stages, while stable execution requires preservation of automated organization without additional intervention. Stability depends on the system's ability to maintain the selected motor program and realize it without distortion.

The proposed model reinterprets both the mechanism of skill formation and the causes of its disruption, and defines follow-through as a key element of feedback ensuring system stabilization and refinement of the motor program.

Conceptually, the results generalize empirically validated principles of the Ukrainian school of archery and interpret them within the framework of hierarchical organization of movement and stage-based skill formation, creating a basis for application to other complex coordination activities (Table 1).

Table 1. Conceptual Model of Motor Skill Organization in Archery

Component	Functional Role	Key Mechanism	Stability Condition	Disruption Mechanism
Aiming system	Formation of spatial and dynamic conditions	Stabilization of system state	Maintains acceptable range of deviations	Becomes trigger for action (visual control → trigger)
Execution system	Realization of motor program	Continuous unfolding of movement	Preserves continuity of action	Interrupted by conscious control
Functional separation	Distribution of regulatory functions	Separation of orientation and execution	No interference between subsystems	Conflict between regulatory levels
Release	Automated event	Emergence at critical system state	Occurs without conscious initiation	Forced initiation disrupts timing
Disruption mechanism	Change in action organization	Transition to reactive mode	Program-based regulation	Trigger-based execution (reactive control)
Follow-through	Stabilization and feedback formation	Continuation of system organization	Preserves motor program integrity	Premature termination of movement

Discussion

The proposed model differs from existing approaches to motor skill analysis by shifting the focus from control of individual technical elements to the organization of interaction between functional subsystems. Within traditional interpretations, performance stability is associated with execution accuracy and control of critical phases, whereas in the present model it is determined by preservation of motor program integrity and functional separation between the aiming and execution systems.

The findings support the interpretation of motor action in archery as an integrated system in which the key determinant is not a specific technical element, but the mode of organization of interactions among its components. This perspective is consistent with theoretical frameworks of hierarchical motor control, where performance stability emerges from coordination across regulatory levels rather than from local control of isolated elements (Bernstein, 1967; Wolpert et al., 2011).

In this context, the principle of functional separation between aiming and execution systems can be understood as a specific operationalization of the distribution of functions within the structure of action. The aiming system performs an orientational role, establishing the conditions for action, whereas the execution system ensures realization of the motor program. This distinction aligns with the concept of the orientational basis of action (Galperin, 1969), which defines conditions for execution but, after automation, should not directly intervene in the final phase of movement. Violation of this separation leads to conflict between regulatory levels and to a transition from program-organized to reactive action (Masters, 1992; Beilock & Carr, 2001).

The interpretation of release as an automated event provides a more precise account of the final phase of movement. Unlike approaches that treat release as a discrete action subject to direct control, the present model conceptualizes it as the outcome of the execution system reaching a critical functional state. This is consistent with the theory of functional systems (Anokhin, 1974), where results emerge as integrated effects of component interaction rather than as responses to discrete triggering signals (Shadmehr & Krakauer, 2008).

Reconsidering skill disruption as a transition to a reactive mode of organization allows a reinterpretation of phenomena such as target panic. In this framework, such phenomena are not isolated psychological responses but reflect a change in the mechanism of regulation, in which visual information assumes the role of a trigger. This interpretation is consistent with research on performance breakdown in precision sports, demonstrating that excessive conscious control and attentional focus on execution destabilize action (Hill et al., 2010; Oudejans et al., 2011; Nieuwenhuys & Oudejans, 2012), as well as with findings in archery highlighting the role of coordination and muscular control in the final phase of the shot (Ertan, 2009).

A critical element of the model is the role of follow-through, which is interpreted as a phase ensuring system stabilization and feedback formation. This is consistent with concepts of sensory regulation and movement representation, according to which preservation of action integrity enables acquisition of information necessary for further refinement (Bernstein, 1967; Schmidt & Lee, 2019).

Overall, the proposed model does not contradict existing theoretical frameworks but specifies them within the context of archery. It clarifies the relationship between orientational and execution components of action, the role of conscious control across different stages, and the mechanisms underlying the transition from stable to disrupted performance.

The model has defined boundaries of applicability. It most adequately describes motor skill formation and execution in complex coordination tasks at stages of stabilization and high-level performance. Its applicability is limited at early stages of learning, where formation of the orientational basis is dominant, and in conditions of high variability requiring continuous restructuring of the motor program.

The proposed framework extends beyond archery and is relevant to other complex coordination activities in which performance depends on precise execution under time constraints and high sensitivity to interference in the final phase. This provides a basis for further research aimed at refining mechanisms of motor control and optimizing learning processes.

Conclusions

Motor skill formation in archery should be understood as the organization of an integrated functional system in which performance outcome is determined by coordinated interaction between subsystems and preservation of the motor program, rather than by the accuracy of individual technical elements.

Shot stability is ensured by functional separation between the aiming and execution systems: the former establishes the conditions for action, while the latter ensures its realization. Violation of this separation leads to regulatory conflict and a transition to a reactive mode of execution.

Release is an automated event that emerges during the realization of the motor program when the execution system reaches a critical state. It cannot be considered a discrete action subject to direct conscious control.

The mechanism of motor skill disruption is associated with the return of voluntary control to the final phase of action and the formation of a reactive principle of organization, resulting in performance instability and persistent disturbances such as target panic.

Follow-through is a functionally necessary phase that ensures system stabilization and feedback formation, which is essential for further refinement of the motor skill.

The proposed model generalizes empirically validated principles of the Ukrainian school of archery and interprets them within contemporary frameworks of motor action organization, providing a basis for application in other complex coordination activities.

Practical implications of the model indicate that training should be directed not toward controlling the moment of release, but toward establishing a stable state of the execution system in which movement completion occurs automatically. Training should clearly separate the functions of aiming and execution, avoiding the use of visual signals as triggers for action initiation. Preservation of follow-through should be considered a mandatory component of training, ensuring the development of adequate sensorimotor feedback.

The presented framework can be used to optimize training processes, particularly in the development of methods aimed at preventing reactive triggering patterns and promoting stable motor skill acquisition.

Ethics Approval

The study is conceptual and theoretical in nature and did not involve human participants or animals; therefore, ethics approval was not required.

Informed Consent

The study did not involve human participants or personal data; therefore, informed consent was not applicable.

Data Availability Statement

The data supporting the findings of this study are available in publicly accessible scientific and methodological sources cited in the reference list. No additional datasets were generated or analyzed.

AI Transparency Statement

Artificial intelligence tools were used in a supportive manner during manuscript preparation for language editing and structuring of the text. AI was not used to generate scientific ideas, interpret findings, or formulate conclusions. The author takes full responsibility for the content of the work.

Conflict of Interest

The author declares no conflict of interest.

References

Ertan, H. (2009). Muscular activation patterns of the bow arm in recurve archery. *Journal of Science and Medicine in Sport, 12*(3), 357–360. https://doi.org/10.1016/j.jsams.2008.01.003

Kalinichenko, A. (2020). Building a bulletproof release. *BOW International,* 143, 28–33.

Hill, D. M., Hanton, S., Matthews, N., & Fleming, S. (2010). Choking in sport: A review. *International Review of Sport and Exercise Psychology, 3*(1), 24–39. https://doi.org/10.1080/17509840903301199

Nieuwenhuys, A., & Oudejans, R. R. D. (2012). Anxiety and perceptual-motor performance: Toward an integrated model of concepts, mechanisms, and processes. *Psychological Research,* 76, 747–759. https://doi.org/10.1007/s00426-011-0384-x

Masters, R. S. W. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology, 83*(3), 343–358. https://doi.org/10.1111/j.2044-8295.1992.tb02446.x

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*(4), 701–725. https://doi.org/10.1037/0096-3445.130.4.701

Kalinichenko, A. (1994). Theoretical basics of archery. The Glade. *International Magazine for Archers, Summer,* 16–20.

Kalinichenko, A. (1994). Theoretical basics of archery. The Glade. *International Magazine for Archers, Autumn,* 63–64.

Kalinichenko, A. (1995). Problems of aiming in archery. The Glade. *International Magazine for Archers, Spring,* 4–10.

Kalinichenko, A. (1995). Aiming. Archery, 1, 53–56.

Bernstein, N. A. (1967). *The co-ordination and regulation of movements.* Pergamon Press.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*(12), 739–751. https://doi.org/10.1038/nrn3112

Galperin, P. Y. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology* (pp. 249–273). Basic Books.

Schmidt, R. A., & Lee, T. D. (2019). *Motor learning and performance: From principles to application* (6th ed.). Human Kinetics.

Kalinichenko, A. (2008). Vplyv liudskoho faktora na funktsionuvannia systemy "luchnyk–luk". *Teorìâ ta Metodika Fìzičnogo Vihovannâ,* 6, 12–17.

Kalinichenko, A. (2009). Kilkisna otsinka skladovykh navedennia zbroi v strilbi z luka. *Teorìâ ta Metodika Fìzičnogo Vihovannâ,* 6, 24–31.

Kalinichenko, A. (2011). Osoblyvosti modeliuvannia navchalno-trenuvalnoi roboty u striletskykh vydakh sportu. *Teorìâ ta Metodika Fìzičnogo Vihovannâ,* 4, 39–45.

Anokhin, P. K. (1974). *Biology and neurophysiology of the conditioned reflex and its role in adaptive behavior.* Pergamon Press.

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research, 185*(3), 359–381. https://doi.org/10.1007/s00221-008-1280-5

Kalinichenko, O. M. (1995). *Formuvannia struktury rukhovykh dii striltsiv z luka za dopomohoiu vykorystannia tekhnichnykh zasobiv navchannia* [Abstract of PhD dissertation]. Kyiv.

Kalinichenko, A. (1999). Initial instruction in archery. Part I. The Glade. *International Magazine for Archers,* 82, 58–62.

Kalinichenko, A. (1999). Initial instruction in archery. Part II. The Glade. *International Magazine for Archers,* 83, 46–50.

Kalinichenko, A. (1999). Initial instruction in archery. Part III. The Glade. *International Magazine for Archers,* 84, 26–30.

Kalinichenko, A. (1999). Method of the initial teaching in archery. Part I. *Archery,* 1, 54–56.

Kalinichenko, A. (1999). Method of the initial teaching in archery. Part II. *Archery,* 3, 46–49.

Kalinichenko, A. (1999). Method of the initial teaching in archery. Part III. *Archery,* 5, 48–49.

Kalinichenko, A. (1999). Method of the initial teaching in archery. Part IV. *Archery,* 7, 48–49.

Kalinichenko, A. (1999). Method of the initial teaching in archery. Part V. *Archery,* 9, 46–47.

Kalinichenko, O. M., & Lopatiev, A. O. (2009). Osoblyvosti tsilovoi tochnosti pry vykonanni postrilu zi striletskykh vydiv zbroi. *Teorìâ ta Metodika Fìzičnogo Vihovannâ,* 8, 45–49.

Kalinichenko, O. M., & Lopatiev, A. O. (2010). Psykhofiziolohichni osoblyvosti tsilovoi tochnosti pid chas vykonannia postrilu z suchasnykh striletskykh vydiv zbroi. *Teorìâ ta Metodika Fìzičnogo Vihovannâ,* 2, 16–18, 35–42.

Kalinichenko, O. M., & Lopatiev, A. O. (2012). Obgruntuvannia psykhofiziolohichnykh mekhanizmiv zastosuvannia metodychnykh pryiomiv vdoskonalennia rukhovykh navychok striltsiv. *Teoriâ ta Metodika Fìzičnogo Vihovannâ,* 1, 37–44.

Kalinichenko, O. M. (2013). Zastosuvannia mekhanizmiv keruvannia finalnymy diiamy typu "rukhy bez mety" yak metodychnyi pryiom vdoskonalennia rukhovykh navychok striltsiv. *Teoriâ ta Metodika Fìzičnogo Vihovannâ,* 2, 34–42.

Kalynichenko, A. (2019). A retrospective of the technique of archers. Part I. *Archery,* 7, 30–31.

Kalynichenko, A. (2019). A retrospective of the technique of archers. Part II. *Archery,* 9, 50–53.

Kalynichenko, A. (2019). The evolution of technological methods. Part I. *Archery,* 11, 50–54.

Kalynichenko, A. (2020). The evolution of technological methods. Part II. *Archery,* 1, 49–53.

Oudejans, R. R. D., Kuijpers, W., Kooijman, C. C., & Bakker, F. C. (2011). Thoughts and attention of athletes under pressure: Skill-focus or performance worries? *Anxiety, Stress, & Coping, 24*(1), 59–73. https://doi.org/10.1080/10615806.2010.481331

Ілюзія «релізу»: концептуальна модель формування рухової навички у стрільбі з лука

Олександр Калиніченко[1ABCDE]

[1] Львівський національний університет ветеринарної медицини та біотехнологій імені Степана Ґжицького

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 7 с., 1 табл., 37 джерел.

Актуальність. У стрільбі з лука завершальна фаза пострілу (реліз) традиційно розглядається як дія, що підлягає безпосередньому контролю. Водночас практика висококваліфікованого виконання свідчить, що спроби свідомого контролю цієї фази знижують стабільність результату та пов'язані з виникненням порушень типу «паніка мішені». Це створює протиріччя між традиційним уявленням про реліз як керовану дію та фактичними механізмами ефективного виконання.

Мета. Розробити концептуальну модель формування рухової навички у стрільбі з лука, яка пояснює механізм виникнення релізу як автоматизованої події, визначає умови стабільної реалізації пострілу та розкриває причини порушення рухової програми.

Матеріали і методи. Дослідження має наративний характер і ґрунтується на аналізі та узагальненні наукових і методичних публікацій з теорії спорту, моторного навчання та психології рухової діяльності, а також на узагальненні практичного досвіду підготовки спортсменів у стрільбі з лука. Використано системний підхід до аналізу рухової дії та метод концептуального моделювання.

Результати. Рухову дію у стрільбі з лука інтерпретовано як цілісну функціональну систему, у якій вирішальним є не точність окремих елементів, а узгодженість взаємодії підсистем і збереження рухової програми. Виділено систему наведення та систему виконання, функціональне розділення яких є умовою стабільності дії. Показано, що реліз не є окремою керованою дією, а виникає як автоматизована подія при досягненні системою виконання критичного стану. Обґрунтовано механізм порушення рухової навички як перехід від програмно-організованої дії до реактивної внаслідок втручання свідомого контролю у фінальну фазу. Визначено роль післядії як фази стабілізації системи та формування зворотного зв'язку.

Висновки. Стабільність пострілу забезпечується збереженням цілісності рухової програми та функціональним розділенням систем наведення і виконання. Реліз виступає індикатором правильно організованої дії, а не об'єктом управління. Порушення рухової навички пов'язане з переходом до реактивного принципу організації дії. Запропонована модель узагальнює положення української школи лучного спорту і створює підґрунтя для їх застосування у інших видах складнокоординаційної діяльності.

Ключові слова: стрільба з лука, моторне навчання, рухова навичка, реліз, сенсомоторний контроль, паніка мішені.

Information about the Authors:

Kalynichenko Oleksandr: kalinarch@gmail.com; https://orcid.org/0000-0001-9556-3727; Stepan Gzhytskyi National University of Veterinary Medicine and Biotechnologies Lviv, Department of Physical Education, Sport and Health, Pekarska St, 50, Lviv, Lviv Oblast, 79000, Ukraine.

Cite this article as: Kalynichenko, O (2026). The "Release Illusion": A Conceptual Model of Motor Skill Formation in Archery. *Journal of Learning Theory and Methodology, 7*(2), 73-79. https://doi.org/10.17309/jltm.2026.7.2.01

Received: 01.05.2026. Accepted: 17.05.2026. Published: 30.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання
ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Methodological Review and Comparative Analysis

Methods for Estimating Sample Size in Structural Equation Modeling

Fernando Almeida[1ABCD]

[1]University of Porto & INESC TEC

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.03

Abstract

Background. Sample size estimation in Structural Equation Modeling (SEM) remains a debated methodological issue. Although minimum sample size rules are widely used, they often lack integration of theoretical, statistical, and practical considerations, particularly in complex research designs.

Objectives. This study aims to provide a synthesis and comparative analysis of different approaches to sample size estimation in SEM, integrating theoretical foundations, statistical criteria, methodological limitations, and applied research contexts.

Materials and Methods. A multi-method approach was employed, consisting of three components: a narrative review of sample size estimation methods, consideration of simulated scenarios reflecting realistic research conditions, and a comparative evaluation of method performance across these scenarios.

Results. Findings show that rule-of-thumb approaches are useful for preliminary estimates but may misrepresent required sample sizes in complex models. Power analysis provides more precise and rigorous estimates, particularly for confirmatory research. Monte Carlo simulations offer robust guidance for complex, longitudinal, and multigroup models. Estimator-based recommendations highlight the importance of aligning sample size with data characteristics and estimation techniques. Model complexity emerged as a key determinant, with advanced SEM models often requiring 400–700+ participants.

Conclusions. A systematic and context-sensitive approach to sample size estimation is essential for ensuring robustness and methodological rigor in SEM studies across disciplines such as management and education.

Keywords: research methodology, structural equation modeling, sample size estimation, statistical power, model complexity.

Introduction

Structural Equation Modeling (SEM) is an advanced statistical approach that has gained enormous relevance in various fields such as engineering, management, and social sciences, mainly due to its ability to deal with complex relationships between observable and latent variables. Nusair and Hua (2010) and Xiong et al. (2015) confirm this view by pointing out that its importance stems from the fact that it allows researchers to test theories robustly, integrating multiple constructs that cannot be measured directly (e.g., attitudes, perceptions, trust, intentions, or skills). Thus, SEM offers a framework that combines confirmatory factor analysis and multiple regression models. This makes it possible to simultaneously assess the quality of measures and the structural relationships between constructs.

The applicability of SEM is particularly evident when researchers deal with multidimensional theoretical models, in which simple linear regression is insufficient. For example, in studies in the field of innovation, entrepreneurship, and organizational management, phenomena often involve interactions between psychological, behavioral, and contextual factors (Allammari et al., 2025; Dulaimi et al., 2005; Ghardashi et al., 2022). Its application allows for testing complex hypotheses, such as direct, indirect, and mediating effects, as well as moderating effects, providing a more holistic view of the processes studied. Furthermore, it makes it possible to compare alternative models and assess which one best fits the data, contributing to the validation of emerging theories. Another relevant aspect is the ability to work with latent variables, which reduce measurement error by aggregating multiple indicators to represent a construct (Kline, 2010). This feature provides greater accuracy to the results and improves the internal validity of the models. SEM also makes it possible to assess the quality of the scales used through metrics such as composite reliability, convergent and discriminant validity, ensuring that the model measures what

JLTM
LLC OVS

Corresponding Author: Fernando Almeida,
e-mail: almd@fe.up.pt

it is intended to measure (Cheung et al., 2024; Newsom, 2023; Panzeri et al., 2024).

Determining the minimum sample size for applying SEM is one of the most critical aspects of ensuring the validity and robustness of the results. The importance of this definition stems from the fact that SEM is a complex statistical technique that involves multiple parameters to be estimated, including factor loadings, covariances, structural paths, and measurement errors (Kline, 2010). The greater the number of parameters, the greater the need for sufficient data to ensure that the estimates are stable and that the model can converge without problems. An inadequate sample size can result in biased estimates, making it impossible to identify the model or leading to weak fit indices that do not adequately reflect the quality of the theory being tested. Furthermore, as Kim (2005) points out, sample size directly influences the statistical power of the model. Small samples may fail to detect significant effects even when they exist, which undermines the ability to validate theoretical relationships. However, determining this minimum size is a complex and challenging task. One of the difficulties lies in the absence of a universal consensus. The diversity of estimators available in SEM adds another layer of complexity, as different methods require different sample sizes. The presence of non-normal variables, complex models, or structures with many indicators also increases the need for larger samples. Another challenge lies in the dependence on the specific context. Two models with the same number of parameters may require different sample sizes depending on the quality of the measurements and the magnitude of the expected effects. Thus, determining the minimum sample size in SEM is not a simple application of rules, but rather a process that requires methodological judgment. In this sense, this study seeks to synthesize the various approaches to sample calculation and explore practical perspectives on their application, considering the context in which they can be used. Although there are studies that discuss minimum sample requirements, such as Hox et al. (2010), Nicolaou & Masoner (2013), and Ramos-Vera (2022), few offer an integrated view that combines theoretical foundations, statistical criteria, methodological limitations, and the practical translation of these guidelines into real research scenarios. This gap provides an opportunity for this study to offer a relevant and distinctive contribution.

The rest of this article is organized as follows: first, a theoretical contextualization of different methods for estimating the minimum sample size in SEM is presented. Next, the methodology adopted in this study is presented, indicating its three phases. After that, the results are presented, and their relevance and practical implications are discussed. Finally, the main conclusions are summarized and suggestions for future work are provided.

Literature Review

Calculating the minimum sample size in SEM involves a diverse set of methods that reflect both statistical advances and practical needs in empirical research. There is no universally accepted criterion, which makes the topic particularly complex and relevant. Literature in this field, such as Wang and Rhemtulla (2021) and Wolf et al. (2013), suggests that the choice of method should incorporate various elements, such as the type of model, the number of constructs considered, the variability of the sample data, and the type of estimator used. It is recognized that these different approaches offer complementary perspectives on what constitutes a sufficiently robust sample to ensure valid results.

One of the most traditional methods is that of practical rules based on ratios between the number of participants and the number of parameters or indicators. Among the best known is the recommendation of at least 10 cases per parameter or indicator, although some authors propose more conservative ratios, such as 15 or 20 cases per variable (Ali Memon et al., 2020; Iacobucci, 2010; Kush et al., 2021). Although simple, this approach has been criticized for not considering the complexity of the model or the specific statistical properties of the data. Even so, it continues to be used as an initial reference, especially in exploratory studies or in preliminary planning stages.

Another relevant approach is statistical power analysis, which seeks to determine the sample size needed to detect effects of a certain magnitude with desirable levels of confidence and power. This method uses previously estimated or expected values, such as factor loadings, correlations, and structural effects, allowing for a more rigorous assessment of the model's ability to detect real relationships in the data (Buchberger et al., 2024; Flora et al., 2025). Power analysis is particularly useful in confirmatory studies, but requires prior knowledge of the parameters, which is not always available.

Monte Carlo simulations represent one of the most advanced and methodologically robust approaches. Traditionally, a Monte Carlo simulation is viewed as a computational method used to understand the behavior of a system or model by running a large number of random simulations (Graf & Korn, 2020; Harrison, 2010). It relies on repeated random sampling to estimate how a model performs under different conditions. It is particularly useful to assess uncertainty, variability, or risk in predictions (Schamberger, 2023). In the context of SEM, Monte Carlo simulations are used to: (i) test how sample size affects parameter estimates and model fit; (ii) examine the impact of non-normal data, missing values, or measurement error; and (iii) evaluate power and stability of complex models, including those with second-order factors, multigroup comparisons, or longitudinal designs. Serang (2020) notes that this method allows for the consideration of complex factors, such as data non-normality and distinct factor loadings. Despite its high accuracy, the use of simulations requires more advanced statistical skills and specific tools, which limit its applicability.

An additional approach involves recommendations based on the type of estimator used. Estimators such as Maximum Likelihood, as discussed by Maydeu-Olivares (2017), tend to require larger samples, especially in non-normal contexts, while robust or asymptotic distribution-based methods can tolerate smaller samples, albeit with potential trade-offs in accuracy. The Weighted Least Squares Mean and Variance Adjusted (WLSMV) estimator, for example, is often recommended for categorical data, but requires relatively large sample sizes to ensure stability (DiStefano & Morgan, 2014).

Another perspective is based on the structural complexity of the model. Models with many constructs, multiple paths, or second-order latent variables require significantly larger samples than simple models. Each of these parameters requires sufficient information to be calculated accurately. When the sample is small relative to the number of parameters, the model tends to exhibit instability, convergence difficulties, and unreliable estimates. In addition, fit indices such as CFI, TLI, or RMSEA become more sensitive to fluctuations, producing results that may mistakenly suggest a poor fit. van Kesteren & Oberski (2019) point out that complexity also increases when there are multiple structural paths between constructs. In models that test direct, indirect, moderating, and mediating effects simultaneously, the amount of information needed to capture these relationships grows exponentially. Indirect effects, for example, are particularly sensitive to sample size because they involve the multiplication of coefficients, which amplifies estimation errors. Thus, it can be concluded that small samples can lead to underestimation or overestimation of these effects, compromising theoretical conclusions. In addition, another factor of complexity arises that is related to second-order latent variables (Koufteros et al., 2009; Yang et al., 2020). These structures, which aggregate several first-order constructs, add new layers to the model, requiring more parameters and increasing the difficulty of stabilizing the model with few observations.

Materials and Methods

This study applies an exploratory methodology that aims to compare different methods for calculating the minimum sample size in SEM. This approach is particularly appropriate in a field where there is no consensus and where recommendations vary substantially depending on the model and research objectives. Thus, the proposed methodology recognizes that calculating sample size is not a purely technical issue, but rather a process that requires coordination between theory, practice, and empirical evaluation.

Figure 1 presents visually the methodological process. The first component of this approach involves a narrative literature review of the main existing methods. This step allows us to map different families of methods, such as practical rules, power analyses, criteria based on the number of parameters, and Monte Carlo simulations. By critically exploring their premises and contexts of applicability, the study provides a solid theoretical basis that helps to understand why different methods produce different recommendations. This review also clarifies gaps and inconsistencies in the literature, justifying the need for additional empirical comparisons.

The second component of the study is centered on the practical application of sample size determination methodologies through the development of structured simulation-based scenarios that reflect realistic research conditions. These scenarios are designed to mirror common challenges encountered in empirical studies, including varying levels of structural complexity such as first- and second-order latent variable models, the presence of mediation effects, multigroup comparisons, non-normal data distributions, and models incorporating categorical or ordinal indicators. However, it is important to clarify that, although the study adopts a simulation-oriented design logic, it does not present full Monte Carlo simulation outputs or empirically generated sampling distributions. Instead, the term "simulation" is used in a conceptual and scenario-building sense, where hypothetical but realistic model configurations are constructed to examine how different sample size estimation approaches behave under diverse methodological conditions. This includes exploring how practical rules, statistical power analysis, Monte Carlo-based reasoning, estimator-specific recommendations, and model complexity considerations may vary when applied to increasingly demanding analytical contexts. Therefore, each scenario functions as a controlled analytical framework rather than a computational simulation experiment, allowing for structured comparison across methodological approaches. Additionally, this study creates a set of analytically grounded situations that reflect the heterogeneity of contemporary empirical research in social sciences, management, education, and engineering.

The third component of the study consists of a structured comparative evaluation of the sample size estimates produced by the different methodological approaches across the defined scenarios. This analytical step goes beyond simple observation of convergence or divergence between methods, aiming instead to examine the magnitude, direction, and consistency of differences under varying model conditions. In practice, the comparison was conducted by applying each estimation approach (e.g., practical rules, statistical power analysis, estimator-based recommendations, and model complexity guidelines) to the same set of scenario specifications. This ensured a common analytical baseline, allowing results to be directly contrasted across methods within identical structural conditions. To enhance analytical rigor and reduce potential bias in interpretation, several strategies were adopted. First, all scenarios were defined a priori and kept constant across methods, preventing post hoc adjustments that could favor a particular approach. Second, comparisons were conducted using a standardized parameter framework, ensuring that differences in sample size recommendations were not driven by inconsistencies in model specification (e.g., number of latent variables, indicators, or structural paths). Third, the analysis emphasized relative comparison patterns rather than absolute numerical differences alone, focusing on whether methods systematically produced more conservative (larger sample requirements) or more liberal (smaller sample requirements) estimates under identical conditions. This assessment is not limited to a numerical comparison of the recommended values; it also involves analyzing the accuracy of the model estimates, the stability of the parameters, the behavior of the goodness-of-fit indices, and the ability of each method to deal with problems reported by Kelava et al. (2008) and Tarka (2017), such as non-normality, multicollinearity, or models with high structural complexity. In addition, this step allows us to assess the sensitivity of the methods to the specific characteristics of the scenarios. For example, if you want to explore how rule-based methods behave when the model contains second-order variables, or how power analysis reacts when the expected effects are small or moderate.

The combination of these three components fully justifies a multi-method approach. As Hammond (2005) points out, a multi-method approach combines two or more methods of data collection, analysis, or interpretation within the same project, with the aim of obtaining a more complete understanding of the phenomenon under study. Thus, rather than relying exclusively on a single method (e.g., whether quantitative, qualitative, or experimental), a multi-method study integrates different methodological strategies to explore a research question from multiple perspectives. Applied to the specific context of this study, it is recognized that no single method can capture all the factors that influence the adequacy of the sample in SEM. Therefore, this multi-method methodology not only enriches the analytical rigor of the study but also increases its usefulness for researchers who need well-founded recommendations that are adaptable to the specific context of their investigations.

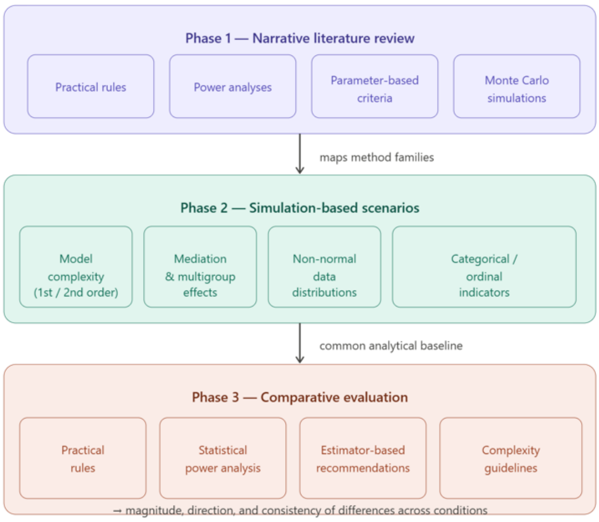


Fig. 1. Phases of the methodological process

Results

Table 1 provides a comparative analysis of different methods to sample size estimation in SEM. Findings reveal that each method for estimating sample size in SEM serves different purposes and fits different research contexts. Practical rules based on participant-to-parameter ratios function mainly as an initial approximation rather than a definitive guideline. They are best used at the earliest planning stage, offering a quick benchmark but lacking the precision required for rigorous studies. Statistical power analysis is more appropriate when the researcher has prior information about expected effect sizes and wishes to ensure sufficient power and confidence. This method is especially valuable in studies where the detection of specific relationships is crucial and where analytical rigor is required. Monte Carlo simulations are particularly suitable for advanced research and complex SEM models. Their strength lies in the ability to empirically test how the model behaves across different sample sizes and data conditions, making them ideal when precision and methodological robustness are priorities. Researchers dealing with innovative, intricate, or high-stakes models often rely on this approach. Recommendations based on the type of estimator used become relevant when the properties of the data are well defined. For instance, models estimated using Maximum Likelihood tend to perform adequately with moderate sample sizes under normality, making this approach suitable when such assumptions are met. Conversely, when data are categorical, the WLSMV estimator becomes the most appropriate choice, though it typically requires considerably larger samples to ensure stability and accuracy. Finally, considering the structural complexity of the model is essential when working with large-scale, multifactorial, or hierarchical SEM designs. As complexity increases, so does the required sample size, making this perspective indispensable for ensuring reliable estimation in sophisticated models.

Figure 2 provides a methodological flowchart that guides researchers in selecting the most appropriate sample-size estimation approach according to model characteristics, data properties, and research objectives. The diagram poses four sequential binary questions related to the researcher's

Table 1. Comparative Analysis of Sample Size Estimation Methods in SEM

Method	Description	Advantages	Limitations
Practical rules (participant-to-parameter ratios)	Based on fixed ratios, such as 10 to 20 participants per parameter or indicator.	Simple, quick, and widely used as an initial reference.	Ignores model complexity and effect sizes. May lead to under- or overestimated samples.
Statistical power analysis	Calculates the sample size needed to detect specific effects with a desired level of confidence and power.	More rigorous and tailored to expected effect sizes.	Requires precise definition of effect magnitudes.
Monte Carlo simulations	Tests model performance under different sample size, variability, and structural scenarios.	Highly robust and empirically grounded. Provides very precise estimates.	Requires software, computation time, and technical expertise. High level of complexity.
Estimator-based recommendations (e.g., ML)	Sample size depends on the characteristics and assumptions of the estimator used.	Can provide estimates better aligned with data characteristics.	Strongly dependent on assumption compliance. Different estimators vary widely in sample demands.
WLSMV for categorical data	Robust estimator for ordinal or categorical variables, with mean and variance adjustments.	Excellent for categorical data. Reduces bias associated with ML in non-continuous variables.	Requires large samples for stable estimates, especially with few categories or many factors.
Model structural complexity	Considers the number of constructs, paths, factor loadings, and presence of second-order latent variables.	Accounts for the actual structural demands of the model. Adjusts sample size to model complexity.	May require very large samples. Difficult to use in isolation without other methodologies.

situation and model characteristics. The first question determines whether the study is still in an early planning stage, directing those who are toward practical rules based on participant-to-parameter ratios. The second question asks whether expected effect sizes are known, routing researchers with that prior knowledge toward statistical power analysis. The third question identifies whether outcome variables are categorical or ordinal, pointing to the WLSMV estimator in such cases. The fourth and final question addresses whether the model is large-scale, multifactorial, or hierarchical, recommending a structural complexity approach when that is true. Researchers who answer no to all four questions are directed to the Maximum Likelihood estimator, suited for continuous data under normality assumptions. It is also recommended the combination of methods for greater precision reflecting the complementary nature of the approaches presented in Table 1.

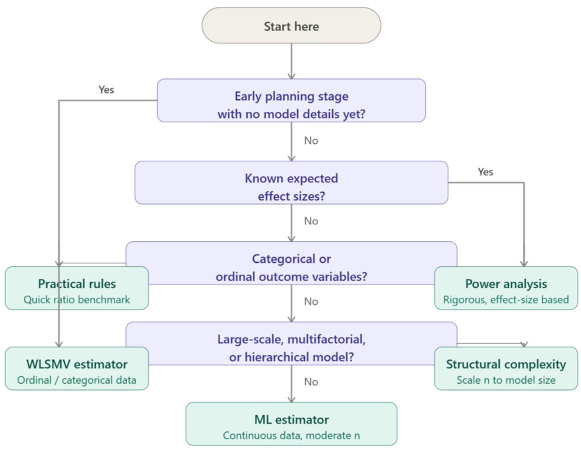


Fig. 2. Methodological flowchart for selecting sample size estimation approaches

The results are further explored considering practical scenarios for the use of SEM in areas such as engineering, management, social sciences, and education. It is important to demonstrate how SEM has broad applicability in multiple areas. Table 2 provides practical guidance on estimating sample sizes for SEM across different methods, scenarios, and fields. However, it is important to clarify that these recommendations are not derived from a single unified empirical dataset or formal simulation study, but rather from a structured synthesis of widely reported guidelines in the SEM literature, combined with established methodological conventions.

The synthesis indicates a relatively stable convergence of recommended minimum sample sizes across SEM traditions, with variability primarily driven by estimator choice and effect size assumptions. Across the reviewed methodological benchmarks, the practical rules approach (10–20 observations per estimated parameter) typically translates into approximate sample sizes ranging from 50–200 (simple models with 5–15 parameters) to 300–400 (more parameter-intensive models). This rule is consistent with early-stage exploratory SEM applications, where precision is secondary to model feasibility and identification.

The statistical power analysis framework, assuming a conventional medium effect size ($f^2 = 0.15$), $\alpha = 0.05$, and power = 0.80, produces more formalized estimates in the range of 150–250 participants for standard SEM models, increasing to 250–400 participants when structural complexity or lower effect sizes are considered. More specifically, this framework derives its robustness from its explicit grounding in Type I and Type II error control, ensuring that the probability of incorrectly rejecting a true model effect (α error) and failing to detect an existing effect (β error) is explicitly regulated. Compared to rule-of-thumb approaches, power analysis incorporates effect size sensitivity, meaning that even moderate reductions in assumed effect magnitude (e.g., from $f^2 = 0.15$ to $f^2 = 0.10$) can lead to substantial increases in required sample size, often exceeding 300–450 participants depending on model specification and number of latent constructs. In SEM contexts, these estimates are further influenced by model complexity parameters. Models with higher parameter density or weaker indicator loadings typically require larger samples to maintain sufficient statistical power for detecting both direct and indirect effects. In mediation models, for instance, indirect effects tend to have lower statistical power than direct effects, often necessitating an additional 20–40% increase in sample size to achieve equivalent detection capability. Moreover, power-based SEM recommendations assume adequate measurement quality, including acceptable reliability levels (e.g., factor loadings ≥ 0.60–0.70). When measurement error increases, the effective power of the model decreases, thereby inflating required sample sizes even further. This reinforces the importance of integrating measurement model assessment with power calculations, rather than treating sample size determination as a purely structural issue.

The Monte Carlo simulation literature (as reported in prior validated studies) indicates more robust and empirically stable requirements, typically ranging from 250–600 participants, with the upper bound associated with conditions of non-normality, higher model misspecification sensitivity, or longitudinal/multigroup structures. These estimates reflect convergence criteria such as parameter bias (<10%), standard error stability, and acceptable convergence rates (>95%), as commonly reported in simulation-based SEM validation studies. Effectively, Monte Carlo approaches derive their strength from the ability to replicate the sampling process under controlled population conditions. Within this framework, sample size requirements are not treated as fixed thresholds but as performance-dependent outcomes, defined by whether predefined statistical criteria are satisfied across a large number of replications (e.g., 1,000–5,000 simulated samples). In SEM applications, acceptable performance is typically evaluated using multiple diagnostics, including relative bias, root mean square error (RMSE), standard error accuracy, coverage probability of confidence intervals (typically ≥ 0.90–0.95), and empirical power stability across replications. When these criteria are jointly considered, sample size recommendations tend to increase as model complexity escalates. In particular, multigroup SEM models and longitudinal growth models tend to require larger samples (often 400–700+ observations) due to the need to simultaneously estimate group-specific or time-varying parameters. Similarly, models involving non-normal indicators or categorical data structures exhibit increased estimation instability at smaller sample sizes, often requiring upward adjustments in the recommended range to maintain convergence reliability and reduce parameter distortion.

From an estimator-specific perspective, Maximum Likelihood (ML) estimation demonstrates adequate performance with 200–300 observations under continuous and approximately normal data conditions, while the WLSMV estimator for ordinal/categorical indicators requires approximately 350–550 observations to maintain parameter stability and robust standard errors. These differences reflect documented sensitivity of categorical estimators to sample size and distributional asymmetry, particularly in models with multiple latent constructs. ML estimation is grounded in the assumption of multivariate normality and sufficient sample size for asymptotic properties to hold, meaning that parameter estimates become increasingly unbiased and efficient as sample size increases. Under conditions of mild to moderate non-normality, ML can still perform adequately, particularly when robust corrections (e.g., Satorra–Bentler scaled statistics or robust standard errors) are applied. However, when sample sizes approach the lower bound of the recommended range (≈200), ML estimates become more sensitive to model misspecification and measurement error, potentially affecting fit indices and standard error precision. In contrast, WLSMV (Weighted Least Squares Mean and Variance adjusted) estimation is specifically designed for ordinal or categorical indicators, where the underlying assumption is that observed responses reflect discretized versions of continuous latent response variables. This estimation approach relies heavily on accurate threshold estimation and polychoric correlations, both of which are highly sensitive to sample size. As a result, smaller samples can lead to unstable threshold estimates, inflated standard errors, and reduced convergence reliability, particularly in models with a large number of categories or indicators per latent construct.

Finally, model structural complexity introduces a consistent upward adjustment factor of approximately +30% to +60% in required sample size, particularly in second-order, multigroup, or longitudinal SEM designs. Under these conditions, recommended samples frequently reach 400–600 participants, with more complex longitudinal trajectories extending to 600–750 participants to ensure stable estimation of growth or group-specific parameters. This upward adjustment is primarily driven by the fact that structural complexity increases the total number of free parameters to be estimated, including additional factor loadings, structural paths, intercepts, variances, covariances, and, in multigroup settings, group-specific parameter sets. As the parameter-to-sample ratio becomes more constrained, estimation precision decreases unless compensated by larger samples. This effect is particularly pronounced in models with latent interactions, hierarchical factor structures, or cross-group equality constraints, where identification conditions require stronger empirical support. Moreover, in second-order SEM models, the introduction of higher-order latent constructs amplifies estimation demands because each higher-order factor aggregates multiple first-order constructs, thereby increasing the dependency structure within the model. This leads to greater sensitivity to sampling variability, especially when first-order factor loadings are moderate or unevenly distributed. Consequently, sample sizes closer to the upper bound (≈500–600) are typically required to maintain stable second-order parameter estimates and acceptable standard errors. Additionally, in multigroup SEM applications, sample size requirements are further influenced by the need to ensure measurement invariance across groups. Each additional group effectively reduces the available sample per estimated

Table 2. Practical Scenarios for the Application of SEM and Estimation the Sample Size

Method	Scenario I	Scenario II	Scenario III
Practical rules (participant-to-parameter ratios)	Exploratory research in startup incubators to map entrepreneurial competencies; ~150–200 participants (assuming 15 parameters).	Pilot with small software teams; ~80–120 participants for ~10 parameters.	Student satisfaction in a small class; ~50–100 participants for ~5–7 parameters.
Statistical power analysis	Leadership training effect on teams; ~200–250 participants to detect medium effect size (f^2 = 0.15) with 0.80 power.	Community program evaluation; ~180–220 participants to detect medium effects with α = 0.05 and power 0.80.	Pedagogical intervention on student skills; ~250–400 participants to detect medium-to-large effects.
Monte Carlo simulations	Complex model that involves culture, networking, and innovation constructs; ~300–500 participants for stable parameter estimates.	Reliability of complex systems; ~250–400 participants depending on non-normality and latent variable structure.	Longitudinal urban mobility study; ~400–600 participants for stable growth trajectories.
Estimator-based recommendations (e.g., ML)	Customer satisfaction surveys with continuous data; ~300–400 participants for moderate complexity.	Industrial process comparison with continuous data from less than 3 sectors; ~200–300 participants.	Well-being factor analysis; ~350–500 participants for stable ML estimation.
WLSMV for categorical data	Likert-scale entrepreneurial competencies; ~400–500 participants due to ordinal items.	School climate evaluation; ~350–450 participants to ensure stability across multiple classes.	Hierarchical public policy perception models; ~450–550 participants for robust WLSMV estimation.
Model structural complexity	Second-order model of motivation, leadership, networking; ~400–600 participants.	Multigroup model comparing departments; ~250–400 participants depending on groups and paths.	Longitudinal model of student competencies; ~500–750 participants for trajectory estimation.

parameter, meaning that total sample size must scale with both the number of groups and the complexity of cross-group constraints.

Discussion

This study offers an analysis and comparison of different methods for estimating sample size in SEM. This is a central theme for ensuring validity, reliability, and methodological rigor in complex quantitative research. Several approaches are explored, ranging from simple practical rules to highly robust Monte Carlo simulations. The results obtained in this study are important for demonstrating how the choice of method directly impacts the accuracy of estimates, the stability of parameters, and statistical power. These are fundamental aspects for different fields of application, including management, engineering, education, and social sciences.

The study shows that traditional methods, such as practical rules based on the participant-parameter ratio, are useful in early stages or in exploratory research with limited resources, offering quick estimates of sample size, usually between 50 and 200 participants. However, current literature indicates that such approaches lack rigor and may underestimate or overestimate sample requirements in complex models (Schuberth et al., 2023). Thus, although useful, these estimates should be used sparingly and not abusively, particularly when the population size we are inferring is significant. Therefore, it can be concluded that the use of fixed rules can lead to underestimation or overestimation of sample size, especially in complex, multifactorial, or longitudinal models. Underestimating the sample can result in unstable parameters, high type I or type II errors, and unreliable conclusions (Akobeng, 2016; Darling, 2024; Rothman, 2010).

Statistical power analysis represents a significant methodological advance over traditional rules of estimation, as it allows one to calculate the sample size needed to detect effects of a specific magnitude with predetermined levels of confidence (usually $\alpha = 0.05$) and power (typically 0.80 or higher). Rather than providing only a rough estimate based on generic ratios, this approach explicitly considers the strength of the expected effects, which is critical to ensuring that statistical tests have a sufficient probability of identifying real relationships between variables, minimizing the risk of type II errors. This approach is particularly useful in confirmatory studies, in which the researcher already has clear hypotheses and wishes to verify specific relationships between constructs. For example, in corporate leadership programs, power analysis allows the calculation of the number of participants needed to detect whether a training intervention actually improves team performance or job satisfaction. Similarly, in educational interventions, such as critical skills teaching programs or active learning methods, power analysis helps define a sample capable of revealing statistically significant effects on academic performance or student motivation. Furthermore, power analysis is aligned with classic recommendations by McQuitty (2004), who emphasizes that studies with insufficient power tend to produce unreliable results, making it difficult to replicate and generalize conclusions. Subsequent SEM research, such as that by Cole (2025) and Irmer et al. (2024), reinforces that the use of power analysis is essential for confirmatory and multifactorial models, as it allows for balancing statistical rigor with practical feasibility, avoiding both undersizing (which compromises the detection of real effects) and oversizing (which can waste resources).

Monte Carlo simulations stand out as an empirically robust method capable of testing model stability in different scenarios of sample size and data distribution. More recent studies emphasize their value in complex or longitudinal models, especially when traditional assumptions of normality are not met (Hsu et al., 2019; Schultzberg & Muthén, 2017). Similarly, recommendations based on the type of estimator, such as Maximum Likelihood or WLSMV, reinforce that the adequacy of the sample size depends on the characteristics of the data. WLSMV, for example, requires considerably larger samples for categorical data, reflecting the need for alignment between model, data, and estimator.

Considering the structural complexity of the model is one of the most critical factors in defining the sample size in SEM, as more elaborate models place additional demands on the data and the accuracy of the estimates. Structural complexity refers to elements such as the number of latent constructs, the number of indicators per construct, the presence of second-order factors, multigroup structures, or longitudinal models that estimate trajectories over time. Each of these elements significantly increases the number of parameters to be estimated, which, in turn, increases the need for a larger sample to ensure statistical stability and reliability. For example, in second-order models, where higher-level latent constructs are formed from first-level factors, the number of paths and factor loadings to be estimated grows nonlinearly in relation to the number of observed variables. This implies that a sample that would be sufficient for a first-order model may be inadequate, resulting in unstable parameters, biased estimates, and wide confidence intervals. In multigroup models, complexity increases even further, as it is necessary to ensure stability in each group analyzed, in addition to allowing comparisons between groups, which require significantly larger samples. Longitudinal models, in turn, require sufficient participants to capture temporal trajectories, considering possible losses due to dropouts and variability over time. The current literature corroborates this need for larger samples in complex models. Studies such as those by Montoya & Edwards (2020) indicate that, for models with multiple factors, indicators, and groups, it is common for the required sample size to exceed 400 to 600 participants, reaching 700 or more in longitudinal or extensive second-order models. Moreover, Savalei (2020) shows that when using robust estimators or categorical data, this requirement may be even greater, as estimators need more information to produce reliable results.

Finally, we recognize the potential of this article in offering practical implications. First, it contributes to greater rigor in defining sample size. The aim is to help researchers plan data collection in a more scientific and informed manner, considering the characteristics of the model, structural complexity, type of data, and estimator used. This rigor avoids arbitrary decisions or those based solely on conventions, such as the generic rule of "10 to 20 participants per parameter." Although the sample size ranges presented in this study are derived from established methodological literature and widely accepted SEM guidelines, they should be viewed as indicative reference values rather than universally applicable thresholds. Second, this study seeks to contribute to better resource allocation. SEM research can involve significant costs, both in terms of money and in terms of time and human effort. More accurate sample planning avoids unnecessarily large data collections, which consume resources without generating proportional gains in accuracy, and also avoids insufficient samples, which would require repeating studies or collecting additional data later. Thus, by adequately estimating the sample size, it is possible to optimize the budget, reduce logistical effort, and increase the operational efficiency of research projects. Furthermore, this situation is highly relevant in applied studies in management, engineering, education, or social sciences, where the availability of participants may be limited. Lastly, it also aims to contribute to ensuring robust and replicable results. An adequate sample size increases the stability of parameter estimates, reduces the risk of type I and type II errors, and improves the generalization of results to the target population. This is essential for confirmatory research and longitudinal studies, in which reliable conclusions depend on the consistency of estimates over time and across different groups. Furthermore, replicability, which has become a central criterion of scientific quality as demonstrated by Chakravorti et al. (2025), Epskamp (2019), and Rahal et al. (2022), is only possible when the sample allows real effects to be detected with statistical confidence. Although adequate sample size is a necessary condition for robust and replicable SEM results, it is not sufficient by itself. Replicability and statistical robustness also depend on factors such as model specification, measurement quality, construct reliability and validity, estimator appropriateness, treatment of missing data, and compliance with distributional assumptions (Guenther et al., 2025; Yuan et al., 2015). Consequently, the recommendations proposed in this study should be interpreted as methodological guidance that supports research planning rather than as guarantees of reproducible findings.

Conclusions

This study highlights the importance of adopting a systematic and careful approach to estimating sample size in SEM, showing that different methods offer varying levels of accuracy and applicability. Traditional rules of estimation are useful in exploratory phases, offering quick estimates, but they have limitations in complex models and can lead to underestimation or overestimation of sample requirements. Statistical power analysis represents an advance, allowing the calculation of samples capable of detecting specific effects with defined levels of confidence and power, and is particularly suitable for confirmatory studies in areas such as management and education. Monte Carlo simulations offer empirical robustness for complex or longitudinal models, while recommendations based on the type of estimator, such as ML or WLSMV, highlight the need to align sample size with data characteristics. Finally, consideration of the structural complexity of the model emerges as a central factor: second-order, multigroup, or longitudinal models require substantially larger samples to ensure parameter stability and reliability of estimates, often above 400–600 participants, reaching 700 or more in extensive models.

Finally, some suggestions for future work are provided. Several avenues can be explored to deepen and expand the findings of this study. First, comparative empirical research between sample size estimation methods in specific application contexts is recommended, with the aim of verifying whether the general recommendations hold true in different populations and data types. Another important line of research would be to evaluate the performance of the methods in highly complex or longitudinal models, including second-order factors and multigroup structures, especially considering categorical or non-normal data, to validate and refine the proposed estimates. In addition, future studies could explore the impact of sample size on measures of validity and reliability in SEM. This type of study could assess how different estimates affect parameters, factor loadings, and fit indices in real-world scenarios. It would also be relevant to develop interactive tools or guides to assist researchers in defining ideal samples, in which various variables such as estimator type, model complexity, and magnitude of expected effects can be incorporated.

Ethics Approval

Not applicable. This study is based on a narrative review and conceptual scenario-based methodological analysis and did not involve human participants, animals, or the collection of identifiable personal data; therefore, ethics committee approval was not required.

Informed Consent

Not applicable. The study did not involve human participants or the collection of personal data.

Data Availability Statement

Data are not applicable because no new empirical dataset was created or analyzed in this study. The article is based on published literature and conceptual methodological scenarios.

AI Transparency Statement

The author did not use AI-assisted tools in the preparation of this manuscript.

Acknowledgements

Not applicable.

Conflicts of Interest

The author declares no conflicts of interest.

References

Nusair, K., & Hua, N. (2010). Comparative assessment of structural equation modeling and multiple regression research methodologies: E-commerce context. *Tourism Management, 31*(3), 314–324. https://doi.org/10.1016/j.tourman.2009.03.010

Xiong, B., Skitmore, M.R., & Xia, B. (2015). A critical review of structural equation modeling applications in construction research. *Automation in Construction,* 49, 59–70. https://doi.org/10.1016/j.autcon.2014.09.006

Allammari, Y., Taqi, A., & El Morabit, N. (2025). Entrepreneurial orientation and SME performance: A sequential mediation analysis of market and learning orientations. *Journal of Applied Structural Equation Modeling, 9*(2), 1–30. https://doi.org/10.47263/JASEM.9(2)02

Dulaimi, M.F., Nepal, M.P., & Park, M. (2005). A hierarchical structural model of assessing innovation and project performance. *Construction Management and Economics, 23*(6), 565–577. https://doi.org/10.1080/01446190500126684

Ghardashi, F., Yaghoubi, M., Teymourzadeh, E., Bahadori, M., & Salesi, M. (2022). The innovation capability model in higher education: A structural equation modelling approach. *African Journal of Science, Technology, Innovation and Development, 15*(4), 473–481. https://doi.org/10.1080/20421338.2022.2125594

Kline, R. (2010). *Principles and Practice of Structural Equation Modeling.* Guilford Press.

Cheung, G.W., Cooper-Thomas, H.D., Lau, R.S., & Wang, L.C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management, 41*(2), 745–783. https://doi.org/10.1007/s10490–023–09871-y

Newsom, J.T. (2023). *Longitudinal Structural Equation Modeling: A Comprehensive Introduction.* Routledge.

Panzeri, A., Castelnuovo, G., & Spoto, A. (2024). Assessing Discriminant Validity through Structural Equation Modeling: The Case of Eating Compulsivity. Nutrients, 16(4), 550. https://doi.org/10.3390/nu16040550

Kim, K.H. (2005). The Relation Among Fit Indexes, Power, and Sample Size in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 12*(3), 368–390. https://doi.org/10.1207/s15328007sem1203_2

Hox, J.J., Maas, C.J.M., & Brinkhuis, M.J.S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica,* 64, 157–170. https://doi.org/10.1111/j.1467–9574.2009.00445.x

Nicolaou, A., & Masoner, M. (2013). Sample size requirements in structural equation models under standard conditions. *International Journal of Accounting Information Systems, 14*(4), 256–274. https://doi.org/10.1016/j.accinf.2013.11.001

Ramos-Vera, C.A. (2022). Beyond sample size estimation in clinical univariate analysis. An online calculation for structural equation modeling and network analysis on latent and observable variables. *Nutrición Hospitalaria, 38*(6), 1310. https://doi.org/10.20960/nh.03751

Wang, Y.A., & Rhemtulla, M. (2021). Power Analysis for Parameter Estimation in Structural Equation Modeling: A Discussion and Tutorial. *Advances in Methods and Practices in Psychological Science, 4*(1), 1–17. https://doi.org/10.1177/2515245920918253

Wolf, E., Harrington, K., Clark, S., & Miller, M. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educ Psychol Meas, 76*(6), 913–934. https://doi.org/10.1177/0013164413495237

Ali Memon, M., Ting, H., Cheah, J.H., Ramayah, T., Chuah, F., & Cham, T.H. (2020). Sample Size for Survey Research: Review and Recommendations. *Journal of Applied Structural Equation Modeling, 4*(2), 1–20. https://doi.org/10.47263/JASEM.4(2)01

Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology, 20*(1), 90–98. https://doi.org/10.1016/j.jcps.2009.09.003

Kush, J.M., Konold, T.R., & Bradshaw, C.P. (2021). The Sampling Ratio in Multilevel Structural Equation Models: Considerations to Inform Study Design. *Educational and Psychological Measurement, 82*(3), 409–443. https://doi.org/10.1177/00131644211020112

Buchberger, E., Ngo, C.T., Peikert, A., Brandmaier, A., & Werkle-Bergner, M. (2024). Estimating statistical power for structural equation models in developmental cognitive science: A tutorial in R. *Behavior Research Methods,* 56, 1–18. https://doi.org/10.3758/s13428–024–02396–2

Flora, D.B., Crone, G., & Bell, S.M. (2025). Effect Size Interpretation in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal, 32*(6), 1069–1076. https://doi.org/10.1080/10705511.2025.2459768

Graf, S., & Korn, R. (2020). A guide to Monte Carlo simulation concepts for assessment of risk-return profiles for regulatory purposes. *European Actuarial Journal,* 10, 273–293. https://doi.org/10.1007/s13385–020–00232–3

Harrison, R.L. (2010). Introduction To Monte Carlo Simulation. *AIP Conf Proc,* 1204, 17–21. https://doi.org/10.1063/1.3295638

Schamberger, T. (2023). Conducting Monte Carlo simulations with PLS-PM and other variance-based estimators for structural equation models: a tutorial using the R package cSEM. *Industrial Management & Data Systems, 123*(6), 1789–1813. https://doi.org/10.1108/IMDS-07–2022–0418

Serang, S. (2020). A Monte Carlo Test for Longitudinal Structural Equation Models in Small Samples. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(2), 165–181. https://doi.org/10.1080/10705511.2020.1756293

Maydeu-Olivares, A. (2017). Maximum Likelihood Estimation of Structural Equation Models for Continuous Data: Standard Errors and Goodness of Fit. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(3), 383–394. https://doi.org/10.1080/10705511.2016.1269606

DiStefano, C., & Morgan, G.B. (2014). A Comparison of Diagonal Weighted Least Squares Robust Estimation Techniques for Ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(3), 425–438. https://doi.org/10.1080/10705511.2014.915373

van Kesteren, E.J., & Oberski, D.L. (2019). Exploratory Mediation Analysis with Many Potential Mediators. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(5), 710–723. https://doi.org/10.1080/10705511.2019.1588124

Koufteros, X., Babbar, S., & Kaighobadi, M. (2009). A paradigm for examining second-order factor models employing structural equation modeling. *International Journal of Production Economics, 120*(2), 633–652. https://doi.org/10.1016/j.ijpe.2009.04.010

Yang, Y., Luo, Y., & Zhang, Q. (2020). A Cautionary Note on Identification and Scaling Issues in Second-order Latent Growth Models. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(2), 302–313. https://doi.org/10.1080/10705511.2020.1747938

Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 4*(2), 51–66. https://doi.org/10.1027/1614–2241.4.2.51

Tarka, P. (2017). An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences. *Quality and Quantity, 52*(1), 313–354. https://doi.org/10.1007/s11135–017–0469–8

Hammond, C. (2005). The Wider Benefits of Adult Learning: An Illustration of the Advantages of Multi-method Research. *International Journal of Social Research Methodology, 8*(3), 239–255. https://doi.org/10.1080/13645570500155037

Schuberth, F., Hubona, G., Roemer, E., Zaza, S., Schamberger, T., Chuah, F., Cepeda-Carrión, G., & Henseler, J. (2023). The choice of structural equation modeling technique matters: A commentary on Dash and Paul (2021). *Technological Forecasting and Social Change,* 194, 122665. https://doi.org/10.1016/j.techfore.2023.122665

Akobeng, A.K. (2016). Understanding type I and type II errors, statistical power and sample size. *Acta Paediatrica,* 105, 605–609. https://doi.org/10.1111/apa.13384

Darling, H.S. (2024). Statistical errors in scientific research: A narrative review. *Cancer Research, Statistics and Treatment, 7*(2), 241–249. https://doi.org/10.4103/crst.crst_283_23

Rothman, K.J. (2010). Curbing type I and type II errors. European Journal of Epidemiology, 25, 223–224. https://doi.org/10.1007/s10654–010–9437–5

McQuitty, S. (2004). Statistical power and structural equation models in business research. *Journal of Business Research, 57*(2), 175–183. https://doi.org/10.1016/S0148–2963(01)00301–0

Cole, D.A., Abitante, G., Kan, H., Liu, Q., Preacher, K.J., & Maxwell, S.E. (2025). Practical Problems Estimating and Reporting Power When Hypotheses Are Embedded in Complex Statistical Models. *Advances in Methods and Practices in Psychological Science, 8*(1), 1–17. https://doi.org/10.1177/25152459241302300

Irmer, J.P. (2024). Model-implied simulation-based power estimation for correctly specified and distributionally misspecified models: Applications to nonlinear and linear structural equation models. *Behavior Research Methods,* 56, 8955–8991. https://doi.org/10.3758/s13428–024–02507-z

Hsu, H.Y., Lin, J., Skidmore, S.T., & Kim, M. (2019). Evaluating fit indices in a multilevel latent growth curve model: A Monte Carlo study. *Behavior Research Methods,* 51, 172–194. https://doi.org/10.3758/s13428–018–1169–6

Schultzberg, M., & Muthén, B. (2017). Number of Subjects and Time Points Needed for Multilevel Time-Series Analysis: A Simulation Study of Dynamic Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(4), 495–515. https://doi.org/10.1080/10705511.2017.1392862

Montoya, A.K., & Edwards, M.C. (2020). The Poor Fit of Model Fit for Selecting Number of Factors in Exploratory Factor Analysis for Scale Evaluation. *Educational and Psychological Measurement, 81*(3), 413–440. https://doi.org/10.1177/0013164420942899

Savalei, V. (2020). Improving Fit Indices in Structural Equation Modeling with Categorical Data. *Multivariate Behavioral Research, 56*(3), 390–407. https://doi.org/10.1080/00273171.2020.1717922

Chakravorti, T., Koneru, S., & Rajtmajer, S. (2025). Reproducibility and replicability in research: What 452 professors think in Universities across the USA and India. *PLoS One, 20*(3), e0319334. https://doi.org/10.1371/journal.pone.0319334

Epskamp, S. (2019). Reproducibility and Replicability in a Fast-Paced Methodological World. *Advances in Methods and Practices in Psychological Science, 2*(2), 145–155. https://doi.org/10.1177/2515245919847421

Rahal, R.M., Hamann, H., Brohmer, H., & Pethig, F. (2022). Sharing the Recipe: Reproducibility and Replicability in Research Across Disciplines. *Research Ideas and Outcomes,* 8, e89980. https://doi.org/10.3897/rio.8.e89980

Guenther, P., Guenther, M., Ringle, C., Zaefarian, G., & Cartwright, S. (2025). PLS-SEM and reflective constructs: A response to recent criticism and a constructive path forward. *Industrial Marketing Management,* 128, 1–9. https://doi.org/10.1016/j.indmarman.2025.05.003

Yuan, K.H., Tong, X., & Zhang, Z. (2015). Bias and Efficiency for SEM With Missing Data and Auxiliary Variables: Two-Stage Robust Method Versus Two-Stage ML. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(2), 178–192. https://doi.org/10.1080/10705511.2014.935750

Методи оцінювання розміру вибірки в моделюванні структурними рівняннями

Фернанду Алмейда[1ABCD]

Університет Порту та Інститут системної та комп’ютерної інженерії, технологій і науки

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 11 с., 2 табл., 2 рис., 48 джерел.

Передумови. Оцінювання розміру вибірки в моделюванні структурними рівняннями (SEM) залишається дискусійним методологічним питанням. Хоча правила щодо мінімального розміру вибірки широко використовуються, вони часто недостатньо інтегрують теоретичні, статистичні та практичні міркування, особливо у складних дослідницьких дизайнах.

Мета. Це дослідження має на меті здійснити синтез і порівняльний аналіз різних підходів до оцінювання розміру вибірки в SEM, інтегруючи теоретичні засади, статистичні критерії, методологічні обмеження та контексти прикладних досліджень.

Матеріали та методи. Було застосовано мультиметодний підхід, що складався з трьох компонентів: наративного огляду методів оцінювання розміру вибірки, розгляду змодельованих сценаріїв, які відображають реалістичні умови дослідження, та порівняльного оцінювання ефективності методів у цих сценаріях.

Результати. Результати показують, що емпіричні правила є корисними для попереднього оцінювання, але можуть некоректно визначати необхідний розмір вибірки у складних моделях. Аналіз статистичної потужності забезпечує точніші та більш обґрунтовані оцінки, особливо для підтверджувальних досліджень. Симуляції Монте-Карло забезпечують надійні орієнтири для складних, лонгітюдних і багатогрупових моделей. Рекомендації, що ґрунтуються на методах оцінювання параметрів, підкреслюють важливість узгодження розміру вибірки з характеристиками даних і методами оцінювання. Складність моделі виявилася ключовим чинником, причому складні моделі SEM часто потребують 400–700 і більше учасників.

Висновки. Систематичний і контекстно орієнтований підхід до оцінювання розміру вибірки є необхідним для забезпечення надійності та методологічної строгості досліджень із використанням SEM у таких галузях, як менеджмент та освіта.

Ключові слова: методологія дослідження, моделювання структурними рівняннями, оцінювання розміру вибірки, статистична потужність, складність моделі.

Information about the Authors:

Almeida Fernando: almd@fe.up.pt; https://orcid.org/0000-0002-6758-4843; University of Porto, Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto, Rua Dr. Roberto Frias s/n. 4200-465 Porto, Portugal.

Cite this article as: Almeida, F. (2026). Methods for Estimating Sample Size in Structural Equation Modeling. *Journal of Learning Theory and Methodology, 7*(2), 80-90. https://doi.org/10.17309/jltm.2026.7.2.03

Received: 06.06.2026. Accepted: 27.06.2026. Published: 13.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Review Article

Augmented Reality in Secondary School Physical Education and Physical Literacy: Literature Review

Trisnar Adi Prabowo[1ABCDE], **Mar'atul Afifah**[1ABD], **Asyidika Vito Indarto**[1CD], **Zakwan Yumna Dhuha Pahlavi**[1BCD] and **Febriansyah Dwi Cahyo**[1ABD]

[1]Universitas Muhammadiyah Brebes

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.04

Abstract

Objectives. Although augmented reality (AR) is increasingly used in physical education, empirical evidence regarding its contribution to secondary school students' physical literacy remains scattered across different research designs, learning contexts, and outcomes. The purpose of this study was to synthesize empirical evidence regarding the use of AR in secondary school physical education and to map its contribution to outcomes conceptually related to physical literacy domains.

Materials and Methods. This study used a literature review design with reference to the PRISMA 2020 guidelines. Article searches were conducted in six databases, namely Scopus, Web of Science, ERIC, SPORTDiscus, PubMed, and Google Scholar. The search was limited to peer-reviewed journal articles published between 2015 and 2025. The keywords used combined terms related to augmented reality, physical education, secondary school, and outcomes related to physical literacy. A total of 1,061 articles were identified. After removing duplicate articles, screening titles and abstracts, assessing full texts, and evaluating eligibility, 7 articles met the inclusion criteria. Data were analyzed using narrative synthesis and thematic synthesis based on four domains of physical literacy: physical competence, motivation and confidence, knowledge and understanding, and physical activity engagement or participation. Methodological quality and risk of bias were assessed using the Mixed Methods Appraisal Tool (MMAT), and the strength of evidence was mapped using explicit coding criteria.

Results. The findings indicate that AR may support several outcomes conceptually related to physical literacy in secondary school physical education. AR-based learning can improve motor skills, sport-specific competencies, motivation, self-confidence, understanding of movement concepts, learning engagement, and physical activity participation. The strongest evidence was found in the domains of physical competence, motivation and self-confidence, and knowledge and understanding. Furthermore, AR-based game-based learning and AR-based outdoor activities also supported increased student engagement and physical activity. However, none of the included studies measured physical literacy as a complete multidimensional construct, so the results should be interpreted as evidence of physical literacy-related outcomes rather than direct evidence of overall physical literacy development.

Conclusions. AR has potential as a supportive pedagogical medium for promoting outcomes related to physical literacy in secondary school physical education. However, its effectiveness is highly dependent on the learning design, teacher preparedness, learning objectives, and the integration of AR with real-life physical activities. Future research is expected to develop more structured and long-term augmented reality interventions that explicitly measure students' physical literacy multidimensionally.

Keywords: augmented reality, physical education, physical literacy, secondary school, motor competence.

Introduction

Physical education plays a crucial role in supporting the development of secondary school students, particularly as adolescence is a transitional period marked by physical, psychological, social, and behavioral changes. In the school context, physical education serves not only to improve physical fitness but also serves as a pedagogical space for developing movement skills, cooperation, discipline, self-confidence, and active lifestyle habits (Prabowo et al., 2025). A study confirms that the provision of physical education, physical activity, and sport in schools has the potential to support adolescents' mental health, well-being, and physical activity experiences (Rocliffe et al., 2024). Therefore, physical

JLTM
LLC OVS

Corresponding Author: Trisnar Adi Prabowo,
e-mail: trisnar.prabowo@umbs.ac.id

education in secondary schools needs to be positioned as an important part of a holistic education that helps students build a foundation for participating in lifelong physical activity.

In modern physical education, physical literacy has become an important orientation because this concept views movement abilities more broadly than simply mastering motor skills. Physical literacy encompasses physical competence, motivation, self-confidence, knowledge, understanding, and engagement in meaningful physical activity. These domains indicate that physical literacy is not a single performance outcome, but a multidimensional construct that integrates physical, cognitive, affective, and behavioral aspects of participation in physical activity (Martínez-López et al., 2022). Scientific studies show that physical literacy has become a significant focus in school physical education and youth sports programs (Liu & Chen, 2021). However, its definition, measurement, and forms of intervention are still evolving, and school-based physical literacy promotion requires learning experiences that integrate movement competence, motivation, knowledge, confidence, and active engagement (Grauduszus et al., 2024; Nita Peranita et al., 2025; Nevitaningrum et al., 2026). Therefore, physical education that is oriented towards physical literacy needs to help students not only be able to perform movements, but also understand, appreciate, and be motivated to actively participate in physical activities.

However, despite its strategic role, its implementation in secondary schools still faces various challenges. Another problem is some students show low motivation, limited confidence, unequal movement competence, or insufficient engagement in movement-based learning. Scientific evidence shows that barriers to physical activity among secondary school and college students include lack of time, low motivation, and limited access to places or facilities that support physical activity (Silva et al., 2022; Alhroub et al., 2024). In addition, student involvement in physical education is greatly influenced by the quality of motivation, learning environment support, and learning strategies used by teachers (Cimenti et al., 2025; Terra et al., 2025). This condition shows that physical education learning requires a pedagogical approach that is more interesting, visual, interactive, and adaptive to the needs of secondary school students.

One approach that has the potential to address these challenges is the use of augmented reality in physical education learning. Augmented reality (AR) is a technology that combines virtual objects with real environments to provide more interactive, contextual, and immersive learning experiences. In education, studies in review articles and meta-analyses have shown that AR has a positive impact on various learning outcomes because it can provide more concrete visualization, interaction, and learning experiences (Chang et al., 2022; Yefremenko et al., 2025; Topalska & Rumenov, 2025). In physical education, AR may support learning through several mechanisms, including three-dimensional movement visualization, interactive exploration, embodied learning experiences, increased learner autonomy, and greater motivational engagement. Previous studies also indicate that virtual reality, augmented reality, and mixed reality technologies have begun to be used to support movement learning, motivation, and engagement in physical education contexts (K. E. Chang et al., 2020; Pérez-Muñoz et al., 2024; Omarov et al., 2024). With these characteristics, AR has the potential to help students understand movement techniques, visualize difficult concepts, increase motivation, and connect conceptual understanding with physical activity practice.

Numerous previous studies and reviews have addressed the use of digital technologies, augmented reality, virtual reality, and mixed reality in physical education. However, the available evidence remains scattered across contexts, educational levels, technology types, learning materials, research designs, and outcomes. This discrepancy suggests that AR studies in physical education still largely focus on instructional design and technology implementation options (Ariffin et al., 2022), while other, broader evidence still discusses and compares the use of virtual reality (VR), augmented reality (AR), and mixed reality (MR) in physical education (Pérez-Muñoz et al., 2024). On the other hand, physical literacy in physical education in schools still tends to focus on the motor skills aspect, while the application of a multidimensional physical literacy framework that includes motivation, self-confidence, knowledge, and long-term engagement is still limited (Ridha et al., 2025; Vieira et al., 2025). Moreover, many empirical AR studies in physical education do not directly measure physical literacy as a complete construct, but instead assess outcomes that are conceptually related to its domains, such as motor competence, motivation, confidence, knowledge, engagement, physical activity, and fitness. Therefore, a synthesis is still needed that specifically connects the use of AR in secondary school physical education with the realm of physical literacy.

Based on this gap, this study aims to conduct a literature review on the use of augmented reality (AR) in physical education at the secondary school level and its contribution to outcomes conceptually related to physical literacy. Specifically, this review aims to identify the forms of AR used, the context of the physical education materials developed, the results studied, and the contribution of AR to physical competence, motivation and self-confidence, knowledge and understanding, and engagement or participation in physical activity. This review does not assume that all included studies directly measured physical literacy as a multidimensional construct; rather, it maps and synthesizes AR-related evidence across domains that are theoretically aligned with physical literacy. This literature review is expected to provide a more comprehensive understanding of AR's position as a pedagogical medium in physical education. In addition, the results of this review are expected to be a basis for physical education teachers, researchers, and learning technology developers in designing AR-based learning that is not only technologically attractive but also meaningful for the development of physical literacy in secondary school students.

Materials and Methods

Research Design

This study employed a literature review design to identify, select, evaluate, and synthesize previous empirical studies on the use of augmented reality in secondary school physical education. The review focused not only on studies that directly measured physical literacy, but also on studies that examined physical literacy-related outcomes, including physical competence, motor competence, motivation, confidence, knowledge and understanding, engagement, physical activity participation, and fitness. The reporting of this literature review refers to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines so that the process of searching, selecting, and synthesizing articles can be carried out transparently, systematically, and can be replicated (Page et al., 2021; Kurniawan et al., 2024).

This literature review was chosen because the available evidence on augmented reality in physical education is still distributed across different learning contexts, intervention designs, educational levels, and outcome measures. Therefore, the review aimed to map the forms of AR used in physical education, the types of learning materials involved, the outcomes measured, and the extent to which these outcomes were theoretically aligned with the domains of physical literacy. In this study, physical literacy is understood as a multidimensional construct that includes physical competence, motivation, self-confidence, knowledge, understanding, and engagement in physical activity. However, because most included studies did not assess physical literacy as a complete construct, the synthesis was framed around physical literacy-related domains rather than direct evidence of physical literacy as a whole.

The article selection process followed the main stages of the PRISMA 2020 flow: identification, screening, eligibility, and inclusion. The initial stage was the identification of articles from all predetermined databases. Next, duplicate articles were removed using a reference management tool. Articles were then screened based on their titles and abstracts to assess their initial relevance to the research topic. Articles deemed relevant were then read in full to ensure compliance with the inclusion and exclusion criteria. This process was conducted to ensure that only studies meeting the predefined criteria regarding AR use, secondary school context, physical education or school-based physical activity, and physical literacy-related outcomes were included in the final synthesis.

Article selection was conducted independently by two reviewers. Any differences in decisions during the selection process were resolved through discussion until consensus was reached. Articles meeting all inclusion criteria were then entered into the data extraction and synthesis phase. The identification, screening, eligibility, and inclusion process is presented in the form of a PRISMA flowchart. This procedure was applied to improve transparency, minimize selection bias, and strengthen the reproducibility of the review process.

Data Sources and Search Strategy

Article searches were conducted in six electronic databases: Scopus, Web of Science, ERIC, SPORTDiscus, PubMed, and Google Scholar. These databases were selected because they encompass multidisciplinary publications in education, physical education, sport science, health, physical activity, and learning technology. Google Scholar was used as a supplementary resource to identify relevant articles that may not be indexed in the primary databases. The article search was limited to publications published between 2015 and 2025. The search was restricted to peer-reviewed journal articles written in English and focused on studies involving secondary school students or adolescent populations within school-based physical education or physical activity contexts.

The search strategy was developed by combining four groups of keywords related to augmented reality, physical education, secondary school context, and physical literacy-related outcomes. The keywords used included: "augmented reality" OR "AR technology" OR "mobile augmented reality" OR "AR application" OR "AR-based learning"; "physical education" OR "school physical education" OR "PE class" OR "physical education learning"; "secondary school" OR "middle school" OR "junior high school" OR "senior high school" OR "high school" OR "adolescent students"; serta "physical literacy" OR "physical competence" OR "motor competence" OR "movement skills" OR "motivation" OR "confidence" OR "knowledge and understanding" OR "physical activity engagement" OR "fitness". The search strings were adapted to the syntax of each database, while maintaining the same conceptual structure across all searches. In Google Scholar, searches are more selective, reviewing the most relevant articles based on title, abstract, and compliance with research criteria. The last search was conducted on June 10, 2026.

Inclusion and Exclusion Criteria

Articles included in this systematic review must meet several inclusion criteria. First, the articles must have been published between 2015 and 2025. Second, the articles must be scientific publications in peer-reviewed journals. Third, the articles must be written in English. Fourth, the articles must discuss the use of augmented reality in the context of physical education or school-based physical activity. Fifth, the research subjects must be secondary school students, such as junior high, senior high, or high school students. Sixth, the articles must report either direct measures of physical literacy or outcomes conceptually related to physical literacy domains. In this review, physical literacy-related outcomes were operationally defined as outcomes associated with physical competence, motor competence, movement skills, motivation, confidence, knowledge and understanding, engagement, physical activity participation, or fitness. Seventh, the articles must be available in full-text or downloadable manuscript form.

The inclusion of physical literacy-related outcomes was necessary because many empirical studies on augmented reality in physical education do not measure physical literacy as a complete multidimensional construct. Therefore, studies were considered eligible when their outcomes could be theoretically mapped onto one or more physical literacy domains, namely physical competence, motivation and confidence, knowledge and understanding, and engagement or physical activity participation. This operational criterion was applied to avoid overclaiming while still allowing a structured synthesis of evidence relevant to physical literacy development.

Articles were excluded from the review if they met any of the following exclusion criteria: (1) did not use augmented reality; (2) only discussed virtual reality, mixed reality, instructional videos, or general digital applications without augmented reality elements; (3) were not in the context of physical education or school-based physical activity; (4) the research subjects were not secondary school students or adolescent students in a school-based setting; (5) the article was in the form of an editorial, opinion piece, book chapter, thesis, dissertation, or conference abstract without full text; (6) did not report results relevant to physical literacy or physical literacy-related outcomes; (7) was a duplicate article; and (8) was not fully accessible.

Data Extraction

Data extraction was conducted systematically using a structured data extraction form developed by the authors. The extracted information included the author and year of publication, article title, study design, sample characteristics, educational level, form of augmented reality used, physical education or school-based physical activity context, intervention characteristics, comparison group, outcome variables, key findings, and relevance to physical literacy-related domains. The extraction process was conducted to ensure that the information obtained from each included study was comparable and relevant to the aims of the review.

The extracted outcomes were not limited to studies that directly measured physical literacy as a complete multidimensional construct. Because most included studies examined specific outcomes related to physical literacy, the extracted outcomes were categorized into four theoretically aligned domains: physical competence, motivation and confidence, knowledge and understanding, and engagement or physical activity participation. This approach was used to avoid overclaiming while still allowing a structured synthesis of how AR may contribute to outcomes conceptually related to physical literacy.

Data extraction was conducted independently by two reviewers. Any discrepancies in extracted data were discussed until consensus was reached. When necessary, the reviewers returned to the full-text articles to verify the accuracy of information related to the intervention, outcome measures, and main findings. The final extracted data were then organized in a synthesis table to support narrative and thematic analysis.

Quality Assessment and Risk of Bias

The methodological quality and risk of bias of the included studies were assessed using the Mixed Methods Appraisal Tool (MMAT) (Liao et al., 2021). MMAT was selected because the included studies used varied empirical designs, including quasi-experimental, comparative, pretest-posttest, and experimental approaches. The use of MMAT allowed the reviewers to evaluate the appropriateness of the study design, sample characteristics, data collection procedures, outcome measures, intervention implementation, and clarity of reported findings according to the methodological characteristics of each study.

The quality assessment was conducted independently by two reviewers. Each study was first assessed using the MMAT screening questions to determine whether the study had clear research questions and whether the collected data were appropriate to address those questions. After that, each article was appraised using the relevant MMAT criteria according to its research design. The appraisal focused on several aspects, including the clarity of participant selection, appropriateness of measurements, completeness of outcome data, clarity of intervention procedures, adequacy of comparison groups, and consistency between findings and conclusions.

The results of the quality assessment were not used to exclude studies from the review. Instead, they were used to support the interpretation of findings and to indicate the level of caution required when drawing conclusions. Studies with clearer intervention procedures, appropriate comparison groups, relevant outcome measures, and more complete reporting were given greater interpretive weight in the synthesis. Conversely, studies with methodological limitations, unclear reporting, or indirect relevance to physical literacy-related domains were interpreted more cautiously.

Coding Procedure for Physical Literacy-Related Domains

A coding procedure was developed to classify the contribution of each included study to the four physical literacy-related domains used in this review: physical competence, motivation and confidence, knowledge and understanding, and engagement or physical activity participation. This coding was necessary because most studies did not directly measure physical literacy as a complete construct, but instead reported specific outcomes that could be theoretically mapped onto one or more physical literacy domains.

Each included study was coded independently by two reviewers based on the outcomes reported, the relevance of those outcomes to the physical literacy-related domains, and the clarity of the study findings. The classification used four categories: strong, moderate, weak, and not reported. A domain was coded as strong when the study directly measured an outcome aligned with that domain and reported clear positive findings. A domain was coded as moderate when the study measured an outcome related to the domain but did not provide comprehensive evidence, or when the findings were positive but limited in scope. A domain was coded as weak when the relationship between the reported outcome and the domain was indirect, limited, or only briefly addressed. A domain was coded as not reported when the study did not include any outcome relevant to that domain.

The coding results were discussed until agreement was reached between reviewers. This procedure was used to make the classification in the evidence mapping table more transparent and reproducible. Therefore, the strong, moderate, and weak classifications in this review should not be interpreted as effect sizes or statistical strength, but as the degree of conceptual and empirical relevance between each included study and the physical literacy-related domains.

The research analysis process using the PRISMA guidelines is shown in Figure 1.

Based on Figure 1, the article selection process was conducted based on the PRISMA 2020 guidelines. In the identification stage, a search of articles through six databases—Scopus, Web of Science, ERIC, SPORTDiscus, PubMed, and Google Scholar—resulting in 1,061 articles. After all articles were collected, a duplication check was performed using a

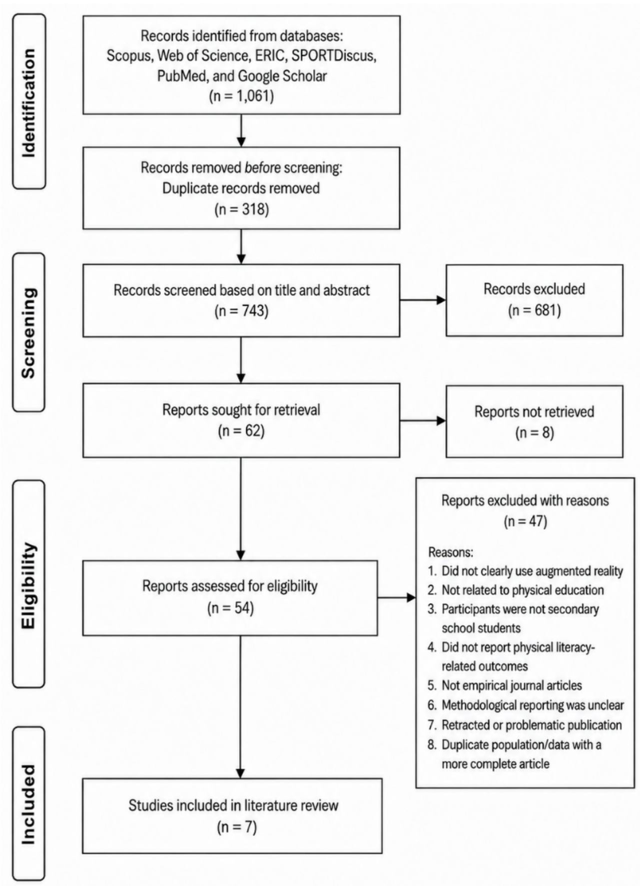


Fig. 1. Research framework based on the PRISMA method

reference management tool. A total of 318 duplicate articles were removed, leaving 743 articles for the screening stage.

In the screening stage, the titles and abstracts of the 743 articles were reviewed for relevance to the research focus, namely the use of augmented reality in physical education at the secondary school level and its relationship to physical literacy or outcomes relevant to physical literacy. From this process, 681 articles were excluded because they did not align with the research topic, did not use augmented reality, were not in the context of physical education, did not involve secondary school students, or did not report outcomes relevant to physical literacy. Therefore, 62 articles proceeded to the full-text search stage.

In the eligibility stage, 62 articles were searched and checked for full-text availability. Of these, 8 articles could not be retrieved in full text, resulting in a complete eligibility assessment of 54 articles. Full-text assessment was conducted based on established inclusion and exclusion criteria. Forty-seven articles were excluded for the following reasons: not explicitly using augmented reality, not in the context of physical education or school-based physical activity, not involving secondary school students, not reporting outcomes related to physical literacy, not being empirical articles, unclear research methods, having publication issues, or sharing data with other, more comprehensive articles.

Based on the entire selection process, seven articles met all inclusion criteria and were included in the final synthesis. These seven articles are Chang et al. (2020), Moreno-Guerrero et al. (2020), Orozco (2022), Kurniawan et al. (2025), Liang et al. (2023), Omarov et al. (2024), and Martínez-López et al. (2022). The identification, screening, eligibility, and inclusion process is presented in the PRISMA flowchart.

Results

Based on the article selection process using the PRISMA 2020 guidelines, seven articles met the inclusion criteria and were analyzed in this systematic literature review. These articles examined the use of augmented reality in secondary school physical education or school-based physical activity contexts and reported outcomes conceptually related to physical literacy. The characteristics of the included studies, forms of AR use, physical literacy-related outcomes, key findings, and synthesis meaning are presented in Table 1.

As shown in Table 1, the included studies varied in design, sample characteristics, AR implementation, learning context, and outcome measures. Most studies used quasi-experimental, comparative, or pretest-posttest designs, while the AR interventions included 3D movement models, mobile AR applications, AR-assisted sport skill learning, AR game-based environments, and outdoor AR activity. Importantly, none of the included studies measured physical literacy as a complete multidimensional construct. Therefore, the findings were interpreted as evidence related to physical literacy-related outcomes rather than direct evidence of overall physical literacy development.

The included studies most frequently reported outcomes related to physical competence, motor skills, learning motivation, knowledge and understanding, engagement, physical activity participation, and fitness. This pattern indicates that AR research in secondary school physical education is still more strongly focused on specific learning or behavioral outcomes than on comprehensive physical literacy assessment. Consequently, the synthesis in this review maps the evidence onto physical literacy-related domains to clarify how each study may contribute conceptually to physical literacy without overstating the findings.

To strengthen methodological transparency, all included studies were assessed for methodological quality and risk of bias. The appraisal focused on the clarity of research questions, appropriateness of collected data, participant selection, outcome measurement, completeness of outcome data, consideration of confounding factors, and clarity of intervention implementation. The results of the methodological quality and risk-of-bias assessment are presented in Table 2.

As shown in Table 2, all included studies had clear research questions and used data that were generally appropriate to address their stated objectives. However, several methodological concerns were identified, particularly related to non-randomized designs, limited reporting of participant allocation, incomplete control of confounding variables, and variation in intervention contexts. Most studies were therefore interpreted as having moderate methodological concerns. One study with a pretest-posttest design without a comparison group was interpreted with higher caution because causal inference was more limited. These appraisal results were not used to exclude studies, but they informed

Table 1. Characteristics of included studies and physical literacy-related outcomes

Title (author and year)	Design and Sample	AR Form	Outcome related to Physical Literacy	Key Findings	Meaning of Synthesis
Applying augmented reality in physical education on motor skills learning (K. E. Chang et al., 2020)	Quasi-experimental; Grade 7 students; comparing AR-assisted instruction with video-assisted instruction.	AR was used to teach movement skills through 3D models combined with books/learning objects. The material included basic running movements and more complex Mark workout movements.	Physical competence, motor skills learning, learning motivation.	AR-assisted instruction is more effective than video-assisted instruction, especially for learning more difficult motor skills.	AR helps students understand movement in a more visual, interactive, and concrete way, contributing to improved movement competency.
Augmented reality as a resource for improving learning in the physical education classroom (Moreno-Guerrero et al., 2020)	Quasi-experimental post-test; 140 high school students in physical education.	AR used the Aurasma application to teach spatial orientation in physical education. The activities were conducted in two learning sessions.	Knowledge-understanding, motivation, autonomy, interaction, problem solving, spatial orientation.	AR-based learning is effective in physical education, particularly for mastering spatially oriented content. The largest difference was found in motivation.	AR not only enhances material comprehension but also strengthens student motivation, learning autonomy, and engagement in Physical Education learning.
Augmented Reality As A Teaching Strategy To Strengthen The Technical Fundamentals Of Volleyball (Orozco et al., 2022)	Comparative study with control and experimental groups; 44 Grade 9 students: 22 control and 22 experimental.	AR was used as a learning strategy for basic volleyball techniques, including overhand passing, underhand passing, serving, blocking, and spiking.	Knowledge-understanding, technical skill knowledge, motivation, sport-specific competence.	The AR group achieved higher post-test scores than the traditional group; the experimental average was 4.41 and the control average was 3.50, with a p-value of 0.001285.	AR strengthens students' understanding of sports technique concepts before they are internalized, thus supporting physical literacy in terms of movement knowledge and sports competency.
Augmented Reality (Ar)-Based Learning on Physical Education Learning Outcomes (Kurniawan et al.,2025)	Quantitative pretest–posttest; 32 junior high school students.	AR-assisted instruction was used in learning basic basketball movements.	Cognitive, affective, psychomotor outcomes, motivation, engagement, conceptual understanding.	There was a significant increase from pre-test to post-test; N-Gain was suggested to be effective. AR provides an immersive learning experience through 3D visualization and increases motivation, engagement, and conceptual understanding.	AR has the potential to enhance three domains of physical education learning outcomes: cognitive, affective, and psychomotor, which align with the concept of physical literacy.
The Effectiveness of Augmented Reality in Physical Sustainable Education on Learning Behaviour and Motivation (Liang et al., 2023)	Quasi-experimental pretest–posttest; two groups: 28 students in the video group and 28 students in the AR group.	AR was developed with Unity 3D and used as a teaching material for movement skills in health and physical education. The control group used video.	Motor skills, learning effect, learning motivation, confidence, satisfaction, learning behaviour.	The AR group demonstrated better motor learning outcomes, motivation, and motor skill performance than the video group.	AR is superior to video because students can manipulate 3D models, view movements from various angles, and learn more independently.
Applying an augmented reality game-based learning environment in physical education classes to enhance sports motivation (Omarov et al., 2024)	Experimental study; two groups of 30 students each; the intervention lasted for one semester.	AR game-based learning environment in physical education/sports classes.	Motivation, physical activity level, engagement, participation.	The AR group demonstrated higher levels of physical activity and motivation than the traditional physical education group.	AR-based games strengthen motivation and active participation, thus supporting the affective and behavioral domains of physical literacy.

Table 1 (continued)

Title (author and year)	Design and Sample	AR Form	Outcome related to Physical Literacy	Key Findings	Meaning of Synthesis
Effects of the augmented reality game Pokémon GO on fitness and fatness in secondary school students (Martínez-López et al., 2022)	Comparative design; 78 students in the Pokémon GO group and 86 students in the control group; students aged 12–15.	Use of the AR Pokémon GO game for 8 weeks in school and family contexts to increase physical activity.	Physical activity, cardiorespiratory fitness, speed/ agility, muscular strength, BMI, body fat.	Pokémon GO players walked a total of 52 km over 8 weeks and played for an average of 40 minutes per day. There was an increase in CRF and a decrease in BMI; inactive students experienced a 22.2% increase in CRF and an 11.3% decrease in body fat.	AR-based outdoor activities can increase physical activity and fitness, but its position as a stronger supporter of the active engagement domain is stronger than formal physical education learning.

Table 2. Methodological quality and risk-of-bias assessment of included studies

Study	Study design	S1	S2	QN1	QN2	QN3	QN4	QN5	Main methodological concerns	Overall appraisal
Chang et al. (2020)	Quasi-experimental	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Non-randomized design; potential confounding was only partly controlled.	Moderate concerns
Moreno-Guerrero et al. (2020)	Quasi-experimental post-test	Yes	Yes	Yes	Yes	Partial	Partial	Yes	Post-test design limits baseline comparison; potential confounding may remain.	Moderate concerns
Orozco et al. (2022)	Comparative study	Yes	Yes	Partial	Yes	Yes	Partial	Yes	Small sample size and limited information on participant allocation.	Moderate concerns
Kurniawan et al. (2025)	Pretest–posttest quantitative study	Yes	Yes	Partial	Yes	Yes	No	Partial	Absence of a comparison group limits causal interpretation.	Higher concerns
Liang et al. (2023)	Quasi-experimental pretest–posttest	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Non-randomized design; some concerns regarding control of confounding variables.	Moderate concerns
Omarov et al. (2024)	Experimental study	Yes	Yes	Partial	Yes	Yes	Partial	Yes	Limited clarity on allocation and control of possible confounders.	Moderate concerns
Martínez-López et al. (2022)	Comparative design	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Outdoor AR activity context may introduce uncontrolled behavioral and environmental factors.	Moderate concerns

Note. S1 = clarity of research questions; S2 = appropriateness of collected data to address the research questions; QN1 = appropriateness of participant selection; QN2 = appropriateness of measurements; QN3 = completeness of outcome data; QN4 = consideration of confounding factors; QN5 = whether the intervention or exposure was implemented as intended. "Yes" indicates that the criterion was sufficiently met, "Partial" indicates that the criterion was partly met or insufficiently reported, and "No" indicates that the criterion was not met. The overall appraisal indicates the level of caution required when interpreting each study and should not be treated as a numerical quality score.

the interpretation of the synthesis and the level of caution applied when drawing conclusions.

The extracted outcomes were then mapped onto four physical literacy-related domains: physical competence, motivation and confidence, knowledge and understanding, and engagement or physical activity participation. Because the included studies did not all directly measure physical literacy, this mapping was used to identify the degree of conceptual and empirical relevance between the outcomes reported in each study and the physical literacy-related domains. The evidence mapping is presented in Table 3.

Table 3 shows that the strongest evidence was concentrated in three domains: physical competence, motivation and confidence, and knowledge and understanding. In the physical competence domain, Chang et al. (2020), Kurniawan et al. (2025), and Liang et al. (2023) provided the strongest evidence because they directly reported motor skill, psychomotor, or movement learning outcomes. These studies suggest that AR may support movement learning by making movement sequences, body positions, and technical information more visible and easier to understand.

In the motivation and confidence domain, Moreno-Guerrero et al. (2020), Kurniawan et al. (2025), Liang et al. (2023), and Omarov et al. (2024) provided strong evidence. These studies reported outcomes such as motivation, confidence, satisfaction, autonomy, and learning behavior. The findings suggest that AR may enhance affective engagement because it provides interactive, novel, and learner-centered experiences that differ from traditional instruction or video-based learning.

Table 3. Evidence mapping of included studies across physical literacy-related domains

Artikel	Physical Competence	Motivation & Confidence	Knowledge & Understanding	Engagement / Physical Activity	Posisi dalam Review
Applying augmented reality in physical education on motor skills learning (K. E. Chang et al., 2020)	Strong	Moderate	Moderate	Moderate	A core article explaining the impact of AR on motor skill learning.
Augmented reality as a resource for improving learning in the physical education classroom (Moreno-Guerrero et al., 2020)	Moderate	Strong	Strong	Moderate	A core article explaining the impact of AR on motivation, autonomy, and spatial orientation understanding.
Augmented Reality As A Teaching Strategy To Strengthen The Technical Fundamentals of Volleyball (Orozco et al., 2022)	Moderate	Moderate	Strong	Moderate	A key article explaining AR in learning specific sport techniques, namely volleyball.
Augmented Reality (Ar)-Based Learning on Physical Education Learning Outcomes (Kurniawan et al.,2025)	Strong	Strong	Strong	Moderate	A key article covering the cognitive, affective, and psychomotor domains.
The Effectiveness of Augmented Reality in Physical Sustainable Education on Learning Behaviour and Motivation (Liang et al., 2023)	Strong	Strong	Medium	Moderate	A core article comparing the effectiveness of AR with video-based learning.
Applying an augmented reality game-based learning environment in physical education classes to enhance sports motivation (Omarov et al., 2024)	Moderate	Strong	Moderate	Strong	A core article explaining AR game-based learning on motivation and physical activity.
Effects of the augmented reality game Pokémon GO on fitness and fatness in secondary school students (Martínez-López et al., 2022)	Moderate	Moderate	Weak	Strong	A supporting article explaining the contribution of AR to students' physical activity and fitness.

Note. Strong = the study directly measured an outcome aligned with the domain and reported clear positive findings. Moderate = the study measured an outcome related to the domain but provided limited, partial, or indirect evidence. Weak = the study only indirectly addressed the domain or provided limited evidence. Not reported = the study did not report outcomes relevant to the domain. These classifications indicate the degree of conceptual and empirical relevance to physical literacy-related domains and should not be interpreted as statistical effect sizes.

In the knowledge and understanding domain, Moreno-Guerrero et al. (2020), Orozco et al. (2022), and Kurniawan et al. (2025) showed strong relevance. These studies indicate that AR can support conceptual understanding in physical education, including spatial orientation, sport technique knowledge, and cognitive learning outcomes. This evidence suggests that AR may function as a bridge between conceptual explanation and movement practice by making abstract or technical content more concrete.

In the engagement or physical activity domain, Omarov et al. (2024) and Martínez-López et al. (2022) provided the strongest evidence. Omarov et al. (2024) showed that AR game-based learning was associated with higher motivation and physical activity levels in physical education classes. Martínez-López et al. (2022) showed that Pokémon GO encouraged physical activity and improved selected fitness-related outcomes. However, this evidence should be interpreted carefully because AR-based outdoor activity does not always represent formal physical education learning with explicit pedagogical objectives.

Overall, the results indicate that AR may support several outcomes conceptually aligned with physical literacy, especially when AR is used to visualize movement, increase motivation, strengthen conceptual understanding, and promote active engagement. Nevertheless, the evidence remains limited because the included studies mostly measured domain-specific outcomes rather than physical literacy as a complete multidimensional construct. Therefore, the findings should be interpreted as evidence of AR's contribution to physical literacy-related outcomes, not as direct evidence that AR improves physical literacy as a whole.

Discussion

The results of this literature review indicate that the use of augmented reality in physical education at the secondary school level has the potential to positively contribute to students' physical literacy-related outcomes. However, the findings should not be interpreted as direct evidence that AR improves physical literacy as a complete multidimensional construct, because most included studies measured specific outcomes related to physical literacy rather than physical literacy itself. This contribution is seen in four main domains: physical competence, motivation and confidence, knowledge and understanding, and engagement or physical activity. These findings suggest that augmented reality functions not only as a technological medium but also as a pedagogical approach that can enrich the movement learning experience. Through visualization, interaction, embodied learning, and learner-centered exploration, AR may create conditions that support outcomes conceptually aligned with physical literacy. These results align with the development of research on immersive technologies in physical education, which shows that AR, VR, and MR are beginning to be used to support student learning, motivation, engagement, and physical activity (Pérez-Muñoz et al., 2024; Jantanukul, 2024).

In the realm of physical competence, review results indicate that AR can help students understand and master movement skills more effectively. Studies by Chang et al. (2020), Liang et al. (2023), Kurniawan et al. (2025), and Orozco (2022) show that the use of AR can support learning in motor skills, sports skills, and psychomotor domains. The advantage of AR lies in its ability to display objects or movement models visually, interactively, and observably from various angles. This mechanism may help students observe movement sequences, recognize body positions, compare correct and incorrect techniques, and internalize technical information before engaging in direct movement practice. This helps students understand movement sequences, body positions, and sports techniques before practicing them directly (Antonopoulos, 2024). Therefore, AR can be a supporting media that is relevant to physical education materials that require movement control, coordination, and mastery of motor skills. Nevertheless, the evidence should be interpreted cautiously because several studies used non-randomized or quasi-experimental designs, meaning that improvements in motor-related outcomes cannot be attributed solely to AR without considering the role of teacher instruction, practice time, and initial student ability.

In the motivation and self-confidence domain, review results indicate that AR can increase student motivation, learning interest, self-confidence, and satisfaction in physical education learning. Findings by Moreno-Guerrero et al. (2020), Liang et al. (2023), Omarov et al. (2024), and Kurniawan et al. (2025) indicate that students who learn using AR tend to have better affective engagement compared to traditional or video-based learning. This may occur because AR provides a more novel, interactive, and enjoyable learning experience. From a pedagogical perspective, AR may influence motivation through several mechanisms, including novelty effects, autonomy in exploring learning materials, immediate visual interaction, game-based challenges, and opportunities to learn at an individual pace. In the context of physical literacy, motivation and self-confidence are important aspects because students not only need to be able to move but also need to have an internal drive to engage in physical activity continuously (Martín-Rodríguez & Madrigal-Cerezo, 2025; Prabowo et al., 2025). Therefore, AR can be a relevant strategy to overcome low student motivation in learning Physical Education, although its motivational benefits may depend on how well the AR activity is integrated with meaningful physical tasks rather than being used merely as a technological attraction.

In the realm of knowledge and understanding, AR plays a role in strengthening students' understanding of movement concepts, sports techniques, spatial orientation, and the relationship between theoretical knowledge and physical practice. A study by Moreno-Guerrero et al. (2020) showed that AR is effective for learning spatial orientation, while Orozco (2022) showed that AR helps students understand basic volleyball techniques. Kurniawan et al. (2025) also showed that 3D visualization in basketball learning can improve students' conceptual understanding. These findings align with the characteristics of AR, which can make abstract concepts more concrete by combining virtual objects with real environments (H. Y. Chang et al., 2022). In this way, AR may serve as a cognitive bridge between verbal explanation, visual representation, and physical execution, allowing students to connect what they see, understand, and perform in movement learning. However, the use of AR still needs to be designed carefully because visualizations that are too complex or unfocused can distract students from the main learning objectives (Chamola et al., 2025; Albishri & Blackmore, 2025). Thus, AR-based learning should include clear instructional cues, simple visual displays, and teacher-guided reflection to prevent unnecessary cognitive load.

In the domain of engagement and physical activity, AR has the potential to increase student engagement in physical activity, especially when developed in the form of game-based learning or outdoor activities. A study by Omarov et al. (2024) showed that AR game-based learning environments can increase students' motivation and physical activity levels compared to traditional physical education. Meanwhile, Martínez-López et al. (2022) showed that using Pokémon GO can encourage students to be more physically active, improve cardiorespiratory fitness, and reduce certain body composition indicators. These findings are important because one of the challenges of physical education in secondary schools is how to keep students active, engaged, and enjoying physical activity (Cimenti et al., 2025; Prabowo et al., 2025). However, the use of outdoor game-based AR needs to be distinguished from formal physical education learning, because not all AR activities automatically have a structured pedagogical goal. Therefore, evidence from outdoor AR games should be interpreted as supporting engagement or physical activity participation rather than as direct evidence of formal physical education learning or complete physical literacy development.

Importantly, the results of this review demonstrate that the effectiveness of AR in physical education is determined not only by technological sophistication but also by the quality of pedagogical design. AR is most effective when used to support clear learning objectives, appropriate movement materials, targeted teacher instruction, and adequate opportunities for physical practice (Ariffin et al., 2022; Antonopoulos, 2024; Ridwan et al., 2025). On the other hand, AR can be less effective if it is only used as an additional medium without strong integration with motor learning (Yudiernawati et al., 2025). The methodological quality assessment also indicates that most included studies had moderate methodological concerns, particularly related to non-randomized designs, limited reporting of participant allocation, and incomplete control of confounding variables. One study with a pretest-posttest design without a comparison group required higher caution because causal inference was more limited. Previous literature also shows that technology integration in education can address challenges such as device access gaps, limited teacher training, digital content quality, cognitive load, and school readiness (Martín-Rodríguez & Madrigal-Cerezo, 2025). Therefore, AR should not be positioned as a substitute for teachers or a substitute for movement practice, but rather as a pedagogical medium to strengthen physical education learning.

The implications of this literature review indicate that physical education teachers can utilize AR to clarify reflection techniques, strengthen understanding of movement concepts, increase motivation, and encourage student engagement in physical activity. For schools, the use of AR requires device support, networking, teacher training, and integration with the physical education curriculum. For learning technology developers, the results of this review emphasize the importance of designing AR applications that are simple, easy to use, appropriate to the characteristics of secondary school students, and support real-life physical activity. For researchers, these findings open up opportunities to develop AR interventions that explicitly measure physical literacy multidimensionally, not just motor learning outcomes or motivation. Thus, AR can be developed as a learning medium that is not only technologically engaging but also relevant to the goals of modern physical education.

This study has several limitations. First, only seven articles met the inclusion criteria, so generalization of the findings should be done with caution. Second, most articles did not measure physical literacy as a whole, but instead used outcomes related to physical literacy, such as motor competence, motivation, knowledge, engagement, physical activity, and fitness. Third, study designs, types of AR, intervention duration, physical education materials, and measurement instruments varied significantly across articles, necessitating a meta-analysis. Fourth, some articles focused more on school-based or outdoor physical activity than on formal learning. Future research is expected to develop more structured and long-term augmented reality interventions that explicitly measure students' physical literacy across multiple dimensions.

Conclusion

This review indicates that augmented reality (AR) has the potential to support outcomes conceptually related to physical literacy in secondary school physical education. Across the seven included studies, AR was associated with improvements in physical competence, motivation and confidence, knowledge and understanding, and engagement or physical activity participation. However, the evidence should be interpreted cautiously because most studies did not directly measure physical literacy as a complete multidimensional construct. Instead, they reported domain-specific outcomes such as motor skills, conceptual understanding, motivation, physical activity, and fitness. The findings suggest that AR may be useful when it is integrated with clear pedagogical objectives, teacher guidance, appropriate movement tasks, and opportunities for real physical practice. Therefore, AR should be positioned as a supportive pedagogical tool rather than a substitute for teachers or direct movement experiences. Future studies should use stronger designs and multidimensional physical literacy measures.

Ethics Approval

Ethical approval was not required because this study was a literature review based exclusively on previously published studies and did not involve the collection of primary data from human participants.

Informed Consent

Informed consent was not applicable because this study did not involve the recruitment of human participants or the collection of primary data.

Data Availability Statement

No new primary data were generated or collected in this study. All data analyzed in this review were derived from the published studies included in the literature review and are reported within the article.

AI Transparency Statement

Generative artificial intelligence tools were used solely for language editing and improving the readability of the manuscript.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Prabowo, T.A., Afifah, M., Cahyo, F.D., Zakaria, A., & Indarto, A.V. (2025). Self-efficacy and motivation student in physical education learning: Scoping review. *Jurnal Porkes, 8*(1), 377–389. https://doi.org/10.29408/porkes.v8i1.29743

Prabowo, T.A., Indarto, A.V., Cahyo, F.D., Afifah, M., & Zakaria, A. (2025). Extracurricular sports as a tool for character development and well-being in secondary school student: Scoping review. *Riyadhoh: Jurnal Pendidikan Olahraga, 8*(1), 72–79. https://doi.org/10.31602/rjpo.v8i1.18623

Prabowo, T.A., Indarto, A.V., Zakaria, A., Cahyo, F.D., & Afifah, M. (2025). The effect of physical fitness on academic achievement through self-confidence in adolescents aged 16–18 years: A scoping review. *Physical Culture, Recreation and Rehabilitation, 4*(1), 50–57. https://doi.org/10.15561/physcult.2025.0105

Rocliffe, P., Adamakis, M., O'Keeffe, B.T., Walsh, L., Bannon, A., Garcia-Gonzalez, L., … MacDonncha, C. (2024). The impact of typical school provision of physical education, physical activity and sports on adolescent mental health and wellbeing: A systematic literature review. *Adolescent Research Review,* 9, 339–364. https://doi.org/10.1007/s40894-023-00220-0

Martínez-López, E.J., López-Serrano, S., De La Torre-Cruz, M., & Ruiz-Ariza, A. (2022). Effects of the augmented reality game Pokémon GO on fitness and fatness in secondary school students. *Health Education Journal, 81*(1), 54–68. https://doi.org/10.1177/00178969211047800

Liu, Y., & Chen, S. (2021). Physical literacy in children and adolescents: Definitions, assessments, and interventions. *European Physical Education Review, 27*(1), 96–112. https://doi.org/10.1177/1356336X20925502

Grauduszus, M., Koch, L., Wessely, S., & Joisten, C. (2024). School-based promotion of physical literacy: A scoping review. *Frontiers in Public Health,* 12, 1322075. https://doi.org/10.3389/fpubh.2024.1322075

Nevitaningrum, N., Syafei, P.M., Mulya, G., Prabowo, T.A., Ramadhanty, O.Y., & Suryani, P.D.I.D. (2026). Healthy lifestyle program to improve physical literacy and physical fitness of junior high school students: A quasi-experimental approach. *Journal of Learning Theory and Methodology, 7*(1), 25–36. https://doi.org/10.17309/jltm.2026.7.1.02

Peranita, N., Semarayasa, I.K., & Muliarta, I.W. (2025). The implementation of physical literacy in physical education learning to support lifelong active lifestyle. *Journal of Physical Education Health and Sport, 12*(2), 645–653. https://doi.org/10.15294/jpehs.v12i2.38015

Alhroub, N., Al-Sarairhe, I., Al Awamleh, R., Ayasreh, I., Alkhawaldeh, A., ALBashtawy, M., … Alkhawaldeh, H. (2024). Physical activity barriers among adolescents in Jordan: A cross-sectional study. *SAGE Open Nursing,* 10, 23779608241272688. https://doi.org/10.1177/23779608241272688

Ferreira Silva, R.M., Mendonça, C.R., Azevedo, V.D., Memon, A.R., Silva Noll, P.R.E., & Noll, M. (2022). Barriers to high school and university students' physical activity: A systematic review. *PLOS ONE, 17*(4), e0265913. https://doi.org/10.1371/journal.pone.0265913

Cimenti, C., Thøgersen-Ntoumani, C., Pfeffer, K., Krustrup, P., Schenk, P.M., & Ntoumanis, N. (2025). Barriers to and enablers of physical education engagement among school students aged 6–16 years: A scoping review. *International Review of Sport and Exercise Psychology.* https://doi.org/10.1080/1750984X.2025.2597932

Terra, L.F., Rezende, L.M.T. de, Ferreira Silva, R.M., Costa, W.P. da, Miñana-Signes, V., Noll, M., & Silva, P.R.E. (2025). Interventions on barriers to the participation of adolescents in physical activity: A systematic review. *International Journal of Environmental Research and Public Health, 22*(6), 881. https://doi.org/10.3390/ijerph22060881

Chang, H.Y., Binali, T., Liang, J.C., Chiou, G.L., Cheng, K.H., Lee, S.W.Y., & Tsai, C.C. (2022). Ten years of augmented reality in education: A meta-analysis of (quasi-) experimental studies to investigate the impact. *Computers & Education,* 191, 104641. https://doi.org/10.1016/j.compedu.2022.104641

Yefremenko, A., Kolokolov, V., Pozdnyakova, M., & Piatysotska, S. (2025). Virtual and augmented reality gaming technologies in physical education: A review of efficiency. *Sports Games, 1*(35), 5–15. https://doi.org/10.15391/si.2025-1.01

Topalska, R., & Rumenov, Y. (2025). Facilitating the visualisation of moving games by implementing augmented reality in primary physical education classes. *TEM Journal, 14*(1), 816–822. https://doi.org/10.18421/TEM141-72

Chang, K.E., Zhang, J., Huang, Y.S., Liu, T.C., & Sung, Y.T. (2020). Applying augmented reality in physical education on motor skills learning. *Interactive Learning Environments, 28*(6), 685–697. https://doi.org/10.1080/10494820.2019.1636073

Pérez-Muñoz, S., Castaño Calle, R., Morales Campo, P.T., & Rodríguez-Cayetano, A. (2024). A systematic review of the use and effect of virtual reality, augmented reality and mixed reality in physical education. *Information, 15*(9), 582. https://doi.org/10.3390/info15090582

Omarov, N., Omarov, B., Azhibekova, Z., & Omarov, B. (2024). Applying an augmented reality game-based learning environment in physical education classes to enhance sports motivation. *Retos,* 60, 269–278. https://doi.org/10.47197/retos.v60.109170

Ariffin, U.H.B., Mohamed Mokmin, N.A.B., & Mohd Hamizi, M.A.A. (2022). Augmented reality technology in physical education: A systematic review in instructional design, and AR implementation option over the last 5 years. *Advanced Journal of Technical and Vocational Education, 6*(1), 13–20. https://doi.org/10.26666/rmp.ajtve.2022.1.3

Ridha, S., Mashud, Warni, H., & Arifin, S. (2025). Physical literacy in Indonesian school physical education: A systematic review of concepts, assessment, and implementation. *Edu Sportivo: Indonesian Journal of Physical Education, 6*(3), 285–301. https://doi.org/10.25299/esijope.2025.vol6(3).17622

Vieira, D., Silva Dias, T., Dudley, D., & Batista, P. (2025). Exploring physical literacy in school contexts: A systematic review of qualitative evidence. *Frontiers in Sports and Active Living,* 7, 1713780. https://doi.org/10.3389/fspor.2025.1713780

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ,* 372, n71. https://doi.org/10.1136/bmj.n71

Kurniawan, F., Hidayatullah, M.F., Kristiyanto, A., Riyadi, S., Ekawati, F.F., Ningrum, N.R., & Prabowo, T.A. (2024). Physical literacy needs in esports: Literature review. *Retos,* 58, 495–505. https://doi.org/10.47197/retos.v58.107885

Liao, X., Hu, R., Li, B., Hu, J., Chu, H., Yin, X., Chen, W., & Zhang, Y. (2021). An introduction to Mixed Methods Appraisal Tool. *Chinese General Practice, 24*(31), 4015–4020. https://doi.org/10.12114/j.issn.1007-9572.2021.00.215

Moreno-Guerrero, A.J., García, S.A., Navas-Parejo, M.R., Campos-Soto, M.N., & García, G.G. (2020). Augmented reality as a resource for improving learning in the physical education classroom. *International Journal of Environmental Research and Public Health, 17*(10), 3637. https://doi.org/10.3390/ijerph17103637

Jiménez Orozco, H.L., Álvarez Araque, W.O., & Sarmiento Fautoque, D.Y. (2022). Augmented reality as a teaching strategy to strengthen the technical fundamentals of volleyball. *Journal of Language and Linguistic Studies, 18*(4), 942–967.

Kurniawan, R., Hakim, S.M., Rangkuti, Y.A., & Pranata, D.Y. (2025). Augmented reality (AR)-based learning on physical education learning outcomes. *Visipena, 16*(2), 345–356. https://doi.org/10.46244/visipena.v16i2.3382

Liang, L., Zhang, Z., & Guo, J. (2023). The effectiveness of augmented reality in physical sustainable education on learning behaviour and motivation. *Sustainability, 15*(6), 5062. https://doi.org/10.3390/su15065062

Jantanukul, W. (2024). Immersive reality in education: Transforming teaching and learning through AR, VR, and mixed reality technologies. *Journal of Education and Learning Reviews, 1*(2), 51–62. https://doi.org/10.60027/jelr.2024.750

Antonopoulos, N. (2024). The effect of augmented reality (AR) on improving motor learning skills in students. *IOSR Journal of Sports and Physical Education, 11*(6), 35–45. https://doi.org/10.9790/6737-1106013545

Martín-Rodríguez, A., & Madrigal-Cerezo, R. (2025). Technology-enhanced pedagogy in physical education: Bridging engagement, learning, and lifelong activity. *Education Sciences, 15*(4), 409. https://doi.org/10.3390/educsci15040409

Chamola, V., Peelam, M.S., Mittal, U., Hassija, V., Singh, A., Pareek, R., … Brown, D.J. (2025). Metaverse for education: Developments, challenges, and future direction. *Computer Applications in Engineering Education,* 33, e70018. https://doi.org/10.1002/cae.70018

Albishri, B., & Blackmore, K.L. (2025). Duality in barriers and enablers of augmented reality adoption in education: A systematic review of reviews. *Interactive Technology and Smart Education, 22*(2), 167–191. https://doi.org/10.1108/ITSE-10-2023-0194

Ridwan, M., Aswanda, N., Sumarno, S., Ahmedov, F., & Pranoto, A. (2025). Augmented reality's impact on learning motivation in physical education: A systematic review. *International Journal of Body, Mind and Culture, 12*(2), 5–12. https://doi.org/10.61838/ijbmc.v12i2.838

Yudiernawati, A., Wahyuni, T.D., & Suryani, P. (2025). Learning innovation using augmented reality for interactive learning experiences. *Jurnal Konseling dan Pendidikan, 13*(1), 305–315. https://doi.org/10.29210/1142300

Доповнена реальність у фізичному вихованні та фізичній грамотності учнів середньої школи: огляд літератури

Тріснар Аді Прабово[1ABCDE], Маратул Афіфа[1ABD], Асьїдіка Віто Індарто[1CD], Закван Юмна Духа Пахлаві[1BCD], Фебріансьях Дві Чахйо[1ABD]

[1]Університет Мухаммадія Бребес

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 13 с., 3 табл., 1 рис., 36 джерел.

Мета. Незважаючи на зростання використання технологій доповненої реальності (AR) у фізичному вихованні, результати емпіричних досліджень щодо їх застосування серед учнів середньої школи є неоднорідними за дизайном досліджень, освітнім контекстом і досліджуваними показниками. Метою дослідження було систематизувати та узагальнити емпіричні дані щодо застосування AR у фізичному вихованні учнів середньої школи та визначити її зв'язок із показниками, що характеризують основні компоненти фізичної грамотності.

Матеріали та методи. У дослідженні використано дизайн огляду літератури з урахуванням рекомендацій PRISMA 2020. Пошук статей проводили в шести базах даних: Scopus, Web of Science, ERIC, SPORTDiscus, PubMed та Google Scholar. Пошук було обмежено рецензованими журнальними статтями, опублікованими в період з 2015 по 2025 рік. Використані ключові слова поєднували терміни, пов'язані з доповненою реальністю, фізичним вихованням, середньою школою та результатами, пов'язаними з фізичною грамотністю. Загалом було виявлено 1061 статтю. Після видалення дублікатів, скринінгу назв і анотацій, оцінювання повних текстів та перевірки відповідності критеріям включення до огляду було включено 7 статей. Дані аналізували за допомогою наративного та тематичного синтезу на основі чотирьох доменів фізичної грамотності: фізична компетентність, мотивація та впевненість, знання та розуміння, а також залученість до фізичної активності або участь у ній. Методологічну якість і ризик упередженості оцінювали за допомогою Mixed Methods Appraisal Tool (MMAT), а силу доказів визначали з використанням чітких критеріїв кодування.

Результати. Результати свідчать, що AR може сприяти покращенню низки показників, концептуально пов'язаних із фізичною грамотністю, у фізичному вихованні учнів середньої школи. Навчання на основі AR може покращувати рухові навички, специфічні для окремих видів спорту компетентності, мотивацію, упевненість у собі, розуміння концепцій руху, залученість до навчання та участь у фізичній активності. Найпереконливіші докази було виявлено в доменах фізичної компетентності, мотивації та впевненості в собі, а також знань і розуміння. Крім того, ігрове навчання та заняття на відкритому повітрі на основі AR також сприяли підвищенню залученості учнів і рівня фізичної активності. Однак у жодному з включених досліджень фізичну грамотність не вимірювали як цілісний багатовимірний конструкт, тому отримані результати слід інтерпретувати як докази щодо показників, пов'язаних із фізичною грамотністю, а не як прямі докази розвитку фізичної грамотності загалом.

Висновки. AR має потенціал як допоміжний педагогічний засіб для сприяння результатам, пов'язаним із фізичною грамотністю, у фізичному вихованні учнів середньої школи. Однак її ефективність значною мірою залежить від дизайну навчання, підготовленості вчителя, навчальних цілей та інтеграції AR із реальною руховою активністю. У майбутніх дослідженнях доцільно розробляти більш структуровані й довготривалі втручання із застосуванням доповненої реальності, які передбачатимуть безпосереднє багатовимірне оцінювання фізичної грамотності учнів.

Ключові слова: доповнена реальність, фізичне виховання, фізична грамотність, середня школа, рухова компетентність.

Information about the Authors:

Prabowo, Trisnar Adi: trisnar.prabowo@umbs.ac.id; https://orcid.org/0000-0001-6977-0503; Study Program of Physical Education, Jl. Pangeran Diponegoro Grengseng No.184, Grengseng, Taraban, Kec. Paguyangan, Kabupaten Brebes, Jawa Tengah 52276, Indonesia

Afifah, Mar'atul: maratul.afifah@umbs.ac.id; https://orcid.org/0009-0004-2849-6543; Study Program of Physical Education, Jl. Pangeran Diponegoro Grengseng No.184, Grengseng, Taraban, Kec. Paguyangan, Kabupaten Brebes, Jawa Tengah 52276, Indonesia.

Indarto, Asyidika Vito: asyidika.vito@umbs.ac.id; https://orcid.org/0009-0005-3210-8379; Study Program of Physical Education, Jl. Pangeran Diponegoro Grengseng No.184, Grengseng, Taraban, Kec. Paguyangan, Kabupaten Brebes, Jawa Tengah 52276, Indonesia.

Pahlavi, Zakwan Yumna Dhuha: zakwanyumna1@gmail.com; https://orcid.org/0009-0008-2127-754X; Study Program of Physical Education, Jl. Pangeran Diponegoro Grengseng No.184, Grengseng, Taraban, Kec. Paguyangan, Kabupaten Brebes, Jawa Tengah 52276, Indonesia.

Cahyo, Febriansyah Dwi: riancahyoqwerty@gmail.com; https://orcid.org/0009-0006-6845-9602; Study Program of Physical Education, Jl. Pangeran Diponegoro Grengseng No.184, Grengseng, Taraban, Kec. Paguyangan, Kabupaten Brebes, Jawa Tengah 52276, Indonesia.

Cite this article as: Prabowo, T. A., Afifah, M., Indarto, A. V., Pahlavi, Z. Y. D., & Cahyo, F. D. (2026). Augmented Reality in Secondary School Physical Education and Physical Literacy: Literature Review. *Journal of Learning Theory and Methodology, 7*(2), 91-103. https://doi.org/10.17309/jltm.2026.7.2.04

Received: 18.06.2026. Accepted: 21.07.2026. Published: 15.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання
ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Original Scientific Article

Development and Validation of ComicSIM on Volcanic Activity With a PISA-Aligned Assessment Tool

Sheila Barajan[1ABCD] and Erdee Cajurao[2ACD]

[1]Buensuerte Integrated School
[2]College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.05

Abstract

Objectives. Comic-based instructional materials demonstrate measurable gains in student engagement and scientific literacy; yet few studies integrate comic formats with Strategic Intervention Materials (SIMs) and PISA-aligned assessment for Earth Science in Philippine secondary schools. This study developed and validated ComicSIM, a comic-based SIM on volcanic activity for Grade 9 Science, alongside a PISA-aligned assessment tool comprising 20 items across eight lessons.

Materials and Methods. A descriptive-developmental design was used. Five expert jurors holding graduate qualifications in science education or curriculum development, with at least five years of professional experience, independently evaluated both instruments. ComicSIM was assessed across seven sub-criteria using a 5-point rubric; inter-rater agreement was measured using Fleiss' Kappa. The PISA-aligned assessment tool was evaluated using a 15-item alignment instrument with a 4-point scale; internal consistency was assessed using Cronbach's Alpha.

Results. ComicSIM achieved a weighted mean of 4.5 (Excellent) and a Fleiss' Kappa of 0.70 (substantial agreement). The PISA-aligned assessment tool obtained an alignment mean of 3.8 (Strongly Agree), a dimension mean of 4.93 (Excellent), and a Kappa of 0.87 (almost perfect agreement). Cronbach's Alpha of 0.88 confirmed strong internal consistency.

Conclusions. Both instruments are expert-validated, reliable, and suitable for use in advancing PISA-aligned scientific literacy instruction in Philippine secondary science classrooms, pending confirmation of classroom effectiveness through empirical research.

Keywords: comicSIM, strategic intervention material, PISA, volcanic activity, Grade 9 science.

Introduction

The Philippines occupies a geographically hazardous position at the convergence of three major tectonic plates, placing it firmly within the Pacific Ring of Fire. Volcanic eruptions constitute one of the most severe natural hazard threats faced by Filipino communities, making scientific understanding of volcanic processes both academically and practically consequential (Aghaei et al., 2018; German et al., 2022; Kurata et al., 2022). Yet despite this pressing local relevance, Filipino students continue to underperform on international science assessments. In the 2018 PISA cycle, the Philippines scored 357 in Science, placing it well below the OECD average and underscoring systemic gaps in science learning outcomes at the basic education level (DepEd, 2018). The need for innovative instructional materials that simultaneously build scientific literacy and sustain student engagement has consequently become a recognised priority in Philippine science education research (Cahyana et al., 2019; Degorio & Langub, 2025).

These outcomes have prompted educators and researchers to explore more effective and engaging instructional approaches. Strategic Intervention Materials (SIMs), defined as structured, teacher-crafted resources designed to remediate least-mastered learning competencies, have become widely adopted in Philippine classrooms as targeted tools for improving student performance (Dacumos, 2016; Gabucan & Sanchez, 2021). Meta-analytic evidence confirms that both paper-based and digital SIMs demonstrate large effect sizes in improving science and mathematics outcomes among Filipino learners (Borabo & Dio, 2025). However, most existing SIMs rely on conventional text-based formats that may not adequately stimulate student engagement or leverage the motivational potential of visual media.

Comics represent an underutilized educational medium in the Philippine context, despite substantial evidence of

JLTM
LLC OVS
Corresponding Author: Erdee Cajurao,
e-mail: eccajurao@debesmscat.edu.ph

their effectiveness in science instruction internationally. Integrating narrative with sequential visual imagery, comics create a dual-coding effect that deepens conceptual understanding and sustains motivation, consistent with Mayer's (2009) Cognitive Theory of Multimedia Learning and with constructivist perspectives on active meaning-making through contextualized narrative (Bruner, 1990; Vygotsky, 1978). Research consistently demonstrates that comic-based science materials improve student engagement, scientific vocabulary acquisition, and conceptual comprehension (Fitria et al., 2023; Hosler & Boomer, 2011; Tatalovic, 2009). In the Philippine setting specifically, comic-based instructional materials combining SIM structures with comic formats have shown positive learning attitudes and significant gains in science achievement (Molina & Cadorna, 2025; Samosa, 2021).

Simultaneously, growing attention to PISA-aligned assessments reflects a broader imperative to equip Filipino students with the scientific literacy competencies demanded by the 21st century. The PISA 2022 science framework assesses students' capacity to explain phenomena scientifically, evaluate and design scientific inquiry, and interpret data and evidence in real-world contexts (OECD, 2023). Integrating these competencies into locally developed instructional materials represents a powerful mechanism for bridging the gap between national curriculum delivery and international benchmarks (Valio et al., 2025).

Comics, defined as a medium combining sequential art with narrative text, have gained increasing recognition as legitimate instructional tools in science education (Eisner, 2008). A systematic review of 20 studies found that comics significantly enhance student motivation, engagement, and conceptual understanding while promoting active participation and fostering critical thinking skills (Rasmet et al., 2025). This position is theoretically grounded in Mayer's (2009) Cognitive Theory of Multimedia Learning, which posits that the simultaneous processing of words and pictures through dual-coding enhances learning outcomes by reducing cognitive load and fostering deeper encoding. Constructivist theory similarly supports comic-based instruction: Bruner (1990) argued that narrative is a fundamental mode through which learners construct meaning, while Vygotsky (1978) emphasized the role of culturally situated tools in scaffolding conceptual development. Comics that embed scientific concepts within culturally resonant storylines align directly with these principles.

Empirical support for comic-based science instruction is well established. Hosler and Boomer (2011) found that students learning from comics outperformed those using conventional text resources in measures of conceptual understanding. Tatalovic (2009) observed that science comics facilitate science literacy by rendering technical vocabulary accessible through narrative contexts. More recent evidence from Southeast Asia corroborates these findings: Fitria et al. (2023) demonstrated that digital comic teaching materials significantly improved scientific literacy scores among elementary students in Indonesia, while Fananta et al. (2019) found that inquiry-based science comics enhanced primary learners' scientific literacy across three dimensions. Constructing socio-scientific issue literacy through comics has further been shown to foster critical thinking and equip students to engage with complex real-world challenges (Rosidah et al., 2022). Valente and Faria (2026) extend this argument to pre-service teacher education, finding that future elementary teachers who engaged with science comics developed more positive science identity attitudes, an outcome with downstream implications for classroom instruction.

In the Philippine context, comic-based SIMs have yielded particularly promising results. Samosa (2021) reported high positive attitudes and significant pre-test to post-test gains among students exposed to CoSIM, a comics-plus-SIM hybrid, in Biology. Molina and Cadorna (2025) demonstrated that comic-based learning materials enhanced understanding, attitude, and skills in Physics among Filipino high school students. These locally grounded findings affirm that comics, when integrated with SIM structures, constitute a culturally resonant and pedagogically effective modality for Philippine science instruction.

Strategic Intervention Materials are teacher-made instructional tools specifically designed to address identified learning gaps in students who have not mastered targeted competencies. Within the Philippine Department of Education framework, SIMs function as supplementary resources structured around a guide card, activity cards, assessment cards, and an enrichment card, providing scaffolded and self-directed remediation (Dacumos, 2016). A 2025 meta-analysis of 15 studies on SIM effectiveness in science and mathematics confirmed that SIMs produce large effect sizes regardless of format, with paper-based SIMs yielding an effect size of 1.67 and digital SIMs yielding 1.34, establishing their value as remedial tools across the K-12 spectrum (Borabo & Dio, 2025).

Development and validation studies of SIMs in Philippine science consistently report expert-validated materials as highly acceptable in content and technical aspects (Lituanas & Dela Cruz, 2024; Samosa, 2021). In instructional material development research, expert validation by qualified jurors constitutes the accepted and methodologically defensible first phase of the development cycle, establishing content validity and design quality before proceeding to learner-facing trials (Branch & Varank, 2009; Gall et al., 1996). This phased approach reflects the understanding that materials with unverified content accuracy or design integrity should not be deployed in classrooms, making expert validation an essential prerequisite rather than a substitute for subsequent effectiveness studies. However, most existing SIMs do not explicitly align their assessment components with international literacy frameworks such as PISA, limiting their utility for preparing students for globally benchmarked competencies.

Earth science topics such as volcanic activity present particular instructional challenges because they require students to reason across multiple spatial and temporal scales, connect abstract geological processes to observable phenomena, and apply scientific evidence to real-world risk decisions (Francek, 2013). Research on geoscience education consistently finds that students hold persistent misconceptions about volcanic processes, including the belief that eruptions are random or unpredictable, and that human activity plays no role in volcanic risk exposure (King, 2010). Narrative-based and contextualized instructional approaches have been shown to be especially effective at addressing these misconceptions by grounding abstract processes in relatable human consequences (Demirci & Okur, 2021; Korur et al., 2025). For Filipino students in particular, volcanic activity is not an abstract academic topic but a lived environmental reality, making contextualized instructional materials such as ComicSIM especially well-suited to this content domain.

The PISA 2022 science framework defines scientific literacy as the capacity to engage with science-related issues and use scientific knowledge and reasoning to explain phenomena, evaluate evidence, and make informed decisions in real-world contexts (OECD, 2023). The Philippines' 2018 PISA result of 357 in Science placed it among the lowest-scoring participants, reflecting systemic challenges in developing the inquiry-based, evidence-evaluating competencies that PISA measures (DepEd, 2018). Valio et al. (2025) argued that integrating PISA-aligned assessments into local instructional materials offers a direct mechanism for building these competencies within the K-12 curriculum. Instructional materials that contextualize PISA competencies within locally relevant phenomena, such as volcanic activity in the Philippine archipelago, address a critical pedagogical need by linking scientific reasoning to students' lived experience. The PISA 2022 framework identifies real-world contexts, including geohazard risk and environmental sustainability, as primary assessment domains, making volcanic activity an ecologically valid and curriculum-aligned content focus (OECD, 2023).

Despite the documented benefits of both SIMs and comics in Philippine science education, no study to date has systematically combined the two within a PISA-aligned framework for Earth science content. This study addresses that gap by developing and validating ComicSIM, a comic-based SIM focused on volcanic activity for Grade 9 Science, along with an accompanying PISA-aligned assessment tool. The study is guided by three research questions: (1) How do expert jurors assess ComicSIM across seven sub-criteria of instructional material quality? (2) What PISA-aligned assessment tool can be developed from the content of ComicSIM? (3) How reliable and valid are ComicSIM and the PISA-aligned assessment tool as evaluated through expert judgment and statistical analysis?

Materials and Methods

Study Participants

Five expert jurors validated both ComicSIM and the PISA-aligned assessment tool. Jurors were identified through purposive sampling based on three criteria: (a) a graduate degree (master's or doctoral level) in science education, curriculum development, or a closely related field; (b) a minimum of five years of professional experience in secondary science teaching, instructional material development, or educational assessment; and (c) familiarity with the Philippine K-12 Science curriculum and DepEd instructional material standards. To manage potential institutional bias, jurors were drawn from schools and institutions external to the researchers' institution. No juror had prior access to the ComicSIM drafts before the validation process. Table 1 presents a summary of juror profiles.

Table 1. Summary of Expert Juror Profiles

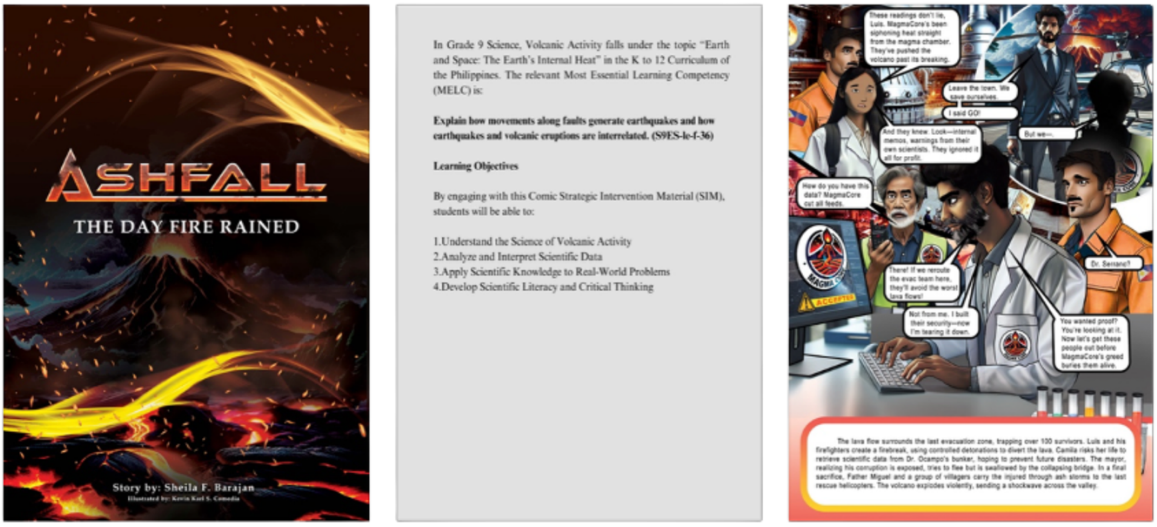
Juror	Highest Qualification	Area of Expertise	Years of Experience
J1	Doctor of Education	Science Education and Curriculum	15
J2	Master of Arts in Education	Instructional Material Development	10
J3	Doctor of Philosophy	Educational Assessment and Measurement	18
J4	Master of Arts in Teaching	Secondary Science Teaching (Earth Science)	8
J5	Doctor of Education	Science Pedagogy and Curriculum Evaluation	12

Study Organization

This study employed a descriptive-developmental research design. The developmental phase involved the systematic creation of ComicSIM, a comic-based Strategic Intervention Material on volcanic activity for Grade 9 Science, together with a 20-item PISA-aligned assessment tool structured across eight thematic lessons. The descriptive phase involved expert evaluation of both instruments through juror validation. Consistent with established practice in instructional material development research, expert validation constitutes the appropriate first phase of the development cycle, generating evidence of content validity and design quality prior to classroom deployment (Branch & Varank, 2009; Gall et al., 1996). A classroom-based effectiveness study involving student participants is identified as the necessary and immediate next phase of this programme of research.

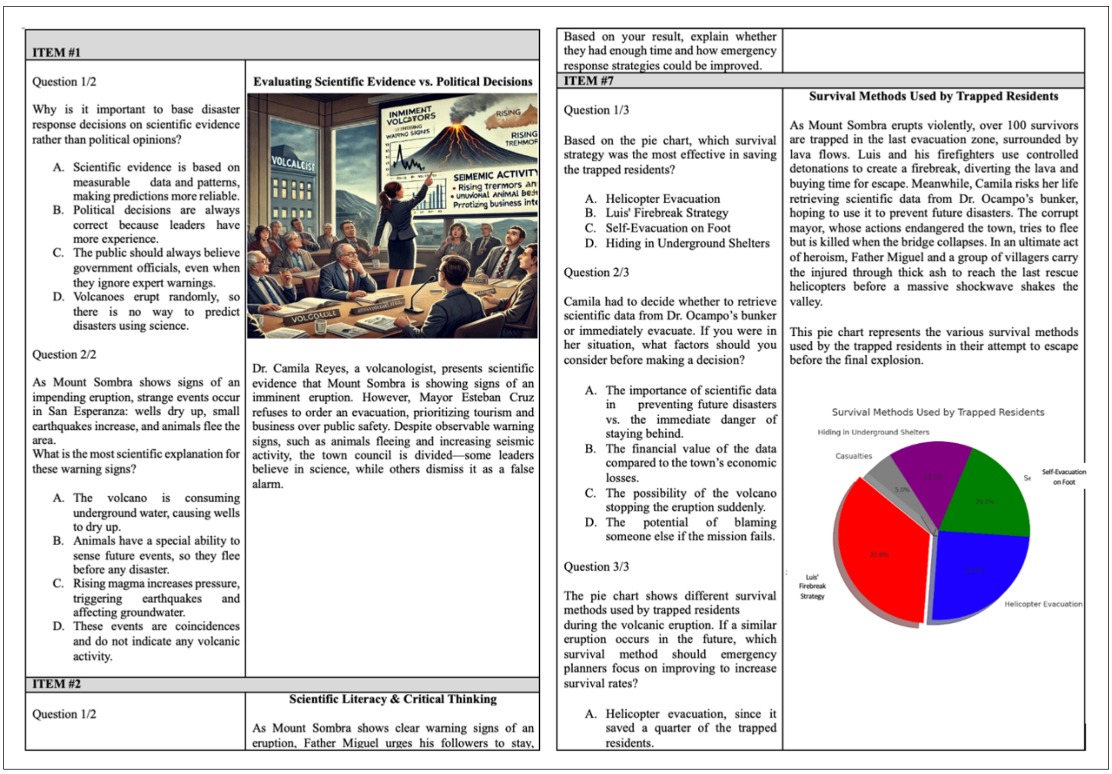
ComicSIM was developed in comic strip format based on the Grade 9 Science Most Essential Learning Competencies (MELCs) related to volcanic activity. The material follows a continuous fictional narrative set in a volcanic hazard context, incorporating characters who navigate scientific, ethical, and social dimensions of volcanic risk. The storyline was designed to integrate PISA scientific competency domains, including applying scientific knowledge to real-world issues, distinguishing scientific evidence from personal belief, and engaging in ethical reasoning under disaster conditions. Eight lessons were embedded within the narrative, each mapped to a specific PISA competency and assessed through corresponding items in the accompanying assessment tool. Figure 1 presents sample pages from the developed ComicSIM.

Two validation instruments were employed. The Comics Evaluation Rubric (CER) is a researcher-constructed 7-criterion evaluation tool developed through a review of existing rubrics for instructional comics (Rasmet et al., 2025; Tatalovic, 2009), SIM evaluation frameworks used in Philippine DepEd practice (Dacumos, 2016; Lituanas & Dela Cruz, 2024), and Mayer's (2009) Cognitive Theory of Multimedia Learning. The seven criteria—(a) content accuracy and relevance, (b) clarity and readability, (c) visual appeal and creativity, (d) interactivity and engagement, (e) pedagogical effectiveness, (f) cultural and contextual relevance, and (g) assessment and feedback—were selected to capture the dual requirements of ComicSIM as both a



(a) (b) (c)

Fig. 1. Structural components of the developed Comic Strategic Intervention Material (comicSIM) on volcanic activity: (a) front cover page showing the Title, (b) guide card displaying the target Grade 9 Science competency and explicit learning objectives, and (c) sample narrative comic panel integrating scientific concepts within the storyline.



ITEM #1

Question 1/2

Why is it important to base disaster response decisions on scientific evidence rather than political opinions?

A. Scientific evidence is based on measurable data and patterns, making predictions more reliable.
B. Political decisions are always correct because leaders have more experience.
C. The public should always believe government officials, even when they ignore expert warnings.
D. Volcanoes erupt randomly, so there is no way to predict disasters using science.

Question 2/2

As Mount Sombra shows signs of an impending eruption, strange events occur in San Esperanza: wells dry up, small earthquakes increase, and animals flee the area.
What is the most scientific explanation for these warning signs?

A. The volcano is consuming underground water, causing wells to dry up.
B. Animals have a special ability to sense future events, so they flee before any disaster.
C. Rising magma increases pressure, triggering earthquakes and affecting groundwater.
D. These events are coincidences and do not indicate any volcanic activity.

Evaluating Scientific Evidence vs. Political Decisions

Dr. Camila Reyes, a volcanologist, presents scientific evidence that Mount Sombra is showing signs of an imminent eruption. However, Mayor Esteban Cruz refuses to order an evacuation, prioritizing tourism and business over public safety. Despite observable warning signs, such as animals fleeing and increasing seismic activity, the town council is divided—some leaders believe in science, while others dismiss it as a false alarm.

ITEM #2

Question 1/2

Scientific Literacy & Critical Thinking

As Mount Sombra shows clear warning signs of an eruption, Father Miguel urges his followers to stay,

Based on your result, explain whether they had enough time and how emergency response strategies could be improved.

ITEM #7

Question 1/3

Based on the pie chart, which survival strategy was the most effective in saving the trapped residents?

A. Helicopter Evacuation
B. Luis' Firebreak Strategy
C. Self-Evacuation on Foot
D. Hiding in Underground Shelters

Question 2/3

Camila had to decide whether to retrieve scientific data from Dr. Ocampo's bunker or immediately evacuate. If you were in her situation, what factors should you consider before making a decision?

A. The importance of scientific data in preventing future disasters vs. the immediate danger of staying behind.
B. The financial value of the data compared to the town's economic losses.
C. The possibility of the volcano stopping the eruption suddenly.
D. The potential of blaming someone else if the mission fails.

Question 3/3

The pie chart shows different survival methods used by trapped residents during the volcanic eruption. If a similar eruption occurs in the future, which survival method should emergency planners focus on improving to increase survival rates?

A. Helicopter evacuation, since it saved a quarter of the trapped residents.

Survival Methods Used by Trapped Residents

As Mount Sombra erupts violently, over 100 survivors are trapped in the last evacuation zone, surrounded by lava flows. Luis and his firefighters use controlled detonations to create a firebreak, diverting the lava and buying time for escape. Meanwhile, Camila risks her life retrieving scientific data from Dr. Ocampo's bunker, hoping to use it to prevent future disasters. The corrupt mayor, whose actions endangered the town, tries to flee but is killed when the bridge collapses. In an ultimate act of heroism, Father Miguel and a group of villagers carry the injured through thick ash to reach the last rescue helicopters before a massive shockwave shakes the valley.

This pie chart represents the various survival methods used by the trapped residents in their attempt to escape before the final explosion.

Fig. 2. Sample assessment layout from the developed PISA-Aligned assessment tool illustrating the integration of scientific literacy and critical thinking items: (left) multiple-choice questions evaluating conceptual understanding of volcanic warning signs, and (right) an integrated data-analysis item requiring students to interpret graphical evidence (pie chart) regarding disaster response strategies.

comic-format material and a structured SIM. Content validity of the CER was established through a two-stage process: (1) independent review of criteria and descriptors by a measurement specialist for relevance, clarity, and construct alignment, and (2) informal piloting with two science educators external to the study prior to formal use. Jurors rated each criterion on a 5-point scale (5 = Excellent, 4 = Good, 3 = Satisfactory, 2 = Needs Improvement, 1 = Poor).

Table 2. Matrix of Item Distribution of the Assessment Tool

Lesson	Lesson Title	PISA Competency	No. of Items	Item Numbers
1	Volcanic Hazards and Predictions	Applying scientific knowledge to real-world issues	2	1, 2
2	Human Responses and Scientific Evidence in Disaster Preparedness	Distinguishing scientific evidence from personal beliefs	2	3, 4
3	Ethical Decision-Making and Disaster Risk Reduction	Scientific reasoning with social and ethical considerations	2	5, 6
4	Survival Skills and Risk Reduction During Eruptions	Explaining phenomena scientifically	3	7, 8, 9
5	Human Activities and Their Impact on Volcanic Systems	Interpreting data and evidence scientifically	1	10
6	Volcanic History, Warnings, and Emergency Response	Evaluating and designing scientific inquiry	4	11, 12, 13, 14
7	Planning for Eruptions: Survival Strategies and Safety Measures	Analyzing data and making evidence-based decisions	3	15, 16, 17
8	Geothermal Energy, Volcanic Risk, and Sustainable Development	Applying scientific principles to real-world situations	3	18, 19, 20

The PISA Science Framework Alignment Scale (PISFAS) is a 15-item evaluation instrument whose items were derived directly from the five core competency domains of the PISA 2022 Science Framework: Scientific Competencies, Scientific Knowledge, Contexts, Science Identity, and Environmental Science Competencies (OECD, 2023). Items were organized under four constructs: (1) relevance to the PISA Science Literacy Framework, (2) cognitive demand and inquiry-based learning, (3) format and engagement, and (4) overall PISA alignment. Content validity of the PISFAS was established through review by the measurement specialist, who verified the alignment of each item against the corresponding PISA framework competency. Ratings used a 4-point scale (4 = Strongly Agree, 3 = Agree, 2 = Disagree, 1 = Strongly Disagree). Internal consistency of the PISFAS in the present administration yielded Cronbach's Alpha = 0.88, confirming strong reliability for this use.

The PISA-aligned assessment tool comprises 20 items distributed across eight thematic lessons embedded within the ComicSIM narrative. Table 2 presents the mapping of lessons to PISA competencies and item distribution. Each lesson was constructed to address a distinct PISA scientific competency, ensuring that the assessment tool evaluates the full spectrum of scientific literacy skills outlined in the PISA 2022 framework. Item contexts draw directly from the ComicSIM narrative, requiring students to apply scientific reasoning to situations, characters, and decisions depicted in the comic (Figure 2). This contextual embedding is intentional: assessment items anchored in a narrative stimulus that students have already engaged with are hypothesized to reduce construct-irrelevant variance from unfamiliar item contexts, a recognized source of measurement error in PISA-style assessments (OECD, 2023).

Statistical analysis

Mean and standard deviation were computed to describe the central tendency and spread of juror ratings. Fleiss' Kappa (Fleiss, 1971) was calculated to determine inter-rater agreement beyond chance, interpreted using the Landis and Koch (1977) benchmarks: 0.00 to 0.20 as slight, 0.21 to 0.40 as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1.00 as almost perfect agreement. Cronbach's Alpha was computed to assess the internal consistency of the jurors' ratings on the PISA alignment evaluation sheet. All computations were performed using Jamovi (Version 2.4; The jamovi project, 2022).

Results

Table 3 presents the juror ratings of ComicSIM across its seven evaluative sub-criteria. All five jurors completed the evaluation independently, and no rating was changed following consultation.

ComicSIM received an overall weighted mean of 4.5 (Excellent) across all seven sub-criteria, with an overall Fleiss' Kappa of 0.70 indicating substantial agreement among the five jurors. The highest-rated dimension was Assessment and Feedback ($M = 4.8$, Excellent; $\kappa = 0.60$, Moderate agreement), reflecting strong consensus on the material's integration of meaningful evaluation tasks and feedback mechanisms. Content Accuracy and Relevance, Interactivity and Engagement, and Cultural and Contextual Relevance each received a weighted mean of 4.6 (Excellent), confirming that ComicSIM delivers scientifically accurate content within an engaging and culturally situated format. Visual Appeal and Creativity and Pedagogical Effectiveness were both rated 4.4 (Good), while Clarity and Readability received a rating of 4.0 (Good).

The pattern of Kappa values warrants careful interpretation. The fair Kappa values observed for most sub-criteria ($\kappa = 0.30$ to 0.40), despite uniformly high means, are consistent with a well-documented ceiling effect in expert validation studies: when all raters cluster at the upper end of a rating scale, chance-corrected agreement statistics such as Kappa are mathematically suppressed even when raters are highly consistent in their overall evaluative judgements (Feinstein & Cicchetti, 1990; McHugh, 2012). This phenomenon, sometimes termed the Kappa paradox, does not indicate disagreement among jurors; rather, it reflects the statistical consequence of near-unanimous high ratings. The Clarity and Readability dimension illustrates this point from

Table 3. Inter-Rater Agreement of ComicSIM Across Sub-Criteria

Criterion	Weighted Mean	Interpretation	Fleiss' Kappa (κ)	Agreement Level
Content Accuracy and Relevance	4.6	Excellent	0.40	Fair
Clarity and Readability	4.0	Good	1.00	Almost Perfect
Visual Appeal and Creativity	4.4	Good	0.30	Fair
Interactivity and Engagement	4.6	Excellent	0.40	Fair
Pedagogical Effectiveness	4.4	Good	0.40	Fair
Cultural and Contextual Relevance	4.6	Excellent	0.40	Fair
Assessment and Feedback	4.8	Excellent	0.60	Moderate
Overall	4.5	Excellent	0.70	Substantial

Table 4. Validation of the Developed PISA-Aligned Assessment Tool

PISA Construct	Items	Mean Score Range	Interpretation
Relevance to PISA Science Literacy Framework	1 to 4	3.6 to 4.0	Strongly Agree
Cognitive Demand and Inquiry-Based Learning	5 to 8	3.8 to 4.0	Strongly Agree
Format and Engagement	9 to 11	3.6	Strongly Agree
Overall Alignment with PISA	12 to 15	3.8 to 4.0	Strongly Agree
Total (All Items)	1 to 15	3.8	Strongly Agree

a different angle: its Kappa of 1.00 indicates that all five jurors assigned exactly the same score to this criterion, resulting in perfect agreement at a mean of 4.0. This pattern confirms that while ComicSIM's readability is consistently rated as Good, it is an area where targeted revision could elevate ratings to the Excellent range observed in other criteria.

Juror qualitative feedback reinforced these quantitative findings. Evaluators commended the material for its scientific accuracy, balance of scientific terminology with narrative storytelling, and its capacity to represent diverse perspectives including ethical and social considerations surrounding volcanic hazards. Areas recommended for refinement included increased localization of the narrative setting to enhance contextual relevance for learners from provincial Philippine communities such as Masbate, and the addition of greater narrative tension in the pre-eruption sequence to heighten dramatic engagement.

Table 4 presents juror ratings on the 15-item PISA alignment evaluation instrument.

The assessment tool received an overall mean rating of 3.8 (Strongly Agree) across all four PISA constructs, confirming strong expert consensus on its alignment with the PISA Science Literacy Framework. Items assessing Relevance to PISA (items 1 and 2) and Overall Alignment with PISA (item 12) received the highest scores (M = 4.0), reflecting unanimous agreement on their framework fidelity. Cognitive Demand and Inquiry-Based Learning items performed strongly (M = 3.8 to 4.0), affirming that the tool promotes higher-order thinking consistent with PISA's competency demands. Format and Engagement items received slightly lower ratings (M = 3.6), with one juror assigning lower scores on items 10 and 11, citing concerns about the visual layout of contextual stimuli alongside test questions. This finding points to a specific and addressable design issue rather than a fundamental validity problem with the tool.

The PISA-aligned assessment tool achieved an overall weighted mean of 4.93 (Excellent) across all five evaluative dimensions, with overall inter-rater agreement of Kappa = 0.87 (Almost Perfect). In contrast to the ceiling-effect pattern observed in the ComicSIM validation, the high Kappa values across PISA dimensions reflect genuine and consistent consensus among jurors at near-ceiling mean levels, indicating that the assessment tool's PISA alignment is clearly evidenced and uniformly recognized by independent evaluators. The Contexts dimension received a perfect score (M = 5.00, κ = 1.00), confirming unanimous consensus that ComicSIM's narrative effectively embeds PISA-type real-world contexts. Scientific Competencies, Scientific Knowledge, and Science Identity dimensions each yielded near-perfect scores (M = 4.93, κ = 0.87). Environmental Science Competencies produced slightly lower inter-rater agreement (M = 4.87, κ = 0.73, Substantial), suggesting minor variability in how jurors interpreted the scope of environmental science content within the volcanic activity context. Cronbach's Alpha of 0.88 confirmed strong internal consistency among the validation instrument's items, further establishing the reliability of the evaluation procedure (Table 5).

Table 5. Juror Assessment of the PISA-Aligned Assessment Tool Across Dimensions

PISA Dimension	Weighted Mean	Interpretation	Fleiss' Kappa (κ)	Agreement Level
Scientific Competencies	4.93	Excellent	0.87	Almost Perfect
Environmental Science Competencies	4.87	Excellent	0.73	Substantial
Scientific Knowledge	4.93	Excellent	0.87	Almost Perfect
Science Identity	4.93	Excellent	0.87	Almost Perfect
Contexts	5.00	Excellent	1.00	Almost Perfect
Overall	4.93	Excellent	0.87	Almost Perfect

Note. Cronbach's Alpha = 0.88. Kappa benchmarks follow Landis and Koch (1977).

Discussion

The findings of this study provide expert validation evidence for ComicSIM as a novel instructional material for Grade 9 Science in the Philippine context. It is worthy to emphasize that this evidence is restricted to expert validation, no classroom-based data on student learning outcomes were collected, and the present study does not constitute empirical evidence of instructional effectiveness. Within the scope of expert evaluation, the overall Excellent rating (M = 4.5) and substantial inter-rater agreement (κ = 0.70) position ComicSIM favorably alongside other expert-validated comic-based instructional materials reported in the literature. Fitria et al. (2023) similarly found strong expert validity ratings for digital science comics among Indonesian elementary students, while Samosa (2021) reported high positive attitudes toward the CoSIM hybrid format among Filipino Biology learners. The current study extends this body of validation evidence by explicitly integrating PISA-aligned competencies into the SIM structure, addressing a gap identified across most prior SIM development studies in the Philippines (Borabo & Dio, 2025).

The Excellent ratings received by Content Accuracy and Relevance, Interactivity and Engagement, and Cultural and Contextual Relevance indicate that expert evaluators judged ComicSIM to present scientifically accurate content within a format designed to engage learners and reflect their sociocultural context. This finding is consistent with Hosler and Boomer's (2011) argument that comics provide a narrative scaffold with potential to deepen engagement with scientific phenomena, and with Tatalovic's (2009) observation that narrative structures are associated with improved acquisition of technical vocabulary. It also aligns with constructivist principles: the comic's storyline is designed as a culturally familiar zone of proximal development through which students encounter volcanic science concepts in a supported and relatable context (Vygotsky, 1978). The highest-rated dimension, Assessment and Feedback (M = 4.8), suggests that embedding PISA-type questions directly within the comic narrative may be a promising design strategy; whether this translates into improved student metacognitive engagement requires empirical testing (Valio et al., 2025).

The Kappa values for most ComicSIM sub-criteria ranged from 0.30 to 0.40, which warrant careful contextualization. As noted in the Results, these values reflect the well-documented Kappa paradox: when expert raters consistently award high scores, the mathematical structure of chance-corrected agreement coefficients produces low Kappa values even when raters are, in practical terms, highly consistent (Feinstein & Cicchetti, 1990). This phenomenon is reported across the instructional material validation literature and does not indicate meaningful disagreement among the jurors in the present study (McHugh, 2012). The Clarity and Readability dimension, which yielded Kappa = 1.00 at a mean of 4.0, represents a distinct and instructive case: perfect agreement among jurors that the material is Good but not yet Excellent in readability. This finding converges with jurors' qualitative recommendations that dialogue and terminology be simplified for lower-performing readers, and identifies a specific, revision-ready improvement target for the next iteration of ComicSIM.

The PISA-aligned assessment tool's near-perfect inter-rater agreement (κ = 0.87) and Cronbach's Alpha of 0.88 establish strong reliability credentials that compare favorably with similar PISA-contextualized instruments in the literature (Rosana et al., 2020; Valio et al., 2025). The perfect score in the Contexts dimension (κ = 1.00, M = 5.00) is particularly noteworthy. It confirms that grounding assessment items within the comic narrative's volcanic hazard contexts effectively operationalizes the real-world situatedness that PISA designates as essential for measuring authentic scientific literacy (OECD, 2023). This result supports the emerging pedagogical argument that local, narrative-rich contexts may be more effective at eliciting PISA-type reasoning than decontextualized test items, because they reduce the cognitive burden of interpreting an unfamiliar stimulus while preserving the scientific reasoning demands of the task (Fananta et al., 2019; Valio et al., 2025).

Persistent misconceptions about volcanic predictability and risk are well documented in the geoscience education literature (Francek, 2013; King, 2010). ComicSIM's narrative design, which positions student-age characters as active reasoners within a volcanic emergency scenario, is theoretically consistent with findings that consequential decision-making contexts support conceptual restructuring in science education (Renouard & Mazabraud, 2018). Whether ComicSIM demonstrably reduces documented misconceptions in practice is a question that expert validation alone cannot answer; direct classroom investigation remains essential.

The slightly lower performance of the Format and Engagement construct in the PISA alignment instrument (M = 3.6) points to a specific and addressable design issue. One juror's lower ratings on items 10 and 11 identified concerns about the visual layout of contextual stimuli alongside test questions. Assessment tools that rely on rich narrative stimuli require careful attention to how comic panels, text, and test questions are spatially arranged on the page: stimulus material that competes visually with question text can increase cognitive load and compromise measurement validity. Revised versions should be formatted with the comic stimulus clearly separated from and preceding each question cluster, consistent with PISA item design conventions (OECD, 2023).

Taken together, the validation findings support further development and classroom testing of ComicSIM as a supplementary instructional resource in Grade 9 Science, particularly for volcanic activity units within the MATATAG curriculum. The expert evaluations indicate that ComicSIM is designed as a contextually resonant and internationally benchmarked instructional resource; however, its effectiveness for student learning must be confirmed through classroom-based research before recommendations for broader adoption can be made. Jurors' recommendation to localize narrative settings further is actionable: future versions of ComicSIM could incorporate local volcanoes such as Mayon or Bulusan, and community-specific disaster risk contexts familiar to students in volcanic hazard-prone municipalities.

Several limitations of the present study should be acknowledged. First, and most significantly, the evidence presented is restricted to expert validation and does not address student learning outcomes; claims about instructional effectiveness cannot be derived from the present data. Second, the validation panel comprised five jurors—consistent with established practice in Philippine SIM development research (Branch & Varank, 2009; Gall et al., 1996)—but this sample size limits the generalizability of the evaluative findings. Third, the CER and PISFAS are researcher-constructed instruments whose validity in this study rests primarily on face validity review by a measurement specialist; independent psychometric validation of these instruments is a priority for future research. Fourth, the study was conducted within a specific provincial context in the Philippines, and contextual relevance ratings may not fully generalize to learners in other settings. Classroom-based effectiveness testing using the validated instruments is identified as the essential and immediate next stage of this research programme.

Conclusions

This study developed and validated ComicSIM, a comic-based Strategic Intervention Material on volcanic activity, alongside a PISA-aligned assessment tool for Grade 9 Science in the Philippines. Expert juror evaluations confirmed that ComicSIM achieved Excellent overall quality ratings with substantial inter-rater agreement, while the PISA-aligned assessment tool demonstrated near-perfect content validity and strong internal consistency. These findings establish both instruments as suitable candidates for the next phase of development research: a classroom-based quasi-experimental study in which student participants engage with ComicSIM under naturalistic instructional conditions, with learning outcomes, scientific literacy skills, and misconception reduction measured through pre-test and post-test administration of the validated assessment tool. The present study provides validation evidence only; empirical demonstration of student learning gains awaits that subsequent phase.

The contribution of the present study to Philippine science education is twofold: it provides the field with an expert-validated, PISA-aligned instructional material for a high-priority Earth science topic, and it offers preliminary evidence that the comics-plus-SIM format is a promising approach to developing internationally benchmarked science learning resources within low-resource school environments. Science teachers and curriculum developers now have access to an expert-validated tool that requires no technology infrastructure, aligns directly with the MATATAG curriculum, and situates scientific reasoning within a volcanic hazard narrative that is both locally relevant and globally benchmarked. The present study does not, however, provide evidence of classroom effectiveness; such evidence must await the quasi-experimental study identified as the next phase of this research programme.

Acknowledgment

None.

Ethics Statement

This study involved voluntary expert evaluation by five qualified science education professionals. Participation was consensual and conducted in the jurors' professional capacity; no signed consent form was required. No student participants, vulnerable populations, or personal data were involved, and the study did not require formal ethics committee review under the guidelines of the authors' institutions. All procedures were conducted in accordance with ethical principles for educational research, including transparency in reporting and confidentiality of individual juror identities.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AI Transparency Statement

Artificial intelligence tools were used to assist with language editing and manuscript preparation. All content, interpretations, and conclusions are the sole responsibility of the authors, who attest to the accuracy and originality of the work.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Aghaei, N., Seyedin, H., & Sanaeinasab, H. (2018). Strategies for disaster risk reduction education: A systematic review. *Journal of Education and Health Promotion, 7*(1), 98. https://doi.org/10.4103/jehp.jehp_31_18

German, J.D., Redi, A.A.N.P., Ong, A.K.S., Prasetyo, Y.T., & Sumera, V.L.M. (2022). Predicting factors affecting preparedness of volcanic eruption for a sustainable community: A case study in the Philippines. *Sustainability, 14*(18), 11329. https://doi.org/10.3390/su141811329

Kurata, Y.B., Prasetyo, Y.T., Ong, A.K.S., Nadlifatin, R., Persada, S.F., Chuenyindee, T., & Cahigas, M.M.L. (2022). Determining factors affecting preparedness beliefs among Filipinos on Taal Volcano eruption in Luzon, Philippines. *International Journal of Disaster Risk Reduction,* 76, 103035. https://doi.org/10.1016/j.ijdrr.2022.103035

Department of Education (DepEd). (2018). *Statement on the Philippines' ranking in the 2018 PISA results.* https://www.deped.gov.ph/2019/12/04/statement-on-the-philippines-ranking-in-the-2018-pisa-results/

Cahyana, U., Supatmi, S., Erdawati, & Rahmawati, Y. (2019). The influence of web-based learning and learning independence toward student's scientific literacy in chemistry course. *International Journal of Instruction, 12*(4), 655–668. https://doi.org/10.29333/iji.2019.12442a

Degorio, J.M.L., & Langub, M.K.C. (2025). Enhancing physics education through comic strips: A development research study. *International Journal of Instruction, 18*(2), 637–652. https://doi.org/10.29333/iji.2025.18234a

Dacumos, L.P.N. (2016). Perspective of secondary teachers in the utilization of science Strategic Intervention Material (SIM) in increasing learning proficiency of students in science education. *AsTEN Journal of Teacher Education, 1*(2). https://doi.org/10.56278/asten.v1i2.293

Gabucan, J.R., & Sanchez, J.M.P. (2021). Strategic Intervention Material (SIM)-based instruction in teaching global warming in 9th grade science. *Formatif: Jurnal Ilmiah Pendidikan, 11*(1), 15–24. https://doi.org/10.30998/formatif.v11i1.6448

Borabo, J.G., & Dio, R.V. (2025). Effectiveness of Strategic Intervention Material (SIM) in improving Filipino students' performance in science and mathematics: A meta-analysis. *Journal of Basic Education Research, 6*(2), 79–89. https://doi.org/10.51405/jber.v6i2.1420

Mayer, R.E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Bruner, J. (1990). *Acts of meaning.* Harvard University Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Fitria, Y., Malik, A., Halili, S.H., & Amelia, R. (2023). Digital comic teaching materials: Its role to enhance student's literacy on organism characteristic topic. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(10), em2333. https://doi.org/10.29333/ejmste/13573

Hosler, J., & Boomer, K.B. (2011). Are comic books an effective way to engage nonmajors in learning and appreciating science? *CBE Life Sciences Education, 10*(3), 309–317. https://doi.org/10.1187/cbe.10-07-0090

Tatalovic, M. (2009). Science comics as tools for science education and communication: A brief, exploratory study. *Journal of Science Communication, 8*(4), A02. https://doi.org/10.22323/2.08040202

Molina, J.T., & Cadorna, E.A. (2025). Enhancing Filipino students' learning outcomes in physics using comic-based learning materials. *International Journal of Learning, Teaching and Educational Research, 24*(11), 185–204. https://doi.org/10.26803/ijlter.24.11.9

Samosa, R.C. (2021). CoSIM (comics cum SIM): An innovative material in teaching biology. *European Journal of Research Development and Sustainability, 2*(4), 19–28. https://files.eric.ed.gov/fulltext/ED618190.pdf

OECD. (2023). *PISA 2022 assessment and analytical framework.* OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

Valio, F.A., Safira, L., Aulliyah, U.A., & Suwarna, I.P. (2025). Development of a science literacy test for junior high school students based on the PISA 2025 framework. *Jurnal Pendidikan MIPA, 26*(3), 1377–1405. https://doi.org/10.23960/jpmipa.v26i3.pp1377-1405

Eisner, W. (2008). *Comics and sequential art: Principles and practices from the legendary cartoonist.* WW Norton & Company.

Rasmet, R.R., Shaafi, N.F., Sharif, S., Nasir, S.M., Salleh, M.F.M., & Ellianawati, E. (2025). The impacts of comic-based learning on student engagement in primary education: A systematic literature review. *Citra Journal of Innovative Education Practices, 1*(1), 26–39. https://doi.org/10.37934/cjiep.1.1.2639

Fananta, M.R., Umbara, T., & Valdez, N.P. (2019). The effect of inquiry-based science comics on primary learners' scientific literacy skills. *Global Journal of Business and Social Science Review, 7*(4), 210–216. https://doi.org/10.35609/gjbssr.2019.7.4(1)

Rosidah, C.T., Putrayasa, I.B., Wesnawa, I.G.A., & Candiasa, I.M. (2022). Thematic comic to cultivate eco-literacy for young learners. *Kasetsart Journal of Social Sciences, 43*(3), 735–740. https://doi.org/10.34044/j.kjss.2022.43.3.27

Valente, B., & Faria, C. (2026). Turning the page on a missed opportunity: Science comics through the eyes of future elementary teachers. *Journal of Education.* https://doi.org/10.1177/00220574261428474

Lituañas, A.B., & Dela Cruz, F.M. (2024). Development and validation of contextualized strategic intervention materials in science for Grade 6 learners. *Randwick International of Education and Linguistics Science Journal, 5*(4), 1363–1383. https://doi.org/10.47175/rielsj.v5i4.1103

Branch, R.M., & Varank, İ. (2009). *Instructional design: The ADDIE approach.* Springer.

Gall, M.D., Borg, W.R., & Gall, J.P. (1996). *Educational research: An introduction.* Longman Publishing.

Francek, M. (2013). A compilation and review of over 500 geoscience misconceptions. *International Journal of Science Education, 35*(1), 31–64. https://doi.org/10.1080/09500693.2012.736644

King, C. (2010). An analysis of misconceptions in science textbooks: Earth science in England and Wales. *International Journal of Science Education, 32*(5), 565–601. https://doi.org/10.1080/09500690902721681

Demirci, T., & Okur, S. (2021). The effect of teaching science through storytelling on students' academic achievement, story writing skills and opinions about practice. *Education Quarterly Reviews, 4*(2), 562–578. https://doi.org/10.31014/aior.1993.04.02.301

Korur, F., Erduran Avcı, D., & Gürkan, Ş. (2025). Comparison of the effects of stories and digital stories containing refutation texts on conceptual understanding and correction of misconceptions in primary science education. *SAGE Open, 15*(4), 1–19. https://doi.org/10.1177/21582440251403736

Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378–382. https://doi.org/10.1037/h0031619

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

The jamovi project. (2022). *Jamovi* (Version 2.4) [Computer software]. https://www.jamovi.org

Feinstein, A.R., & Cicchetti, D.V. (1990). High agreement but low Kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology, 43*(6), 543–549. https://doi.org/10.1016/0895-4356(90)90158-L

McHugh, M.L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

Rosana, D., Widodo, E., Setyaningsih, W., & Setyawarno, D. (2020). Developing assessment instruments of PISA model to measure students' problem-solving skills and scientific literacy in junior high schools. *Jurnal Pendidikan Sains Indonesia, 8*(2), 292–305. https://doi.org/10.24815/jpsi.v8i2.17468

Renouard, A., & Mazabraud, Y. (2018). Context-based learning for Inhibition of alternative conceptions: The next step forward in science education. Npj Science of Learning, 3(1), 10. https://doi.org/10.1038/s41539-018-0026-9

Розроблення та валідація КомікСІМ з теми вулканічної активності з інструментом оцінювання, узгодженим із ПІЗА

Шейла Бараян[1ABCD], Ерді Кахурао[2ACD]

[1]Інтегрована школа Буенсуерте
[2]Коледж освіти, Державний меморіальний коледж сільського господарства і технологій імені доктора Еміліо Б. Еспіноси-старшого

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 10 с., 5 табл., 2 рис., 38 джерел.

Мета. Навчальні матеріали на основі коміксів демонструють вимірюване підвищення залученості учнів і рівня природничо-наукової грамотності; проте лише в небагатьох дослідженнях формат коміксів інтегрується зі стратегічними інтервенційними матеріалами (SIM) та оцінюванням, узгодженим із PISA, у викладанні наук про Землю в середніх школах Філіппін. У цьому дослідженні було розроблено та валідовано ComicSIM — стратегічний інтервенційний матеріал у форматі коміксу з теми вулканічної активності для природничих наук у 9 класі, а також узгоджений із PISA інструмент оцінювання, що містить 20 завдань за вісьмома уроками.

Матеріали та методи. Використано описово-розробницький дизайн дослідження. П'ять експертів, які мали наукові ступені у галузі природничо-наукової освіти або розроблення навчальних програм і щонайменше п'ять років професійного досвіду, незалежно оцінювали обидва інструменти. ComicSIM оцінювали за сімома підкритеріями з використанням 5-бальної шкали; узгодженість оцінок експертів визначали за допомогою каппи Флейса. Інструмент оцінювання, узгоджений із PISA, оцінювали за 15-пунктовим інструментом відповідності з використанням 4-бальної шкали; внутрішню узгодженість визначали за допомогою альфи Кронбаха.

Результати. ComicSIM отримав середньозважену оцінку 4,5 («відмінно») та значення каппи Флейса 0,70 (суттєва узгодженість). Інструмент оцінювання, узгоджений із PISA, мав середню оцінку відповідності 3,8 («повністю згоден»), середню оцінку за вимірами 4,93 («відмінно») та значення каппи 0,87 (майже повна узгодженість). Значення альфи Кронбаха 0,88 підтвердило високу внутрішню узгодженість.

Висновки. Обидва інструменти пройшли експертну валідацію, є надійними та придатними для використання з метою розвитку природничо-наукової грамотності, узгодженої з PISA, у викладанні природничих наук у середніх школах Філіппін; водночас їхня ефективність у класі потребує підтвердження подальшими емпіричними дослідженнями.

Ключові слова: ComicSIM, стратегічний інтервенційний матеріал, PISA, вулканічна активність, природничі науки у 9 класі.

Information about the Authors

Barajan Sheila: sheila.flaviano@deped.gov.ph; https://orcid.org/0009-0003-6249-6398; Buenasuerte Integrated School, Department of Education, Schools Division of Masbate Province, National Road, Buenasuerte, Pio V. Corpuz, Masbate, Philippines.

Cajurao Erdee: eccajurao@debesmscat.edu.ph; https://orcid.org/0000-0001-5695-2850; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology (DEBESMSCAT), Cabitan, Mandaon, Masbate, Philippines.

Cite this article as: Barajan, S., & Cajurao, E. (2026). Development and Validation of ComicSIM on Volcanic Activity With a PISA-Aligned Assessment Tool. *Journal of Learning Theory and Methodology, 7*(2), 104-113. https://doi.org/10.17309/jltm.2026.7.2.05

Received: 11.06.2026. Accepted: 18.07.2026. Published: 16.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Original Scientific Article

Data Literacy and Computational Thinking Among Secondary Science Teachers: A Quantitative Study of Levels, Differences and Instructional Integration

John Carlo Mortel[1ABCD] and Erdee Cajurao[2ACD]

[1]Federico A. Estipona Memorial High School
[2]College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.06

Abstract

Objectives. This study examined the levels of data literacy (DL) and computational thinking (CT) skills among secondary science teachers, the extent of their integration into science instruction, and the relationships between these competencies and instructional integration.

Materials and Methods. Using a descriptive-correlational design, 53 secondary science teachers in Masbate, Philippines, completed three researcher-developed Likert-type instruments measuring DL, CT, and DL-CT integration into science instruction. Data were analyzed using means and standard deviations, the Wilcoxon Signed-Rank Test, and Spearman rank-order correlation.

Results. Teachers rated themselves Proficient in both DL ($M = 3.92$, $SD = 0.63$) and CT ($M = 3.84$, $SD = 0.64$), with a statistically significant difference between the two, $W = 288.00$, $p = .040$, $r = .28$. Teachers reported a Highly Integrated level of DL-CT instructional integration ($M = 3.90$, $SD = 0.82$). Spearman correlations revealed strong positive associations of both DL ($r_s = .740$) and CT ($r_s = .757$) with integration, both $p < .001$, with algorithmic thinking emerging as the single strongest correlate ($r_s = .785$).

Conclusion. Secondary science teachers demonstrated proficient but uneven data literacy and computational thinking skills, with abstraction and access to computational tools emerging as relative weaknesses. Both competencies played a substantial and complementary role in supporting integrated science instruction, and targeted professional development addressing these weaker areas is recommended.

Keywords: data literacy, computational thinking, science education, instructional integration, secondary teachers.

Introduction

The 21st century has transformed the learning competencies that students must possess to thrive in an era defined by data and technology. Among the most consequential of these competencies are data literacy and computational thinking, two interconnected skill sets that have gained substantial recognition in educational policy, curriculum reform and pedagogical scholarship worldwide. In the domain of science education, these competencies are not merely supplementary skills but fundamental epistemic tools that align directly with the nature of scientific inquiry, which is inherently quantitative, data intensive and systematic (National Research Council, 2012; Weintrop et al., 2016).

Ridsdale et al. (2015) defined data literacy as the ability to collect, manage, evaluate and apply data in a critical and meaningful manner. It is increasingly recognized as a foundational competency for informed citizenship and scientific participation. In the science classroom, students who are data literate can engage meaningfully with empirical evidence, interpret graphs and tables, evaluate the validity and reliability of scientific claims and draw evidence-based conclusions (Kjelvik & Schultheis, 2019). These capabilities mirror the authentic practices of scientists and are central to the scientific inquiry standards embedded in contemporary science curriculum frameworks (NGSS Lead States, 2013).

Meanwhile, computational thinking (CT), as articulated by Wing (2006) in her landmark formulation, refers to the cognitive processes involved in formulating problems and their solutions in ways that can be carried out by an information-processing agent. CT encompasses decomposition (breaking complex real-world problems into manageable parts), pattern recognition (identifying regularities and trends), abstraction (filtering relevant from irrelevant information) and algorithmic thinking (designing step-by-step solutions).

Corresponding Author: John Carlo Mortel,
e-mail: johncarlomortel17@gmail.com

Broadly speaking, computational thinking entails mathematical reasoning, design-based thinking and the ability to reason at several levels of abstraction (Kramer, 2007; Wing, 2008). Additionally, it involves applying these types of reasoning to contextual problem solving (Guzdial, 2008; Wing, 2008). In science education, CT provides students with powerful strategies for modeling natural phenomena, analyzing large datasets, simulating complex systems and designing controlled experiments (Grover & Pea, 2013; Weintrop et al., 2016).

The integration of data literacy and computational thinking into science instruction has received significant impetus from several influential policy documents and curriculum frameworks. The Next Generation Science Standards (NGSS) in the United States explicitly position data analysis, mathematical and computational thinking and the use of models as core scientific and engineering practices (NGSS Lead States, 2013). Similarly, the National Research Council's Framework for K-12 Science Education identifies computational tools as essential instruments for contemporary scientific practice. These frameworks reflect a broader recognition that science education in the 21st century must prepare students not only to understand scientific content but to engage in the processes and practices that constitute authentic scientific work.

Despite this growing attention, empirical evidence suggests a significant gap between the theoretical articulation of these competencies in curriculum documents and their actual realization in classroom practice. Barr and Stephenson (2011) documented widespread unpreparedness among teachers to integrate computational thinking into their instructional practice, indicating that most K-12 teachers had received little or no professional preparation in this area. Lee (2025) stressed that teachers often recognize and appreciate the value of data but remain uncertain due to low self-efficacy in independently analyzing data or using complex computational tools. Similarly, Carlson et al. (2011) found that even in higher education, students' data information literacy skills were frequently insufficient for the demands of research and evidence-based learning.

In the Philippines, the K to 12 Basic Education Curriculum reform, implemented by the Department of Education (DepEd) beginning in 2012, explicitly articulates the goal of developing scientifically, technologically and mathematically literate Filipinos who are capable of critical thinking and problem solving (Department of Education, 2016). The Senior High School Science curriculum includes components that implicitly demand data literacy and computational thinking, such as research-based learning, data-driven argumentation and laboratory investigations requiring quantitative analysis. However, the extent to which Filipino science teachers are equipped to facilitate the development of these competencies in their students, and the degree to which these competencies are systematically addressed in science instruction, remain empirically underexplored.

This study therefore seeks to provide empirical evidence on the self-assessed level of data literacy and computational thinking skills among science teachers, the extent of their integration into science instruction and the relationship between these two core competencies in the context of Philippine secondary science education. By employing a quantitative, descriptive-correlational research design, this study aims to generate data that are both statistically rigorous and practically actionable for curriculum developers, school administrators and science educators and replicable of future researchers.

Specifically, the study sought to determine the level of data literacy of secondary science teachers along the dimensions of data collection and management, data analysis and interpretation, data visualization and data-based decision making; the level of their computational thinking skills along the dimensions of decomposition, pattern recognition, abstraction and algorithmic thinking; whether a statistically significant difference exists between their self-assessed data literacy and computational thinking skills; the extent to which data literacy and computational thinking skills are integrated into science instruction; and the correlation between this instructional integration and each of the two competencies.

Review of Related Literature

Data Literacy in Science Education

The concept of data literacy has evolved substantially over the past two decades. Ridsdale et al. (2015) proposed a comprehensive data literacy framework identifying core competencies, namely data collection, data management, data evaluation, data analysis, data visualization and data-based decision making. This multidimensional view positioned data literacy not as a single homogeneous skill but as a complex competency that develops progressively and must be supported by intentional instructional and institutional conditions. Calzada Prado and Marzal (2013) offered a complementary perspective from the information literacy tradition, emphasizing the critical and evaluative dimensions of data literacy, particularly the capacity to assess the quality, reliability and limitations of data.

In the context of teacher professional development, Pastore et al. (2025) conducted a scoping review of teacher data literacy research published between 2020 and 2024 and found that most studies reported moderate to high data literacy levels among in-service teachers who had received structured professional development. Mandinach and Schildkamp (2021) further argued that when teachers view data as actionable evidence for improving instruction, decision-making naturally becomes the most strongly developed data literacy dimension. However, Sandoval-Ríos et al. (2025) noted a persistent skills gap in data literacy training within teacher education programs, particularly in developing-country contexts where statistical and data skills are often excluded from pre-service curricula.

Within data literacy, visualization has been identified as a particularly challenging dimension. Laramee (2023) emphasized that visualization remains cognitively complex, requiring learners to simultaneously navigate statistical, contextual and communicative layers of meaning. Fernandez Nieto et al. (2022) similarly noted that data visualization tasks can exceed the professional skill set of educators without strong analytical foundations. Wilkerson et al. (2025) found that teachers with higher visualization literacy were more effective in facilitating student-centered data visualization discussions in science classrooms.

Computational Thinking in Science Education

Wing's (2006) seminal framework conceptualized computational thinking as a universally applicable cognitive process extending across all disciplines. She argued that CT involves formulating problems and their solutions in ways that make them tractable for computational processing, emphasizing that decomposition, pattern recognition, abstraction and algorithm design are fundamental intellectual tools. Grover and Pea (2013) subsequently elaborated this framework for K-12 contexts, while Weintrop et al. (2016) further operationalized CT practices for STEM education, namely data practices, modeling and simulation practices, computational problem-solving practices and systems-thinking practices.

Research has consistently identified abstraction as the most challenging CT dimension. Oser et al. (2025), in a systematic review, concluded that abstraction remains the least proficiently developed CT skill in K-12 education, noting significant inconsistencies in how it is conceptualized and taught. Asif et al. (2024) observed that while teachers can engage in surface-level abstraction, deeper forms remain challenging without dedicated professional learning. Asif et al. (2023) further found that teachers often simplify abstraction to merely ignoring details, missing its full computational meaning as level-switching and information suppression.

Mertens and Colunga (2025) identified decomposition as the most cognitively accessible CT component, as it maps naturally onto scientific problem-solving processes such as experimental design. In contrast, Kong et al. (2020) and Rich et al. (2021) reported that even after targeted professional development, many K-12 teachers continued to demonstrate inconsistent CT proficiency. Hsu et al. (2024), in their systematic review, documented that despite growing momentum in CT professional development, there remains a notable absence of consensus on how to assess and build teachers' CT knowledge.

Integration of Data Literacy and Computational Thinking into Instruction

Coenraad et al. (2022) analyzed CT integration in elementary science lesson plans and found that the most successful integration sites were those connected to authentic data analysis tasks embedded in real-world science contexts. Moon et al. (2023) observed that teachers who integrated CT through contextual, problem-based science tasks demonstrated stronger instructional coherence than those who taught CT as a standalone activity. However, the UNESCO Global Education Monitoring Report (2023) documented that digital tool use in science classrooms remains limited even in high-income countries, with simulation and modeling software available to fewer than one-third of students in many contexts.

Angeli et al. (2016) cautioned that the pedagogical integration of computational tools requires dedicated institutional support, including infrastructure provisioning, technical assistance and subject-specific training. Kadioglu-Akbulut et al. (2023) highlighted similar concerns in their study of science teachers' TPACK (Technological Pedagogical Content Knowledge) levels. Lim et al. (2023) found that even when teachers held strong CT beliefs and pedagogical knowledge, technology-tool integration lagged behind due to insufficient discipline-specific training. Cui and Zhang (2022) argued that teacher data competencies must be integrated with pedagogical content knowledge rather than treated as standalone technical skills, a position supported by the broader TPACK-integrated data literacy framework.

Theoretical Framework

This study is grounded in the convergence of three theoretical foundations, namely Constructivism, Wing's Computational Thinking Framework and the Data Literacy Framework.

Constructivism, as elaborated in the foundational works of Piaget (1952) and Vygotsky (1978), posits that learners actively construct knowledge through their interactions with the environment and with other learners, rather than passively receiving information transmitted by teachers. Piaget's theory emphasizes the role of active engagement, assimilation and accommodation in building increasingly sophisticated cognitive schemas. Vygotsky's sociocultural constructivism extends this view by highlighting the central role of social interaction, language and cultural tools, including computational and data tools, in mediating higher-order cognitive development. In the context of science education, constructivism provides the theoretical basis for viewing data literacy and computational thinking not as isolated technical skills to be taught demonstratively, but as higher-order cognitive capacities developed through active, inquiry-based engagement with authentic data and computational problems. Vygotsky's Zone of Proximal Development is particularly relevant, as it underscores the importance of teachers' own competence in scaffolding students' development of these complex skills.

Wing's (2006) computational thinking framework and its subsequent elaborations by Grover and Pea (2013) and Weintrop et al. (2016) provide the conceptual structure for the CT construct in this study. These operationalizations directly inform the measurement of decomposition, pattern recognition, abstraction and algorithmic thinking as applied to science education contexts.

The data literacy framework articulated by Ridsdale et al. (2015), complemented by Calzada Prado and Marzal's (2013) emphasis on the evaluative dimensions of data literacy, provides the conceptual foundation for the DL construct measured in this study. Together, these three frameworks generate the expectation that teachers who possess stronger data literacy and computational thinking skills will be better equipped to design, implement and sustain instruction that integrates these competencies authentically into science learning.

Materials and Methods

Study Participants

This study utilized purposive sampling (Etikan et al., 2016) to select respondents based on criteria directly aligned with the research objectives. A total of 53 secondary science teachers from selected high schools in Masbate, Philippines participated in the study. All respondents were currently teaching science subjects in Grades 7 to 12 under the K to 12 curriculum and were assigned to teach Earth Science, Biology, Chemistry, Physics or Integrated Science. All were employed by the Department of Education (DepEd) in the province of Masbate during the data gathering period. Teachers were included regardless of their years in the teaching profession, which allowed the sample to represent a wide range of teaching experience rather than a single career stage. It should be acknowledged that the use of purposive sampling with a modest sample size (n = 53) limits the generalizability of the findings to the broader population of Filipino secondary science teachers, and the findings are therefore intended to be suggestive of patterns warranting further investigation with larger, probability-based samples from multiple provinces and regions.

Study Organization

This study employed a descriptive-correlational research design, appropriate when a study seeks to measure and describe existing phenomena and examine statistical relationships among variables without manipulation or intervention (Creswell, 2014). The study utilized three researcher-developed instruments, namely the Data Literacy Scale (DLS), the Computational Thinking Skills Scale (CTSS) and the Data Literacy and Computational Thinking Skills Integration in Science Instruction Scale (DLCTSIS), developed based on the Data Literacy Framework (Calzada Prado & Marzal, 2013; Ridsdale et al., 2015) and the Computational Thinking Framework (Weintrop et al., 2016; Wing, 2006). All three instruments were subjected to content validation by a panel of science education experts and pilot tested with 20 science teachers who were not included in the actual study sample. The DLS is a 16-item Likert-type scale measuring data collection and management, data analysis and interpretation, data visualization and data-based decision making, with items rated from 5 (Expert) to 1 (Novice) and a pilot Cronbach's alpha of 0.93. The CTSS is a 16-item Likert-type scale measuring decomposition, pattern recognition, abstraction and algorithmic thinking, rated on the same scale, with a pilot Cronbach's alpha of 0.93. The DLCTSIS is a 10-item Likert-type scale with indicators adapted from the DL framework, the CT framework, the Next Generation Science Standards (NGSS Lead States, 2013) and the statistical thinking framework of Wild and Pfannkuch (1999), rated from 5 (Fully Integrated) to 1 (Not Integrated At All), with a pilot Cronbach's alpha of 0.91 (all interpreted as Excellent per George & Mallery, 2019). Data collection proceeded as follows. An official letter of permission was first secured from the DepEd Masbate Division Office. Survey questionnaires, together with a Participant Information Sheet and a consent form, were then distributed to all qualified respondents through Google Forms, and completed questionnaires were gathered within one week of distribution. All respondents were informed of the purpose of the study, the nature of the data collected, the voluntary nature of participation, their right to withdraw at any time and the confidentiality and data protection measures in place, and the survey proceeded only upon affirmative consent. Participation was entirely voluntary, with no incentive, coercion or pressure applied, and all data were stored on an access-controlled server to be deleted following the prescribed retention period. This study was conducted in compliance with Republic Act No. 10173, the Data Privacy Act of 2012 of the Philippines. Finally, all valid responses were encoded into a statistical analysis software database for analysis.

Statistical Analysis

The data analysis procedure utilized both descriptive and inferential statistics. The Shapiro-Wilk test was used to assess the normality of data distributions, which informed the selection of parametric or non-parametric procedures. To address the levels of data literacy, computational thinking skills and their integration into science instruction, means, standard deviations and verbal interpretations were computed. To analyze whether a significant difference exists between data literacy and computational thinking skills, the Wilcoxon Signed-Rank Test was employed as the appropriate non-parametric alternative to the paired t-test, given that both variables violated the normality assumption. To determine the correlations between data literacy, computational thinking skills and their integration into science instruction, Spearman's rank-order correlation (rs) was used given the non-normal distributions. The following interpretation thresholds were applied. Values from 0.00 to 0.19 indicate negligible correlation, 0.20 to 0.39 weak correlation, 0.40 to 0.69 moderate correlation, 0.70 to 0.89 strong correlation and 0.90 to 1.00 very strong correlation.

Results

Level of Data Literacy of Secondary Science Teachers

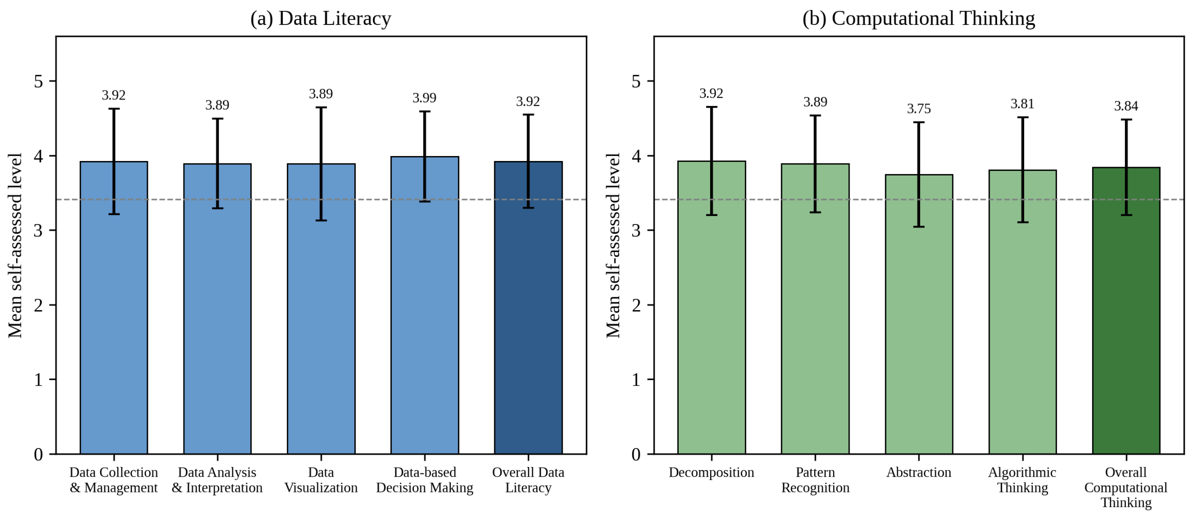
Table 1 and Figure 1(a) present the domain-level means and standard deviations for all 53 respondents across the four data literacy dimensions. All four dimensions were rated at the Proficient level (M range = 3.89 to 3.99). Data-based decision making received the highest mean (M = 3.99, SD = 0.60), followed by data collection and management (M = 3.92, SD = 0.71), data analysis and interpretation (M = 3.89, SD = 0.60) and data visualization (M = 3.89, SD = 0.76). The overall data literacy mean of 3.92 (SD = 0.63) falls within the Proficient range (3.41 to 4.20).

Table 1. Level of Data Literacy of Secondary Science Teachers

Domain	n	M	SD	Level
Data Collection and Management	53	3.92	0.71	Proficient
Data Analysis and Interpretation	53	3.89	0.60	Proficient
Data Visualization	53	3.89	0.76	Proficient
Data-based Decision Making	53	3.99	0.60	Proficient
Overall Data Literacy	53	3.92	0.63	Proficient

Level of Computational Thinking Skills of Secondary Science Teachers

Table 2 and Figure 1(b) present the domain-level means and standard deviations for computational thinking skills across all 53 respondents. All four dimensions were rated at the Proficient level. Decomposition registered the highest



Note. Dashed line marks the lower bound of the Proficient range (M = 3.41). Error bars represent SD.

Fig. 1. Domain-Level Means for (a) Data Literacy and (b) Computational Thinking Skills

mean (M = 3.92, SD = 0.73), followed by pattern recognition (M = 3.89, SD = 0.65), algorithmic thinking (M = 3.81, SD = 0.71) and abstraction (M = 3.75, SD = 0.70). The overall computational thinking skills mean (M = 3.84, SD = 0.64) places respondents within the Proficient level.

Table 2. Level of Computational Thinking Skills of Secondary Science Teachers

Domain	n	M	SD	Level
Decomposition	53	3.92	0.73	Proficient
Pattern Recognition	53	3.89	0.65	Proficient
Abstraction	53	3.75	0.70	Proficient
Algorithmic Thinking	53	3.81	0.71	Proficient
Overall Computational Thinking	53	3.84	0.64	Proficient

Difference Between Data Literacy and Computational Thinking Skills

The Shapiro-Wilk test of normality was conducted for all variables given the sample size (n = 53). Both variables violated the normality assumption, with data literacy showing W = 0.943, p = .014 and computational thinking skills showing W = 0.926, p = .003. Given these violations, the Wilcoxon Signed-Rank Test was employed as the appropriate non-parametric alternative. As shown in Table 3, the test revealed a statistically significant difference between the overall data literacy (M = 3.92) and computational thinking skills (M = 3.84) scores, W = 288.00, p = .040, r = 0.28.

Table 3. Wilcoxon Signed-Rank Test Comparing Data Literacy and Computational Thinking Skills

Comparison	Wilcoxon W	p-value	Effect Size (r)
DL vs. CT	288.00	.040	0.28 (Small)

Level of Data Literacy and Computational Thinking Skills Integration into Science Instruction

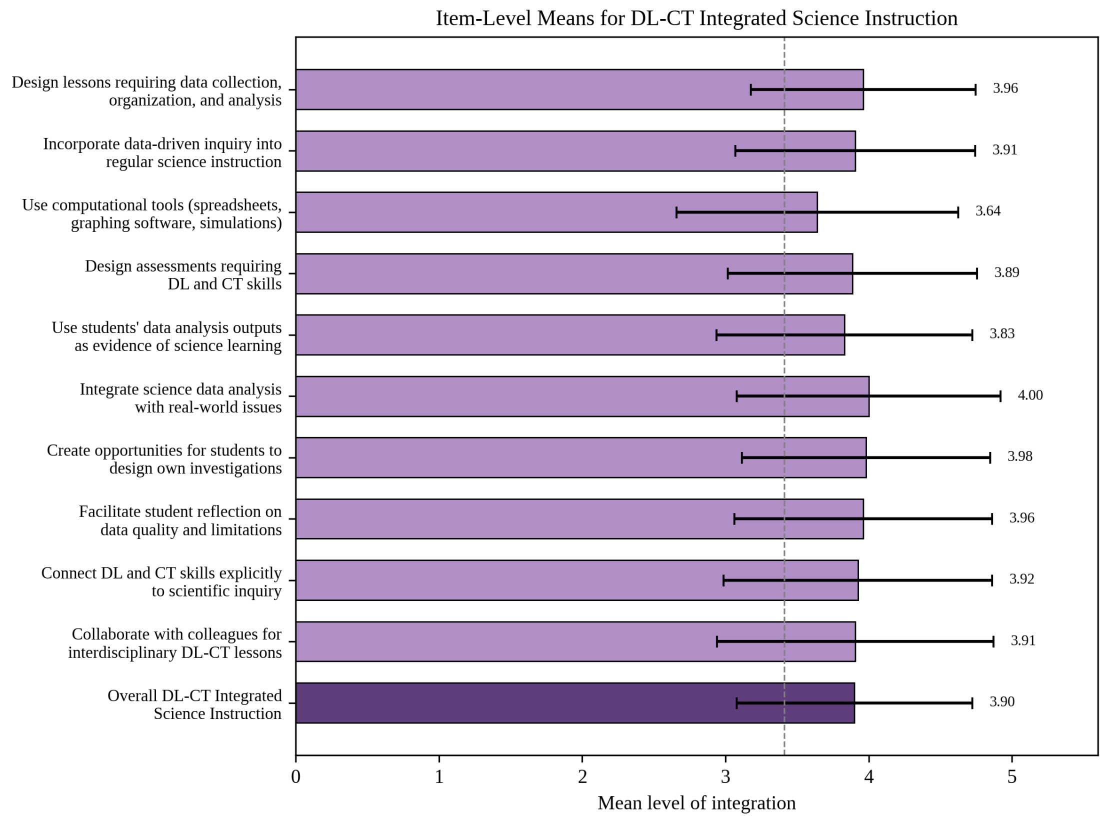
Table 4 and Figure 2 present item-level means for the DL-CT Integrated Science Instruction scale. The overall mean was M = 3.90 (SD = 0.82), placing it at the Highly Integrated level. The highest-rated item was integrating science data analysis with real-world issues (M = 4.00, SD = 0.92), while the lowest-rated item was the use of computational tools such as spreadsheets, graphing software and simulations in science lessons (M = 3.64, SD = 0.98).

Table 4. Level of Data Literacy and Computational Thinking Skills Integration into Science Instruction

Indicator	M	SD	Interpretation
Design lessons requiring data collection, organization, and analysis	3.96	0.78	Highly Integrated
Incorporate data-driven inquiry into regular science instruction	3.91	0.84	Highly Integrated
Use computational tools (spreadsheets, graphing software, simulations)	3.64	0.98	Highly Integrated
Design assessments requiring DL and CT skills	3.89	0.87	Highly Integrated
Use students' data analysis outputs as evidence of science learning	3.83	0.89	Highly Integrated
Integrate science data analysis with real-world issues	4.00	0.92	Highly Integrated
Create opportunities for students to design own investigations	3.98	0.87	Highly Integrated
Facilitate student reflection on data quality and limitations	3.96	0.90	Highly Integrated
Connect DL and CT skills explicitly to scientific inquiry	3.92	0.94	Highly Integrated
Collaborate with colleagues for interdisciplinary DL-CT lessons	3.91	0.97	Highly Integrated
Overall DL-CT Integrated Science Instruction	3.90	0.82	Highly Integrated

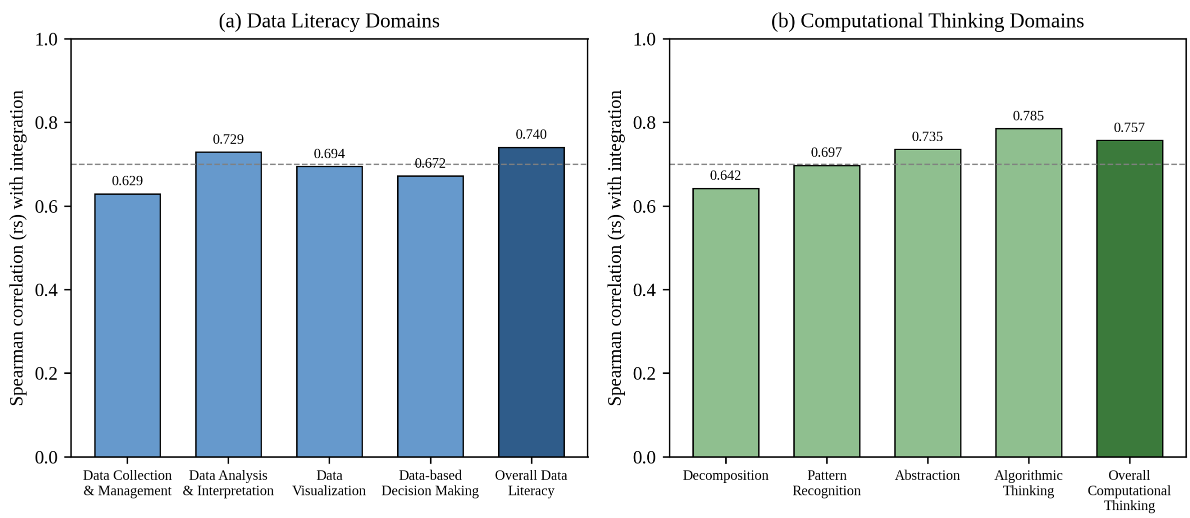
Correlation Between Data Literacy and Integrated Science Instruction

Spearman rank-order correlations were computed between all DL domains and the DL-CT integration score, given the violation of normality. As shown in Table 5, all four DL domains



Note. Dashed line marks the lower bound of the Highly Integrated range (M = 3.41). Error bars represent SD.

Fig. 2. Item-Level Means for DL-CT Integrated Science Instruction



Note. Dashed line marks the lower bound of a strong correlation (rs = .70). All correlations significant at p < .001.

Fig. 3. Spearman Correlations Between (a) Data Literacy Domains and (b) Computational Thinking Domains With DL-CT Integrated Science Instruction

showed statistically significant strong positive correlations with DL-CT integrated instruction. The overall correlation between data literacy and integration was rs = .740 ($p < .001$). At the domain level, data analysis and interpretation emerged as the strongest correlate (rs = .729), followed by data visualization (rs = .694), data-based decision making (rs = .672) and data collection and management (rs = .629).

Correlation Between Computational Thinking Skills and Integrated Science Instruction

As shown in Table 6 and Figure 3, all four CT domains showed statistically significant strong positive correlations with DL-CT integrated science instruction. Algorithmic thinking was the strongest correlate (rs = .785), followed by abstraction (rs = .735), pattern recognition (rs = .697) and decomposition (rs = .642). The overall CT-integration correlation was rs = .757 ($p < .001$).

Table 5. Spearman Correlation Between Data Literacy and DL-CT Integrated Science Instruction

DL Domain	Spearman rs	p-value	Interpretation
Data Collection and Management	.629	< .001	Moderate
Data Analysis and Interpretation	.729	< .001	Strong
Data Visualization	.694	< .001	Moderate
Data-based Decision Making	.672	< .001	Moderate
Overall Data Literacy	.740	< .001	Strong

Table 6. Spearman Correlation Between Computational Thinking Skills and DL-CT Integrated Science Instruction

CT Domain	Spearman rs	p-value	Interpretation
Decomposition	.642	< .001	Moderate
Pattern Recognition	.697	< .001	Moderate
Abstraction	.735	< .001	Strong
Algorithmic Thinking	.785	< .001	Strong
Overall CT Skills	.757	< .001	Strong

Discussion

The overall pattern of findings is broadly consistent with the study's guiding expectation, grounded in Constructivist theory and the Data Literacy and Computational Thinking frameworks, that secondary science teachers who possess stronger data literacy and computational thinking skills would also report stronger integration of these competencies into science instruction. The following sections discuss each set of findings in relation to prior research, consider their practical implications for science teaching and outline directions for further inquiry.

Data Literacy of Secondary Science Teachers

The finding that secondary science teachers demonstrated an overall self-assessed Proficient level of data literacy is broadly consistent with a growing body of literature. Pastore et al. (2025) found that most studies reported moderate to high data literacy levels among in-service teachers who had received structured professional development, a finding that aligns with the present results. Data-based decision making emerged as the highest-rated domain (M = 3.986), resonating with Mandinach and Schildkamp (2021), who argued that when teachers view data as actionable evidence for improving instruction, decision-making naturally becomes their most developed data literacy dimension. This is further supported by Sandoval-Ríos et al. (2025), who highlighted that structured training models emphasizing decision-making significantly produced teachers with the highest proficiency in translating data into instructional choices. In the context of science teaching, where teachers routinely interpret experimental results and student outputs, data-based decision making is arguably the most practiced competency (Amini & Sinaga, 2021).

Data visualization, however, registered the lowest mean (M = 3.887, SD = 0.757). This finding is consistent with Laramee (2023), who emphasized that visualization remains a cognitively complex and pedagogically challenging dimension. Fernandez Nieto et al. (2022) similarly noted that data visualization tasks can exceed the professional skill set of some educators. The relatively higher standard deviation for this domain further suggests greater variability in teacher competency, indicating uneven readiness across the sample.

However, the present results should be interpreted cautiously in light of the self-report methodology. Earlier studies employing performance-based measures have documented lower levels of data literacy among public school science teachers. Sandoval-Ríos et al. (2025) noted a persistent skills gap in data literacy training, particularly in developing-country contexts. The discrepancy between the present self-assessed proficiency and findings from performance-based studies may be attributable to social desirability bias inherent in self-report instruments, the professional development background of the participants or the institutional context within which the respondents operate. Future studies should incorporate objective, performance-based assessments to triangulate self-report findings.

From an implications standpoint, schools and teacher development programs should invest in visualization-specific professional development, particularly in graph literacy, infographic creation and the use of digital tools, to elevate this relatively weaker sub-domain (Wilkerson et al., 2025).

Computational Thinking Skills of Secondary Science Teachers

This finding is notable given the relatively recent introduction of CT into K-12 science education curricula and teacher professional development. The result is broadly supportive of emerging literature showing that teachers who have participated in CT-oriented professional development can reach proficiency across CT dimensions (Mumcu et al., 2023; Yadav et al., 2014). Decomposition registered the highest mean, suggesting that secondary science teachers are most adept at breaking complex science problems into manageable components, a finding that aligns with Mertens and Colunga (2025), who identified decomposition as the most cognitively accessible CT component for novice-to-intermediate learners.

Abstraction yielded the lowest mean (M = 3.745), a finding consistent with a substantial body of literature. Oser et al. (2025) concluded that abstraction remains the least proficiently developed CT skill in K-12 education. Asif et al. (2024) observed that deeper forms of abstraction remain challenging without dedicated professional learning, and Asif et al. (2023) found that teachers often simplify abstraction to merely ignoring details. Rodrigues et al. (2024) corroborated this by finding that abstraction sub-skills showed non-significant improvement in a semester-long CT intervention.

However, these findings partially contrast with Kong et al. (2020) and Rich et al. (2021), who reported that even after targeted professional development, many K-12 teachers continued to demonstrate inconsistent CT levels. The relatively higher self-assessed proficiency in the present study may reflect the specialized science background of the respondents, whose training in scientific modeling and systematic investigation provides a foundation overlapping with several CT dimensions. It is also possible that self-report inflation accounts for part of this discrepancy.

Abstraction and algorithmic thinking, as the lowest-scoring CT dimensions, should be the focus of targeted professional development. Abstraction benefits from structured, problem-based learning contexts where teachers are prompted to identify what is essential and what can be omitted in complex science scenarios. Algorithmic thinking development can be advanced through unplugged activities, flowcharting and lesson-plan design that requires teachers to sequence instructional steps explicitly (Tariq et al., 2025).

Difference Between Data Literacy and Computational Thinking Skills

The effect size (r = 0.28) falls within the small range according to Cohen's (1988) benchmarks (small = 0.10, medium = 0.30, large = 0.50), indicating that while the difference is statistically significant, its practical magnitude is modest. It should be noted that comparing means across two different constructs measured by different instruments carries inherent conceptual limitations. Data literacy and computational thinking are distinct constructs, and a mean difference does not necessarily indicate that teachers are more capable in one than the other in an absolute sense. The comparison is meaningful primarily as a relative indicator of perceived proficiency and as a guide for prioritizing professional development resources.

This finding reflects an asymmetry in the depth of secondary science teachers' professional preparation. Science teachers' professional formation has historically embedded data collection, analysis and interpretation as core competencies within laboratory practices, scientific inquiry cycles and assessment literacy frameworks. Computational thinking, by contrast, is a comparatively newer curricular expectation, and its formal integration into teacher education programs and K-12 science curricula is still emergent (Bocconi et al., 2022; Yadav et al., 2014). This finding is supported by Hsu et al. (2024), who documented that despite growing momentum, there remains a notable absence of consensus on how to assess and build teachers' CT knowledge. The Philippine context, where CT integration into science education is still at an early curriculum stage, may explain the observed gap. This has direct policy implications, as the K-12 curriculum framework in the Philippines could benefit from a more explicit articulation of CT expectations within science learning competencies, supported by targeted CT professional development for in-service teachers.

Integration of Data Literacy and Computational Thinking Skills into Instruction

The finding that secondary science teachers rated their DL-CT integrated science instruction at a Highly Integrated level suggests that these teachers perceive themselves as having moved beyond surface-level familiarity with both skill sets and as embedding both meaningfully into their instructional practice. Their highest-rated capability, integrating science data analysis with real-world issues, reflects a pedagogical orientation aligned with authentic and context-driven science learning. This result is consistent with Coenraad et al. (2022), who found that the most successful integration sites were those connected to authentic data analysis tasks in real-world contexts, and with Moon et al. (2023), who observed stronger instructional coherence among teachers integrating CT through problem-based science tasks.

However, the lowest-rated item, the use of computational tools, warrants careful attention. This result echoes the UNESCO Global Education Monitoring Report (2023), which documented limited digital tool use in science classrooms even in high-income countries. In the Philippine secondary school setting, limited access to computers, software licenses and reliable internet connectivity may constitute structural barriers to computational tool use, a concern highlighted by Kadioglu-Akbulut et al. (2023). Angeli et al. (2016) cautioned that pedagogical integration of computational tools requires dedicated institutional support, including infrastructure provisioning, technical assistance and subject-specific training. The gap between conceptual and tool-mediated integration reinforces Lim et al. (2023), who found that technology-tool integration lagged behind despite strong CT beliefs and pedagogical knowledge.

School administrators and curriculum planners should therefore distinguish between conceptual DL-CT integration, where teachers show strong self-assessed performance, and technology-tool-mediated CT integration, where a meaningful gap persists. Resource allocation, laboratory provisioning and professional development should prioritize the latter.

Relationships Between Competencies and Integrated Instruction

This affirms the theoretical proposition that data literacy is a foundational component of integrated science instruction. This finding is strongly supported by Mandinach and Schildkamp (2021), who argued that data-literate teachers are significantly more likely to design, implement and assess inquiry-based lessons. This pattern is consistent with research on teachers' instructional data use, which has consistently identified analytical and interpretative skills, rather than data collection skills, as the critical bridge between data literacy and instructional action (Filderman et al., 2022; Pastore et al., 2025). The moderate-to-strong correlation for data visualization is consistent with Wilkerson et al. (2025), who found that teachers with higher visualization literacy were more effective in facilitating student-centered data visualization discussions. These findings extend the TPACK-integrated data literacy framework of Cui and Zhang (2022), providing empirical support for the proposition that data literacy is not merely an administrative competency but a pedagogically active one that shapes instructional design.

The overall CT-integration correlation (rs = .757) was marginally higher than the DL-integration correlation (rs = .740). However, it is important to note that the difference of .017 is small, and a formal test for comparing dependent correlations, such as Steiger's Z-test, was not conducted. Therefore, it cannot be concluded with statistical confidence that computational thinking is a significantly stronger correlate of integration than data literacy. The pattern is nonetheless suggestive in nature and warrants further investigation. This observation is broadly consistent with Yadav et al. (2014), who argued that CT, when appropriately adapted for science classrooms, provides the procedural and structural scaffolding through which data literacy practices are operationalized in instruction.

Algorithmic thinking emerged as the single strongest correlate among all CT and DL domains (rs = .785). Algorithmic thinking, defined as the ability to design logical and sequential procedures for solving complex problems, maps directly onto lesson planning, procedural scaffolding and scientific inquiry design. Teachers who think algorithmically may be more inclined to design structured investigations, explicit data analysis protocols and scaffolded argumentation sequences, all of which are hallmarks of integrated DL-CT science instruction (Mumcu et al., 2023).

These results are broadly consistent with the integrated STEM-CT literature reviewed by Tariq et al. (2025), which emphasized that CT integration in science teaching is most successful when teachers deploy all four CT domains in concert rather than in isolation. It should be emphasized, however, that correlation does not establish causation. Future studies employing regression analysis or structural equation modeling would be needed to determine whether computational thinking skills, data literacy or their combination serve as significant predictors of instructional integration.

A further methodological consideration is the potential for common method variance, given that all three variables were assessed through self-report instruments completed by the same respondents at the same time. This shared method may artificially inflate the observed correlations. Future research should employ multi-method designs, such as combining self-report surveys with classroom observation protocols or performance-based assessments, to mitigate this threat.

Conclusions

Secondary science teachers demonstrated self-assessed Proficient levels in both data literacy and computational thinking skills, with all sub-domains falling within the same proficiency band. A statistically significant difference was found between the two skill sets, indicating that teachers perceive data literacy as a slightly stronger area of competence than computational thinking, a pattern reflecting the longer curricular history of data practices in science education relative to the more recently introduced CT competencies. Within data literacy, data-based decision making was the strongest domain, while data visualization was the weakest, with the latter showing greater variability in teachers' graphical and visual data representation skills. Within computational thinking, decomposition was the highest-rated dimension, while abstraction was the lowest, consistent with literature identifying abstraction as the most cognitively demanding and least developed CT skill among K-12 educators. Teachers rated their DL-CT integrated science instruction at a Highly Integrated level, with real-world data contextualization rated highest and use of computational tools rated lowest, pointing to a structural gap between conceptual integration and technology-mediated instruction. Both data literacy and computational thinking showed strong positive correlations with integrated instruction. Algorithmic thinking emerged as the single strongest correlate across all domains. While the CT-integration correlation was marginally higher than the DL-integration correlation, this difference was small and was not formally tested for statistical significance. Both skill sets appear to play substantial and complementary roles in supporting integrated instructional practice.

Limitations

Several limitations of this study should be acknowledged. First, the reliance on self-report instruments introduces the possibility of social desirability bias and response inflation, meaning that teachers may perceive and report higher competence than what objective, performance-based measures would reveal. Second, the modest sample size (n = 53) drawn from a single province through purposive sampling limits the generalizability of the findings. Third, the cross-sectional design precludes causal inference; the observed correlations cannot be interpreted as evidence of causal relationships.

Recommendations

Based on the findings and limitations of this study, the following recommendations are offered. Schools and teacher development programs should design targeted interventions addressing the two weakest sub-domains identified, namely abstraction within computational thinking and data visualization within data literacy. For abstraction, structured problem-based tasks that require teachers to identify essential features, suppress irrelevant detail and build generalizable models should be incorporated into professional learning. For data visualization, training focused on graph literacy, infographic creation and the pedagogical use of digital visualization tools should be embedded into science-specific professional development.

Given the lowest integration rating for computational tool use, school administrators should differentiate between conceptual DL-CT integration, where teachers demonstrate strong self-assessed performance, and technology-tool-mediated CT integration, where a meaningful gap persists. Resource allocation should prioritize hardware provisioning, software licensing and subject-specific training to ensure that science teachers can confidently deploy spreadsheets, graphing software and simulations as computational tools in inquiry-based lessons.

Curriculum planners and teacher educators should treat computational thinking as a co-equal priority alongside data literacy. The Philippine K-12 science curriculum framework should articulate explicit, progressive CT learning competencies covering all four dimensions across grade levels. Pre-service science education programs should incorporate both data literacy and computational thinking as core, assessable content in science methods courses.

Given the strong correlation between algorithmic thinking and integrated science instruction, professional development that builds teachers' capacity to sequence instruction explicitly, design structured investigations and scaffold data analysis is recommended. Unplugged activities, flowcharting and explicit lesson-design training aligned with scientific inquiry processes are suggested entry points.

Finally, future research should address the limitations identified in this study. Longitudinal or experimental designs should be employed to test whether targeted DL and CT professional development produces measurable improvements in instructional quality. Performance-based measures, classroom observation protocols and student outcome data should be incorporated to complement self-report findings. Factor analytic studies of the instruments are recommended to establish construct validity. Larger, probability-based samples from multiple provinces should be used to enhance generalizability. Additionally, regression or structural equation modeling approaches should be employed to determine whether data literacy and computational thinking serve as significant predictors of instructional integration.

Acknowledgment

The authors gratefully acknowledge the secondary science teachers from the Department of Education School Division of Masbate Province, who voluntarily participated and shared their time and insights.

Funding

This research received no external funding.

Ethics Approval

This study did not require review by a formal institutional ethics or review board. As the respondents were public school teachers rather than students or another vulnerable population, and the study posed no more than minimal risk, formal ethics committee approval was not required under the applicable Philippine research regulations. Ethical clearance was instead secured through an official permission letter from the Department of Education Masbate Division Office, and the study adhered to standard ethical safeguards throughout, including an informed consent prior to participation, voluntary participation with the right to withdraw at any time and compliance with Republic Act No. 10173, the Data Privacy Act of 2012 of the Philippines.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

AI Disclosure Statement

The authors declare that artificial intelligence (AI) tools were used in language polishing. All aspects of the research, including data collection, analyses, interpretation and manuscript preparation, were carried out entirely by the authors without the assistance of AI-based technologies.

Conflict of interest

The authors declare no conflict of interest.

References

National Research Council. (2012). *A framework for K-12 science education: Practices, disciplinary core ideas, and crosscutting concepts.* The National Academies Press. https://doi.org/10.17226/13165

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., & Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis report.* Dalhousie University. https://doi.org/10.13140/RG.2.1.1922.5044

Kjelvik, M.K., & Schultheis, E.H. (2019). Getting messy with authentic data: Exploring the potential of using primary scientific data in K-12 education. *CBE—Life Sciences Education, 18*(2), Article es2. https://doi.org/10.1187/cbe.18-11-0220

NGSS Lead States. (2013). *Next generation science standards: For states, by states.* The National Academies Press. https://doi.org/10.17226/18290

Wing, J.M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Kramer, J. (2007). Is abstraction the key to computing? *Communications of the ACM, 50*(4), 36–42. https://doi.org/10.1145/1232743.1232745

Wing, J.M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366*(1881), 3717–3725. https://doi.org/10.1098/rsta.2008.0118

Guzdial, M. (2008). Education paving the way for computational thinking. *Communications of the ACM, 51*(8), 25–27. https://doi.org/10.1145/1378704.1378713

Grover, S., & Pea, R. (2013). Computational thinking in K-12: A review of the state of the field. *Educational Researcher, 42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science education community? *ACM Inroads, 2*(1), 48–54. https://doi.org/10.1145/1929887.1929905

Lee, V.R. (2025). *Advancing data science education in K-12: Foundations, research, and innovations.* Routledge.

Carlson, J., Fosmire, M., Miller, C.C., & Sapp Nelson, M.R. (2011). Determining data information literacy needs: A study of students and research faculty. *Portal: Libraries and the Academy, 11*(2), 629–657. https://doi.org/10.1353/pla.2011.0022

Department of Education. (2016). *K to 12 curriculum guide: Science (Grades 3-10).* Department of Education, Republic of the Philippines.

Calzada Prado, J., & Marzal, M.Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123–134. https://doi.org/10.1515/libri-2013-0010

Pastore, S. (2026). Teacher data literacy: A scoping review (2020–2024). *Cogent Education, 13*(1), Article 2592397. https://doi.org/10.1080/2331186X.2025.2592397

Mandinach, E.B., & Schildkamp, K. (2021). Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational*

Evaluation, 69, Article 100842. https://doi.org/10.1016/j.stueduc.2020.100842

Sandoval-Ríos, F., Gajardo-Poblete, C., & López-Núñez, J.A. (2025). Role of data literacy training for decision-making in teaching practice: A systematic review. *Frontiers in Education,* 10, Article 1485821. https://doi.org/10.3389/feduc.2025.1485821

Bach, B., Keck, M., Rajabiyazdi, F., Losev, T., Meirelles, I., Dykes, J., Laramee, R.S., AlKadi, M., Stoiber, C., Huron, S., Perin, C., Morais, L., Aigner, W., Kosminsky, D., Boucher, M., Knudsen, S., Manataki, A., Aerts, J., Hinrichs, U., Roberts, J.C., & Carpendale, S. (2024). Challenges and opportunities in data visualization education: A call to action. *IEEE Transactions on Visualization and Computer Graphics, 30*(1), 649–660. https://doi.org/10.1109/TVCG.2023.3327378

Fernandez Nieto, G., Kitto, K., Drachsler, H., & Buckingham Shum, S. (2022). Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling. *British Journal of Educational Technology, 53*(2), 249–270. https://doi.org/10.1111/bjet.13165

Wilkerson, M.H., Kim, J., Lee, H.S., Stokes, D.J., & Ferrell, M. (2025). How teachers envision using data visualization discussion tasks in classroom instruction. *International Journal of Science and Mathematics Education, 23*(7), 2653–2687. https://doi.org/10.1007/s10763-024-10521-y

Oser, M., Geyer, C., & Gruber, H. (2025). Teaching and assessing abstraction in K-12 computational thinking education: A systematic literature review. *Computer Applications in Engineering Education, 33*(1), Article e70073. https://doi.org/10.1002/cae.70073

Asif, A.D., Malik, H., Orrill, C.H., Balasubramanian, R., & Kayumova, S. (2024). Computational thinking: Teachers' practice of abstraction, Part II. In Proceedings of the Society for Information Technology & Teacher Education International Conference (pp. 1234–1240). Association for the Advancement of Computing in Education (AACE).

Asif, A.D., Malik, H., Orrill, C.H., Balasubramanian, R., & Kayumova, S. (2023). *Computational thinking: Teachers' practice of abstraction (NSF-funded research report).* National Science Foundation. https://par.nsf.gov/

Mertens, A.J., & Colunga, E. (2025). Assessing cognitive components of computational thinking. *Frontiers in Psychology,* 16, Article 1434453. https://doi.org/10.3389/fpsyg.2025.1434453

Kong, S.C., Lai, M., & Sun, D. (2020). Teacher development in computational thinking: Design and learning outcomes of programming concepts, practices and pedagogy. *Computers & Education,* 151, Article 103872. https://doi.org/10.1016/j.compedu.2020.103872

Rich, K.M., Yadav, A., & Larimore, R.A. (2021). Teacher implementation profiles for integrating computational thinking into elementary mathematics and science instruction. *Education and Information Technologies, 26*(4), 3169–3196. https://doi.org/10.1007/s10639-020-10424-5

Hsu, T.C., Chang, C., & Hung, Y.T. (2024). Supporting K-12 teachers in integrating computational thinking into classrooms: A systematic review. *International Journal of STEM Education,* 11, Article 47. https://doi.org/10.1186/s40594-024-00510-6

Coenraad, M., Cabrera, L., Killen, H., Plane, J., & Ketelhut, D.J. (2022). Computational thinking integration in elementary teachers' science lesson plans. *In Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1 (pp. 11–18). Association for Computing Machinery.* https://doi.org/10.1145/3507951.3519283

Moon, P.F., Himmelsbach, J., Weintrop, D., & Walkoe, J. (2023). Developing preservice teachers' intuitions about computational thinking in a mathematics and science methods course. *Journal of Pedagogical Research, 7*(2), 15–32. https://doi.org/10.33902/JPR.202318636

UNESCO. (2023). *Technology in education: A tool on whose terms?* Global education monitoring report 2023. https://gem-report-2023.unesco.org/technology-in-education/

Angeli, C., Voogt, J., Fluck, A., Webb, M., Cox, M., Malyn-Smith, J., & Zagami, J. (2016). A K-6 computational thinking curriculum framework: Implications for teacher knowledge. *Educational Technology & Society, 19*(3), 47–57.

Kadioglu-Akbulut, C., Cetin-Dindar, A., Acar-Sesen, B., & Küçük, S. (2023). Predicting preservice science teachers' TPACK through ICT usage. *Education and Information Technologies, 28*(9), 11269–11289. https://doi.org/10.1007/s10639-023-11657-0

Lim, B., Lake, V., Beisly, A., & Ross-Lightfoot, R. (2023). Preservice teachers' TPACK growth after technology integration courses in early childhood education. *Early Education and Development, 35*(1), 114–131. https://doi.org/10.1080/10409289.2023.2224219

Cui, Y., & Zhang, H. (2022). Integrating teacher data literacy with TPACK: A self-report study based on a novel framework for teachers' professional development. *Frontiers in Psychology,* 13, Article 966575. https://doi.org/10.3389/fpsyg.2022.966575

Piaget, J. (1952). *The origins of intelligence in children.* International Universities Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Etikan, I., Musa, S.A., & Alkassim, R.S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Creswell, J.W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Wild, C.J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference* (16th ed.). Routledge.

Amini, S., & Sinaga, P. (2021). Inventory of scientific literacy ability of junior high school students based on the evaluation of PISA framework competency criteria. *Journal of Physics: Conference Series, 1806*(1), Article 012017. https://doi.org/10.1088/1742-6596/1806/1/012017

Mumcu, F., Uslu, N.A., & Yildiz, B. (2023). Teacher development in integrated STEM education: Design of lesson plans through the lens of computational thinking. Education and Information Technologies, 28(3), 3443–3474. https://doi.org/10.1007/s10639-022-11342-8

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J.T. (2014). Computational thinking in elementary and secondary teacher education. *ACM Transactions on Computing Education, 14*(1), 1–16. https://doi.org/10.1145/2576872

Rodrigues, R.N., Costa, C., & Martins, F. (2024). Integration of computational thinking in initial teacher training for primary schools: A systematic review. *Frontiers in Education,* 9, Article 1330065. https://doi.org/10.3389/feduc.2024.1330065

Tariq, R., Aponte Babines, B.M., Ramirez, J., Alvarez-Icaza, I., & Naseer, F. (2025). Computational thinking in STEM education: Current state-of-the-art and future research directions. *Frontiers in Computer Science,* 6, Article 1480404. https://doi.org/10.3389/fcomp.2024.1480404

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Bocconi, S., Chioccariello, A., Kampylis, P., Dagienė, V., Wastiau, P., Engelhardt, K., Earp, J., Horvath, M.A., Jasutė, E., Malagoli, C., Masiulionytė-Dagienė, V., & Stupurienė, G. (2022). *Reviewing computational thinking in compulsory education.* Publications Office of the European Union. https://doi.org/10.2760/126955

Filderman, M.J., Toste, J.R., Didion, L., & Peng, P. (2022). Data literacy training for K–12 teachers: A meta-analysis of the effects on teacher outcomes. *Remedial and Special Education, 43*(5), 328–343. https://doi.org/10.1177/07419325211054208

Грамотність у роботі з даними та обчислювальне мислення вчителів природничих дисциплін середньої школи: кількісне дослідження рівнів, відмінностей та інтеграції у навчальний процес

Джон Карло Мортел[1ABCD], Ерді Кахурао[2ACD]

[1]Меморіальна середня школа імені Федеріко А. Естіпони
[2]Педагогічний коледж, Меморіальний державний коледж сільського господарства і технологій імені доктора Еміліо Б. Еспіноси-старшого

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 12 с., 6 табл., 3 рис., 38 джерел

Мета. У дослідженні вивчено рівні грамотності у роботі з даними (DL) та навичок обчислювального мислення (CT) серед учителів природничих дисциплін середньої школи, ступінь їх інтеграції у викладання природничих дисциплін, а також взаємозв'язки між цими компетентностями та їх інтеграцією у навчальний процес.

Матеріали та методи. У дослідженні описово-кореляційного дизайну взяли участь 53 учителі природничих дисциплін середніх шкіл провінції Масбате, Філіппіни. Учасники заповнили три розроблені авторами інструменти типу шкали Лайкерта для оцінювання DL, CT та інтеграції DL-CT у викладання природничих дисциплін. Дані аналізували за допомогою середніх значень і стандартних відхилень, критерію знакових рангів Вілкоксона та рангової кореляції Спірмена.

Результати. Учителі оцінили свій рівень як компетентний щодо DL ($M = 3{,}92$; $SD = 0{,}63$) і CT ($M = 3{,}84$; $SD = 0{,}64$); різниця між цими показниками була статистично значущою ($W = 288{,}00$; $p = 0{,}040$; $r = 0{,}28$). Учителі повідомили про високий рівень інтеграції DL-CT у навчальний процес ($M = 3{,}90$; $SD = 0{,}82$). Кореляційний аналіз Спірмена виявив сильні позитивні зв'язки інтеграції як із DL ($r_s = 0{,}740$), так і з CT ($r_s = 0{,}757$), в обох випадках $p < 0{,}001$. Алгоритмічне мислення мало найсильніший окремий зв'язок з інтеграцією ($r_s = 0{,}785$).

Висновки. Учителі природничих дисциплін середньої школи продемонстрували компетентний, але неоднорідний рівень грамотності у роботі з даними та обчислювального мислення, при цьому відносно слабшими аспектами були абстрагування та доступ до обчислювальних інструментів. Обидві компетентності відіграють суттєву та взаємодоповнювальну роль в інтегрованому викладанні природничих дисциплін. Рекомендовано цільове професійне вдосконалення, спрямоване на розвиток виявлених слабших аспектів.

Ключові слова: грамотність у роботі з даними, обчислювальне мислення, природнича освіта, інтеграція у навчальний процес, учителі середньої школи.

Information about the Authors

Mortel, John Carlo: johncarlomortel17@gmail.com; https://orcid.org/0009-0002-7491-9063; Federico A. Estipona Memorial High School, Department of Education, Masbate, Philippines.

Cajurao, Erdee: eccajurao@debesmscat.edu.ph; https://orcid.org/0000-0001-5695-2850; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology, Mandaon, Masbate, Philippines.

Cite this article as: Mortel, J. C., & Cajurao, E. (2026). Data Literacy and Computational Thinking Among Secondary Science Teachers: A Quantitative Study of Levels, Differences and Instructional Integration. *Journal of Learning Theory and Methodology, 7*(2), 114-125. https://doi.org/10.17309/jltm.2026.7.2.06

Received: 03.07.2026. Accepted: 13.08.2026. Published: 17.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання
ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Original Scientific Article

When Usefulness Outweighs Norms: Predictors of Generative Artificial Intelligence Use in a Mobile-First Teacher Education Context

RP Reel Acosta[1ABCD], Liza Angel Almanzor[1ABCD], Irish Cierto[1ABCD], Arisa Mae Toledo[1ABCD] and Erdee Cajurao[1ACD]

[1]College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.07

Abstract

Background. Generative artificial intelligence (GenAI) tools are increasingly integrated into higher education, yet the factors driving their adoption among pre-service teachers remain underexplored in Southeast Asian contexts.

Objectives. This study investigated predictors of generative AI use among 118 pre-service science teachers at a state college in Masbate, Philippines. Drawing on the Technology Acceptance Model, the Unified Theory of Acceptance and Use of Technology, and the Theory of Planned Behavior, five constructs were examined: Perceived Usefulness, Attitude Toward AI, Effort Expectancy, Facilitating Conditions, and Subjective Norms.

Materials and Method. A quantitative descriptive-predictive design with stratified random sampling was used. Data were collected using a validated 32-item instrument and analyzed using four multiple linear regression models.

Results. Generative AI use was rated high ($M = 3.67$, $SD = 0.72$). Perceived Usefulness ($M = 3.79$), Attitude Toward AI ($M = 3.71$), and Effort Expectancy ($M = 3.75$) were rated high, while Facilitating Conditions ($M = 3.44$) and Subjective Norms ($M = 3.35$) were rated moderate. Across all four models ($F = 50.47$ to 67.14, $p < .001$), Perceived Usefulness was the sole predictor consistently significant at $p < .001$ ($\beta = 0.443$ to 0.479), explaining 63.9 to 64.6% of the variance ($R^2 = 0.639$ to 0.646). Attitude Toward AI and Effort Expectancy reached significance in selected model configurations. Multicollinearity was within acceptable limits ($VIF = 1.40$ to 4.76).

Conclusion. These correlational findings, drawn from a single institutional sample, point to the primacy of perceived usefulness in technology adoption. As a hypothesis for future confirmation, teacher education programs may benefit from demonstrating the practical value of generative AI in science instruction.

Keywords: generative artificial intelligence, perceived usefulness, pre-service science teachers, science education, technology acceptance.

Introduction

Artificial intelligence has begun reshaping educational systems at a rate that few observers anticipated even a decade ago. Recent bibliometric and narrative reviews confirm the accelerating pace of this integration across educational levels and disciplines (Limna et al., 2022; Zhang & Aslan, 2021). The emergence of generative AI tools, defined broadly as systems capable of producing original text, images, code and other content in response to user prompts, has introduced a new category of cognitive technology into the learning environment (Crompton & Burke, 2023; UNESCO, 2023). Unlike earlier AI applications that functioned primarily as passive recommenders or adaptive content distributors, generative AI enables learners and instructors to engage in open-ended, conversational interactions that can support writing, reasoning, research synthesis and problem-solving at scale (Su & Yang, 2023; Zhao et al., 2026). Empirical evidence for these instructional benefits has begun to accumulate: AI-enabled personalized recommendations have improved learner engagement and performance in flipped-classroom settings (Huang et al., 2023), and adaptive natural-language-processing tools have improved the quality of student argumentative writing (Wambsganss et al., 2020). The rapid adoption of platforms such as ChatGPT has placed a central question at the forefront of educational research, specifically what determines whether a learner chooses to use these tools, and how consistently.

Corresponding Author: Erdee Cajurao,
e-mail: eccajurao@debesmscat.edu.ph

Pre-service teachers occupy a particularly consequential position in this context. As individuals who are simultaneously learners in higher education and future professionals who will shape how the next generation relates to technology, their dispositions toward AI carry implications that extend well beyond their own academic performance. A pre-service teacher who understands and actively uses generative AI is far more likely to integrate it meaningfully into classroom practice after graduation than one who merely encounters it as a peripheral option (Mishra & Koehler, 2006; Sun et al., 2025; Wang et al., 2025). This upstream influence extends to ethical and literacy dimensions as well: interventions that build ethics-aware AI literacy during teacher preparation may shape how future K-12 students themselves come to use these tools responsibly (Zhang et al., 2023). This upstream influence on educational technology integration makes pre-service teacher populations a high-value subject group for adoption research.

Within the science education context, generative AI presents both distinctive opportunities and genuine tensions. Science inquiry depends on evidence evaluation, claim-warrant reasoning, and iterative model revision, competencies that can in principle be scaffolded by conversational AI, yet also potentially short-circuited by it if students accept AI-generated explanations uncritically (Holmes et al., 2019; Luan et al., 2020; Ouyang & Jiao, 2021). Realizing the benefits of generative AI without inviting this risk depends on students and instructors developing the digital competencies needed to use these tools critically and effectively (Kohnke et al., 2023). Understanding what drives science pre-service teachers to use or avoid these tools is therefore not simply a question of technology access or preference; it is a curricular and professional development concern with long-term consequences for science teaching quality.

Despite growing interest in AI adoption among university students broadly, empirical studies focused specifically on pre-service science teachers in the Philippines remain scarce. The country presents a context of particular interest because it combines high rates of smartphone penetration and social media engagement with persistent infrastructure disparities in internet reliability, especially in rural and island provinces (Ittefaq et al., 2025; Maheshwari, 2024; He & Ren, 2025). Pre-service teachers trained in state universities in these regions may encounter generative AI through mobile-first access patterns that differ substantially from those observed in studies conducted in urban or high-income settings. Theories developed and tested primarily in North American, East Asian, or Western European contexts may not transfer directly to this population without empirical verification.

The present study is grounded in three established frameworks for understanding technology adoption. The Technology Acceptance Model, first articulated by Davis (1989), identifies Perceived Usefulness and perceived ease of use as the two primary determinants of individual technology acceptance. Perceived Usefulness captures the degree to which a person believes that using a particular system will enhance their performance, and it has consistently emerged as the stronger of the two predictors across prior empirical tests (Haq et al., 2025; Venkatesh et al., 2003). The Unified Theory of Acceptance and Use of Technology, proposed by Venkatesh et al. (2003), extended this framework by incorporating social and organizational variables, including Subjective Norms, Facilitating Conditions, and Effort Expectancy, the latter conceptually aligned with perceived ease of use. The Theory of Planned Behavior (Ajzen, 1991; Fishbein & Ajzen, 1975) complements both models: it holds that behavioral intentions are jointly shaped by attitudes, subjective social norms, and perceived behavioral control. This tripartite structure aligns well with the multi-construct approach adopted in the present study.

The Diffusion of Innovations framework (Rogers, 2003) adds a further theoretical lens: it positions technology adoption as a social process in which the perceived relative advantage of an innovation, an analog of Perceived Usefulness, is among the strongest predictors of adoption speed and breadth. From a pedagogical standpoint, the Technological Pedagogical Content Knowledge framework (Mishra & Koehler, 2006) draws attention to the integrative competency teachers must develop to use technology not merely as a convenience but as a purposeful instructional resource. Together, these frameworks motivate the five-predictor model tested in this study. Acceptance constructs such as Perceived Usefulness capture a learner's subjective judgment about a tool, not a direct measure of instructional benefit. This study therefore also draws on Cognitive Load Theory (Sweller, 1988) and self-regulated learning frameworks (Zimmerman, 2002) to interpret what a high usefulness rating plausibly reflects about how generative AI is used in practice, for example as a means of offloading extraneous cognitive demand or supporting planning and self-monitoring during science coursework. This learning-theoretic lens is intended to complement rather than substitute for the acceptance-theoretic model that anchors the present analysis; its implications are established further in the Discussion.

Despite the theoretical richness of this literature, several empirical gaps remain. This pattern echoes a broader gap identified across the AIEd literature, where technological application has often outpaced integration with educational theory (Chen et al., 2020). First, most studies examining AI adoption among pre-service teachers have treated technology acceptance as a unidimensional or bivariate outcome, rarely comparing the relative predictive strength of multiple constructs across competing regression models (Suson, 2025; Wu et al., 2025). Second, the specific role of social influence through Subjective Norms in collectivist educational cultures such as the Philippines has received limited attention in the generative AI literature. Third, infrastructure constraints such as internet reliability and device type, which may moderate the relationship between facilitating conditions and actual use, have seldom been reported alongside behavioral outcome data (Haq et al., 2025; He & Ren, 2025).

A further complication in the Philippine context is the uneven distribution of digital literacy competencies across teacher education programs. Students entering pre-service programs from diverse socioeconomic backgrounds may have encountered technology-mediated instruction at dramatically different rates in secondary school, producing cohorts with varying degrees of prior exposure to digital tools. Prior exposure has been shown to moderate technology acceptance in prior studies by shaping baseline Perceived Usefulness and ease-of-use expectations before formal instruction begins (Bandura, 1977; Rogers, 2003). Understanding whether and how this heterogeneous prior exposure interacts with in-program AI experiences is a question that cannot be answered without first establishing the baseline predictors of current use behavior.

The present study also responds to a methodological gap in the existing literature. Most single-study examinations of AI adoption predictors test one or two constructs rather than systematically comparing multiple predictor combinations within the same sample. This study instead estimates four competing multiple regression models that vary in their inclusion of the five predictor variables. This model-comparison approach identifies which predictors are significant in isolation and whether that significance holds when competing variables are partialled out. The within-sample comparison reduces the risk of over-attributing predictive power to any single construct and provides a more conservative test of theoretical claims.

This study addresses these gaps by examining five predictor variables across four competing multiple regression models in a sample of 118 pre-service science teachers drawn from a rural Philippine state university. The research questions guiding this inquiry are as follows. First, what is the technology profile of pre-service science teachers in terms of internet reliability, devices used, and generative AI tools used? Second, how do pre-service science teachers rate the five predictor variables and their generative AI use? Third, which predictors significantly explain variance in generative AI use behavior?

The contribution of this study is best understood as a boundary-condition test rather than the introduction of a new predictor construct or theoretical model. TAM, UTAUT and TPB relationships have been established predominantly in North American, East Asian and urban high-income settings. This study examines whether those relationships hold, weaken or fail in a population and infrastructure context where they have rarely been tested jointly: mobile-first, connectivity-constrained pre-service science teachers in a rural Philippine state college. Culture and infrastructure have previously been identified as boundary conditions that can alter which UTAUT constructs retain predictive power (Bayaga & du Plessis, 2024; Venkatesh & Zhang, 2010), and the present multi-model comparative design extends this boundary-condition logic to generative AI specifically. As detailed in the Discussion, Facilitating Conditions and Subjective Norms failed to predict use in any of the four models, despite the Philippines' collectivist social structure and the theoretical importance ordinarily attributed to these constructs. This null result is itself a substantive empirical contribution: it suggests that in mobile-first, infrastructure-constrained contexts, individual utility judgments may outweigh social and institutional influences to a degree not anticipated by UTAUT's original specification. The finding qualifies established technology acceptance theory rather than simply confirming it.

Materials and Methods

Research Design

This study employed a quantitative descriptive-predictive design. The descriptive component characterized the technology profile and mean scores of study variables, while the predictive component used multiple linear regression to identify the constructs that significantly explained variance in generative AI use. This design is consistent with recent technology adoption research in higher education settings (Haq et al., 2025; Suson, 2025; Wu et al., 2025) and was selected because it enables systematic comparison of competing predictor models without requiring experimental manipulation.

Participants and Sampling

The population consisted of all pre-service science teachers enrolled at a state college in the Philippines during the academic year 2024 to 2025. A total of n = 118 participants were recruited using stratified random sampling, with year level serving as the stratification variable. This approach ensured proportional representation across all four undergraduate year levels. Inclusion criteria required that participants (a) were officially enrolled in the Bachelor of Secondary Education major in Science program, (b) had prior experience using at least one digital device for academic purposes, and (c) provided informed consent. Participation was voluntary and no incentives were offered.

Instrument

The survey instrument contained 32 items distributed across six constructs, namely Generative AI Use (dependent variable), Perceived Usefulness, Attitude Toward AI, Subjective Norms, Facilitating Conditions, and Effort Expectancy. Items were rated on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The instrument was adapted from validated scales used in prior technology acceptance research (Cabero-Almenara et al., 2024; Venkatesh et al., 2003) and underwent content validation by a panel of five experts with backgrounds in educational technology and science education. Internal consistency for the full instrument was Cronbach alpha = 0.797, which falls within the acceptable range for social science research.

Mean scores were interpreted using a five-level rubric in which 4.21 to 5.00 was classified as Very High, 3.51 to 4.20 as High, 2.51 to 3.50 as Moderate, 1.81 to 2.50 as Low, and 1.00 to 1.80 as Very Low. This descriptive scale, commonly applied in Philippine higher education research, provided a standardized and interpretively consistent basis for comparing mean scores across constructs and relating them to prior work that used similar scales.

Data Collection

Data were collected through a self-administered Google Form distributed to participants during scheduled class sessions in February to March 2025. Section one gathered technology profile information including internet reliability, primary device type, and generative AI tools used. Section two contained the 32 Likert-scale items. Completion required approximately 15 to 20 minutes. All responses were anonymized before analysis.

Data Analysis

Descriptive statistics including means, standard deviations, and frequency distributions were computed for all variables. Four multiple linear regression models were estimated to examine different predictor combinations and assess the robustness of each variable's contribution. Model 1 included PU, ATT, FC and EE. Model 2 included PU, ATT, SN and FC. Model 3 included PU, ATT, SN and EE. Model 4 included PU, ATT and EE. This model comparison approach allowed the identification of which predictors retained significance when competing variables were held constant.

Regression assumptions were tested prior to interpretation. Normality of residuals was assessed using the Shapiro-Wilk test ($W = 0.980$, $p = .076$), and homoscedasticity was confirmed via the Breusch-Pagan test ($BP = 1.910$, $df = 2$, $p = .385$). Multicollinearity was evaluated using Variance Inflation Factor values and tolerance statistics. All analyses were conducted using R 4.3 and SPSS 26. Statistical significance was set at alpha = .05.

Results

Technology Profile of Respondents

In terms of internet reliability, approximately half of the respondents (51.7%) rated their connection as moderate, while 40.7% reported good reliability. Very poor and very good ratings were minimal (0.0% and 1.7%, respectively), and only 5.9% reported poor connectivity. These data suggest that most respondents operate within a connectivity environment that is functional but not uniformly robust.

With respect to primary devices, the overwhelming majority of respondents (88.1%) used smartphones as their main access point for digital academic activities, compared to only 11.9% who relied on laptops or desktop computers. This mobile-first profile is consistent with patterns observed in other low-to-middle income country settings and has implications for how generative AI tools must function to be practically usable by this population. Table 1 details the technology profile of the respondents.

Regarding generative AI tool use, ChatGPT was by far the most frequently used platform (86.4%), followed by Gemini (38.1%), Meta AI (32.2%), QuillBot (28.0%) and Grammarly (26.3%). The dominance of ChatGPT mirrors findings in broader Southeast Asian student populations and reflects the platform's early market saturation and interface familiarity (Maheshwari, 2024).

Table 1. Technology Profile of Pre-Service Science Teachers (n = 118)

Variable	Category	n	%
Internet Reliability	Very Poor	0	0.0
	Poor	7	5.9
	Moderate	61	51.7
	Good	48	40.7
	Very Good	2	1.7
Primary Device Used	Smartphone	104	88.1
	Laptop/Desktop	14	11.9
Generative AI Tools Used (Multiple Response)	ChatGPT	102	86.4
	Gemini	45	38.1
	Meta AI	38	32.2
	QuillBot	33	28.0
	Grammarly	31	26.3

Descriptive statistics of study variables

Table 2 presents the mean scores and standard deviations for all six study variables. Perceived Usefulness obtained the highest mean score ($M = 3.788$, $SD = 0.702$), followed by Effort Expectancy ($M = 3.746$, $SD = 0.636$), Attitude Toward AI ($M = 3.710$, $SD = 0.654$), and Generative AI Use itself ($M = 3.669$, $SD = 0.717$). All four of these variables fell within the High interpretation range (3.51 to 4.20). Facilitating Conditions ($M = 3.444$, $SD = 0.635$) and Subjective Norms ($M = 3.347$, $SD = 0.492$) were rated in the Moderate range, indicating that peer and institutional influences were perceived less strongly than internal utility and ease evaluations. Figure 1 illustrates the mean scores and variability across all six constructs.

Table 2. Descriptive Statistics of Study Variables (n = 118)

Variable	Mean	SD	Interpretation
Generative AI Use	3.669	0.717	High
Perceived Usefulness (PU)	3.788	0.702	High
Attitude Toward AI (ATT)	3.710	0.654	High
Subjective Norms (SN)	3.347	0.492	Moderate
Facilitating Conditions (FC)	3.444	0.635	Moderate
Effort Expectancy (EE)	3.746	0.636	High

Multiple linear regression results

Table 3 presents the standardized regression coefficients across four models. Perceived Usefulness consistently emerged as the only predictor significant at $p < .001$ in every model, with beta coefficients ranging from 0.443 in

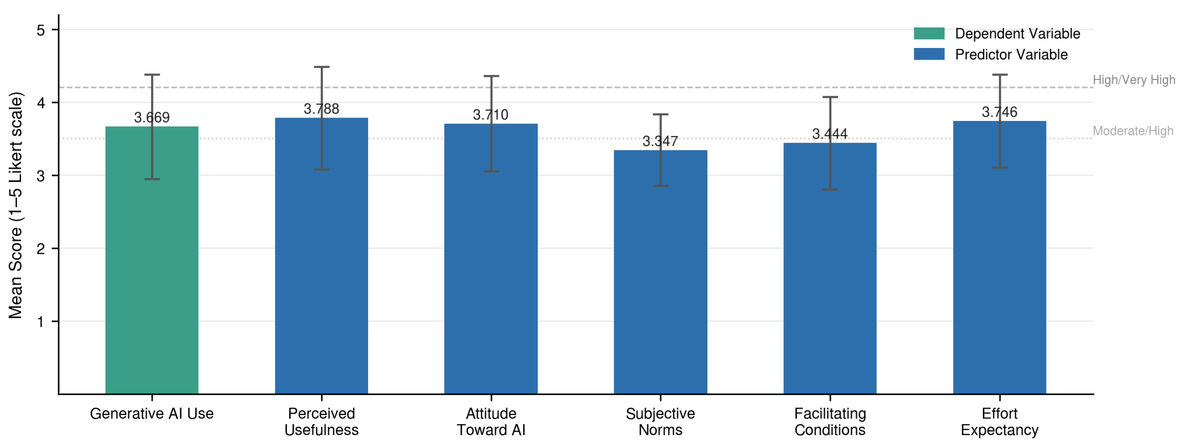


Fig. 1. Mean scores and standard deviations of study variables.

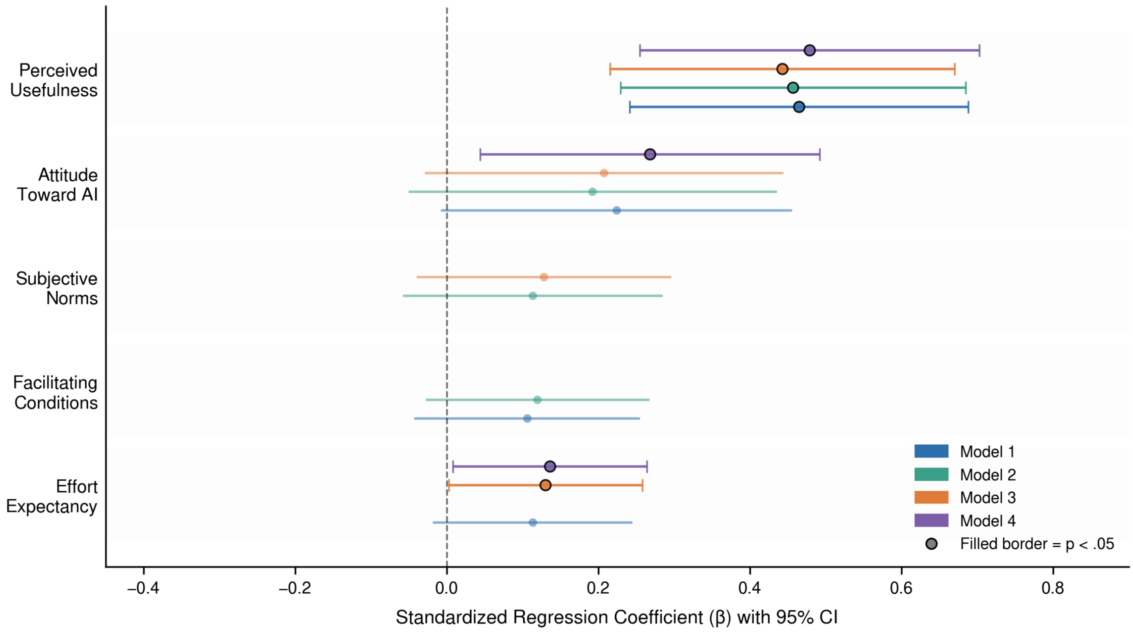


Fig. 2. Forest plot of standardized regression coefficients (β).

Table 3. Standardized Regression Coefficients Across Four Models (n = 118)

Predictor	Model 1		Model 2		Model 3		Model 4	
	β	p	β	p	β	p	β	p
PU	0.4650	<.001	0.4571	<.001	0.4431	<.001	0.4789	<.001
ATT	0.2240	.058	0.1925	.119	0.2075	.085	0.2684	.019
SN	—	—	0.1137	.192	0.1285	.133	—	—
FC	0.1061	.161	0.1199	.111	—	—	—	—
EE	0.1134	.090	—	—	0.1304	.045	0.1361	.037
R2	0.645		0.641		0.646		0.639	
Adjusted R2	0.632		0.628		0.633		0.629	
F	51.29	<.001	50.47	<.001	51.50	<.001	67.14	<.001

Note: Dashes indicate that the predictor was not included in that model. PU = Perceived Usefulness; ATT = Attitude Toward AI; SN = Subjective Norms; FC = Facilitating Conditions; EE = Effort Expectancy.

Model 3 to 0.479 in Model 4. The total variance explained by the models was similar, with R-squared values of 0.645, 0.641, 0.646, and 0.639 for Models 1 through 4, respectively, indicating that each model accounted for approximately 64 percent of variance in generative AI use.

Attitude Toward AI produced positive beta coefficients across all models, ranging from 0.193 to 0.268, reaching statistical significance in Model 4 (β = 0.268, $p = .019$). Effort Expectancy achieved statistical significance in Model 3 (β = 0.130, $p = .045$) and Model 4 (β = 0.136, $p = .037$), while its coefficients in Models 1 and 2 fell short of the significance threshold. Subjective Norms (Models 2 and 3) and Facilitating Conditions (Models 1 and 2) were non-significant in all configurations in which they were included.

Figure 2 presents a forest plot of the standardized regression coefficients with 95% confidence intervals for each predictor, plotted separately for all four models. The dashed vertical reference line at zero represents the null hypothesis of no effect. Filled-border markers denote statistically significant estimates ($p < .05$). The confidence intervals for Perceived Usefulness do not cross zero in any model, confirming the reliability and precision of its effect. By contrast, the intervals for Attitude Toward AI, Subjective Norms, Facilitating Conditions, and Effort Expectancy all straddle zero or rest near it, indicating that their apparent positive effects carry substantial uncertainty. The forest plot makes the contrast in estimation precision visually explicit and reinforces the conclusion that the PU effect is qualitatively different in nature from the effects of the other four predictors.

Regression assumption diagnostics

Table 4 presents multicollinearity diagnostics. Tolerance values ranged from 0.210 (Attitude Toward AI) to 0.713 (Effort Expectancy), and VIF values ranged from 1.40 to 4.76, all well below the conventional cutoffs of 0.10 and 10.0 respectively. The highest VIF was observed for Attitude Toward AI (4.76), reflecting its strong theoretical and empirical overlap with Perceived Usefulness ($r = 0.864$). Perceived Usefulness itself had a VIF of 4.21. Nevertheless, all values remained within acceptable bounds and do not indicate problematic multicollinearity.

Table 4. Multicollinearity Diagnostics Across Predictor Variables

Predictor	Tolerance	VIF	Min VIF	Max VIF
Perceived Usefulness (PU)	0.237	4.21		
Attitude Toward AI (ATT)	0.210	4.76		
Subjective Norms (SN)	0.423	2.36		
Facilitating Conditions (FC)	0.539	1.85		
Effort Expectancy (EE)	0.713	1.40		
Range across models			1.40	4.76

Note: VIF = Variance Inflation Factor. All VIF values are below the conventional threshold of 10.0, and all tolerance values are above 0.10.

Residual normality was confirmed by the Shapiro-Wilk test (W = 0.980, p = .076), and homoscedasticity was supported by the Breusch-Pagan test (BP = 1.910, df = 2, p = .385). These diagnostics collectively indicate that the regression models meet standard assumptions and that the coefficient estimates are reliable.

Discussion

Technology Access and Tool Preferences

The finding that nearly nine in ten respondents relied primarily on smartphones to access digital tools reflects the broader mobile-first orientation of students in provincial Philippine higher education. While smartphone ownership is widespread in this context, the moderate internet reliability ratings reported by more than half of the respondents introduce practical constraints on tool use. Generative AI applications, particularly text-based interfaces like ChatGPT, are lightweight in bandwidth requirements compared with video- or image-heavy platforms. This may explain why these students were still able to adopt and use such tools regularly despite connectivity limitations. This observation is consistent with Maheshwari (2024), who noted that Southeast Asian students adapt their technology use pragmatically to available infrastructure conditions.

The dominance of ChatGPT (86.4%) over other platforms points to the role of early adoption network effects and platform visibility. Students who begin using a tool because peers or instructors mention it are likely to recommend it in turn, creating a cascade that concentrates use around one or two dominant platforms even when alternatives exist. Grammarly (26.3%) and QuillBot (28.0%) saw lower uptake. Both function primarily as writing-assistance tools rather than conversational AI systems, which suggests that respondents prioritized interactive reasoning support over grammar correction when choosing among generative AI platforms.

Perceived Usefulness and Favorable Attitudes

The high mean scores for Perceived Usefulness (M = 3.788) and Attitude Toward AI (M = 3.710) suggest that the respondents entered the study with broadly positive dispositions toward generative AI. These findings extend a pattern observed in comparable studies. Haq et al. (2025) found that university students in higher education settings across multiple countries similarly rated AI usefulness positively, while Cabero-Almenara et al. (2024) reported favorable AI acceptance scores among Spanish-speaking university students using the AI4U scale. Maheshwari (2024) identified positive attitudes as a significant antecedent to ChatGPT behavioral intention among Vietnamese students, reinforcing the cross-contextual generalizability of this finding within Southeast Asia.

The alignment between high usefulness perceptions and high use behavior in the present sample supports the core prediction of the Technology Acceptance Model (Davis, 1989), that individuals adopt technologies they perceive as useful for their work or learning. For pre-service science teachers specifically, usefulness is likely tied to concrete applications such as simplifying complex scientific concepts for lesson planning, generating practice problems, drafting explanatory texts for diverse learners, and supporting research writing. These are all high-frequency tasks in teacher education programs that generative AI can plausibly reduce the cognitive load of completing. Recognizing this, teacher educators who explicitly frame AI tools around these discipline-specific applications may find that usefulness perceptions, and with them adoption rates, respond favorably.

What the present data cannot resolve is whether the high PU scores in this sample reflect well-calibrated perceptions grounded in actual productive use, or optimistic anticipatory beliefs that have not yet been tested against the real limitations of generative AI in science educational contexts. Distinguishing calibrated from aspirational usefulness perceptions would require a longitudinal design that tracks PU scores alongside objective indicators of AI-assisted task quality over an academic term. This distinction matters because interventions designed to sustain rather than merely initiate adoption would differ, and would need to ensure that early enthusiasm is supported by structured experiences in which AI tools demonstrably deliver on the usefulness they are perceived to offer. An unmet usefulness expectation, if it accumulates, could depress PU scores and adoption rates in later cohorts, creating a trough in the adoption curve following the initial enthusiasm wave characteristic of emerging technology diffusion (Rogers, 2003).

Moderate Social Influence and Institutional Support

The moderate ratings for Facilitating Conditions (M = 3.444) and Subjective Norms (M = 3.347) deserve careful interpretation. In collectivist societies such as the Philippines, group conformity and social referents typically exert strong behavioral influence (Ajzen, 1991). Here, however, peer and institutional normative pressure yielded only moderate scores. This suggests that the social milieu around generative AI at this institution has not yet crystallized into a clear descriptive norm. Peer use of AI tools may not yet be visible enough, or sufficiently encouraged by faculty, to register as a potent social force. This interpretation is consistent with He and Ren (2025), who found that social influence on AI tool adoption among university students was weaker in samples characterized by lower institutional AI policy clarity.

The moderate Facilitating Conditions score likewise points to infrastructural and institutional gaps. The respondents could access generative AI tools on their smartphones, but the absence of formal training programs, institutional guidelines on academic integrity in AI use, guidelines that would ideally be grounded in established ethical frameworks for AIED (Nguyen et al., 2023), and faculty-modeled integration of AI into science pedagogy likely constrained their sense that conditions were fully in place to support that use. Selwyn (2022) cautioned that access alone does not generate productive use and that students need scaffolded environments that show how AI tools connect to learning goals. The present data suggest that the institution, like many regional state universities, has yet to fully provide that scaffolding at the institutional level.

Perceived Usefulness as the Dominant Predictor

The most important finding of this study is the consistent significance of Perceived Usefulness across all four regression models at $p < .001$. With beta coefficients ranging from 0.443 to 0.479 and significance sustained regardless of which competing variables were included, PU demonstrated a degree of predictive robustness that was not matched by any other variable. Attitude Toward AI and Effort Expectancy achieved significance in select model configurations, suggesting secondary roles that are condition-dependent rather than universal. This pattern is theoretically coherent and carries direct practical implications.

Theoretically, the finding replicates and extends the central prediction of Davis (1989) into a generative AI context. Earlier TAM applications in education already established that usefulness perceptions predict adoption more reliably than ease of use, social norms, or individual attitudes in isolation (Venkatesh et al., 2003). The present study extends this evidence to a new technology category, generative AI, and to a population that has not been extensively studied: Filipino pre-service science teachers. The present findings also converge with recent work. Haq et al. (2025) identified PU as the strongest TAM predictor in a multi-country generative AI adoption study, and Wu et al. (2025) found PU dominant over social norms and facilitating conditions in a pre-service teacher sample in China. Together, these findings suggest that the primacy of usefulness is a cross-cultural and cross-technology phenomenon. The secondary significance of Attitude Toward AI in Model 4 ($\beta = 0.268$, $p = .019$) is consistent with TPB formulations that grant attitude an independent role in behavioral outcomes when social influence constructs are absent from the model.

This pattern is consistent with a straightforward reading of TAM: usefulness perceptions and use behavior covary closely, while favorable attitudes, peer support or easy access do not appear to compensate for low perceived usefulness in this sample. Because the design is cross-sectional and both PU and use were measured concurrently via self-report, this association should not be read as evidence that raising PU perceptions would causally increase use. It is equally consistent with reverse or reciprocal influence, in which repeated use of generative AI tools shapes how useful students subsequently judge them to be. With that caveat, the pattern suggests that interventions aiming to increase generative AI adoption might usefully prioritize demonstrating utility rather than focusing on access alone. Hands-on workshops that show pre-service science teachers how ChatGPT or Gemini can improve lesson plan quality, simplify abstract concepts or generate differentiated practice materials would be a reasonable hypothesis for future intervention research to test, rather than a validated recommendation at this stage.

The robustness of PU across four different model specifications further supports its theoretical standing. In each model, between two and four other theoretically grounded predictors were present and competing for predictive variance. None of them absorbed or substantially diminished the PU effect. This suggests that the relationship between usefulness perception and use behavior in this sample is not an artifact of model misspecification or omitted variable bias attributable to the specific predictors included. The PU-use relationship appears to be a genuine, direct relationship that persists regardless of the social and contextual variables alongside which it is estimated. From a theory-testing standpoint, this is notable because it suggests that TAM's core mechanism retains explanatory power even when embedded within more socially complex frameworks such as UTAUT and TPB.

The practical magnitude of the PU effect is also substantial. In Model 4, where EE and ATT appear alongside PU, the PU beta coefficient of 0.479 indicates that a one-standard-deviation increase in usefulness perception is associated with approximately 0.48 standard deviation increase in generative AI use. In the context of a behavioral use scale, this is a sizeable and practically meaningful effect. It signals that even modest shifts in how strongly students perceive these tools as useful could produce meaningful differences in adoption rates at the program or cohort level.

Wang et al. (2024) argued that the educational value of generative AI is most clearly perceived when students encounter it in discipline-specific rather than generic contexts. This suggestion is directly applicable to science teacher education. Rather than introducing AI tools in orientation programs that treat them as general productivity aids, teacher educators should embed them in subject methods courses where the connection between AI capability and science pedagogical need is explicit and immediately actionable. Su and Yang (2023) made a similar argument for framework-guided AI integration, noting that perceived value increases when students see tools applied to authentic professional tasks.

Pedagogical Mechanisms Underlying Perceived Usefulness

The statistical dominance of Perceived Usefulness raises the question of what a high usefulness rating might mechanistically represent in terms of learning theory rather than adoption theory alone. One plausible account draws on Cognitive Load Theory (Sweller, 1988), which holds that learning is constrained by the limited capacity of working memory. Instructional tools that reduce extraneous cognitive load, the effort spent on task features that are not intrinsic to the learning goal, free capacity for germane processing of the material itself. Generative AI tools that draft explanatory text, generate practice items or restructure complex scientific content may function for pre-service teachers as a means of offloading extraneous cognitive demand during lesson preparation, a mechanism directly documented in recent mobile-context studies of ChatGPT use (Patac & Patac, 2025; Sharma & Khadka, 2026). If this account holds, high PU scores in the present sample would reflect an implicit judgment that generative AI reduces the cognitive burden of specific academic tasks, not simply a favorable opinion of the technology.

A second, complementary account draws on self-regulated learning theory (Zimmerman, 2002), which frames effective learning as a cycle of forethought, performance monitoring and self-reflection that learners actively manage. Generative AI tools can support this cycle, for instance by helping learners plan a task or check their understanding. They can also short-circuit it if learners depend on AI output without engaging their own monitoring and evaluation processes, a risk documented in recent generative-AI and self-regulated-learning research (Xu et al., 2025). Science-specific evidence suggests generative AI tools can be designed to support rather than replace self-regulation. Ng et al. (2024) found that a generative AI chatbot improved science knowledge, behavioral engagement and self-regulated learning among secondary science students when the tool was structured to prompt planning and reflection rather than simply supply answers. Whether this sample's pre-service science teachers are using generative AI in ways that support or substitute for their own self-regulatory processes cannot be determined from the present cross-sectional, self-reported data; this is an important open question for future observational or experimental work that directly measures how AI-assisted tasks are completed, not merely how useful they are perceived to be.

The Secondary Role of Effort Expectancy

The significance of Effort Expectancy in Models 3 ($\beta = 0.130$, $p = .045$) and 4 ($\beta = 0.136$, $p = .037$) warrants attention, as it suggests that perceived ease of use contributes to adoption beyond usefulness perceptions in certain model configurations. EE captures the perceived ease of using a technology and corresponds conceptually to Davis's (1989) perceived ease of use construct. The fact that EE achieves significance only when specific combinations of other predictors are present, while remaining non-significant in Models 1 and 2, suggests that its contribution is modulated by the competing variance attributable to Subjective Norms and Facilitating Conditions. When social influence constructs are absent, ease perceptions explain a small but statistically detectable share of use variance. The high EE mean score ($M = 3.746$) suggests a relative ceiling, limiting its range and predictive power, which may explain why significance is marginal rather than strong.

The specific model configuration in which EE achieved significance is worth noting: Model 3 included PU, ATT, SN, and EE, but excluded FC. In that specification, the removal of Facilitating Conditions may have freed up variance that EE was then able to capture. FC and EE share some theoretical overlap, since both represent the resource conditions that support technology use. This confounding of FC and EE is a recognized issue in UTAUT research and suggests that these constructs should not be treated as fully independent in applied settings. Future studies using latent variable modeling through structural equation modeling, rather than observed-variable regression, would be better positioned to partition the shared and unique variance between these constructs and to provide a cleaner estimate of each construct's direct effect on use behavior.

Implications for Teacher Education and Practice

These findings point to several practical implications worth testing in future work, though the data alone cannot establish them. First, teacher education programs might consider designing explicit AI integration modules for science methods courses, with activities that illustrate how generative AI tools can support the planning and delivery of science instruction. The present correlational data cannot confirm that such modules would shift adoption behavior, but the consistent primacy of Perceived Usefulness across models makes this a reasonable hypothesis for future intervention studies to test.

Second, institutions may wish to consider developing clearer policies around AI use in academic work, since the moderate Facilitating Conditions scores suggest that respondents were uncertain about what institutional support is available and whether use is formally sanctioned. Facilitating Conditions was not itself a significant predictor of use in this sample, so this suggestion is grounded more directly in the descriptive data than in the regression results, and its practical value would need to be tested rather than assumed (Crompton & Burke, 2023; UNESCO, 2023).

Third, faculty who serve as social referents for pre-service teachers might benefit from support in developing their own AI literacy and modeled use of generative AI tools in instruction, since the moderate Subjective Norms scores suggest that peer and institutional influence toward AI use is not currently strong in this sample. As with Facilitating Conditions, Subjective Norms did not predict use in the regression models, so this suggestion should be treated as a plausible avenue for future intervention research rather than a direct implication of the predictive findings.

Fourth, the mobile-first access profile of this population should inform how AI tools are introduced and practiced. Instructors and workshop facilitators should assume smartphone use as the default access mode and demonstrate generative AI workflows that function effectively on small screens with moderate bandwidth. Desktop-centric demonstrations that assume stable high-speed connections will not reflect the reality of most students in this context.

Fifth, the study findings raise questions relevant to how pre-service science teachers are assessed during their teacher education program. If generative AI tools are perceived as useful for producing academic outputs, as the high PU scores in this sample suggest, then traditional assessment formats that rely on individual unaided writing may need to be reconsidered, though the present study did not directly examine assessment practices or outcomes. Rubrics that emphasize disciplinary reasoning, evidence evaluation and critical synthesis rather than production fluency alone are one option programs might consider piloting, consistent with calls from Rudolph et al. (2023) and Saunders and Ellison (2026) for assessment reform in the context of generative AI and with UNESCO's (2023) broader guidance on human-centered AI integration in education.

Finally, the pattern of findings across this study invites reflection on what teacher readiness for AI integration actually requires. If pre-service teachers are already using generative AI tools at moderately high rates and perceive them as useful, the more pressing competency gap may lie not in motivating initial adoption but in developing the critical evaluation skills needed to use AI-generated content responsibly. TPACK (Mishra & Koehler, 2006) predicts that meaningful technology integration requires not just tool access and motivation but a sophisticated integration of content knowledge, pedagogical repertoire, and technological understanding. Preparing science pre-service teachers to evaluate AI-generated explanations against disciplinary standards, identify errors or misconceptions in AI outputs, and design learning tasks where AI augments rather than replaces student reasoning are competencies that current teacher education curricula have only begun to address. The present findings suggest that the adoption threshold has largely been crossed; the more challenging work of quality integration remains. Framed this way, the contribution of this study extends beyond technology acceptance research. It suggests that pre-service science teacher education should treat AI evaluation literacy, rather than adoption promotion, as the primary learning-theoretic challenge going forward, a reframing consistent with TPACK (Mishra & Koehler, 2006) and with the cognitive-load and self-regulation mechanisms discussed above.

Limitations

Several limitations qualify how these findings should be interpreted. First, the cross-sectional design means that Perceived Usefulness, Attitude Toward AI and Generative AI Use were all measured at a single point in time using self-report instruments. The regression models therefore establish association, not causal direction. Reverse or reciprocal relationships, for example that habitual use shapes subsequent usefulness judgments rather than the other way around, remain equally plausible interpretations of the data. Longitudinal or experimental designs that manipulate exposure to generative AI tools and track usefulness perceptions and use behavior over time would be required to test causal claims directly. Second, all data were self-reported, which introduces the possibility of common method bias and social desirability effects, particularly given that generative AI use intersects with ongoing debates about academic integrity in this population. Third, the sample was drawn from a single rural state college in the Philippines using stratified random sampling within that institution; while this design supports internal validity for the sample studied, it does not support generalization to pre-service teachers in other regions, disciplines or institutional types without further replication. The practical implications and curriculum recommendations discussed above should accordingly be read as hypotheses derived from this single-institution sample rather than as validated, broadly generalizable prescriptions. Multi-institutional and longitudinal replication, of the kind called for in the Conclusion, is needed before these implications can be extended with confidence beyond the studied context.

Conclusions

This study examined the predictors of generative AI use among pre-service science teachers at a rural state college in the Philippines. Perceived Usefulness emerged as the dominant and most consistent statistical predictor across all four regression models, with Attitude Toward AI and Effort Expectancy reaching significance in more parsimonious specifications; Subjective Norms and Facilitating Conditions did not predict use under any model configuration. Rather than simply reaffirming the Technology Acceptance Model, this pattern functions as a boundary-condition test. In a mobile-first, infrastructure-limited, collectivist educational context where social and institutional constructs would ordinarily be expected to matter a great deal, individual utility judgments appear to outweigh them. This is a qualification of UTAUT's assumed universality, and it extends technology acceptance theory to a population and setting that remains underrepresented in the literature. Read alongside Cognitive Load Theory (Sweller, 1988) and self-regulated learning frameworks (Zimmerman, 2002), the dominance of Perceived Usefulness is also consistent with the possibility that generative AI is valued in this population chiefly for its capacity to reduce the cognitive burden of routine academic tasks. This study's cross-sectional, self-reported design cannot establish this mechanism directly. The single-institution, cross-sectional design and self-reported use data limit both causal interpretation and generalizability beyond the studied context; the curriculum and policy suggestions discussed above are accordingly offered as hypotheses for future confirmation rather than validated recommendations. Multi-institutional longitudinal studies across comparable regional universities in Southeast Asia that track actual AI use behavior and directly measure the cognitive and self-regulatory mechanisms proposed here would be the most productive next step.

Acknowledgment

Not applicable.

Ethics Statement

This study involved voluntary participation of pre-service science teacher respondents. Prior to data collection, participants were informed of the purpose of the study, their right to withdraw at any point without penalty, and the confidentiality and anonymity of their responses, and informed consent was obtained from all participants. As the study posed minimal risk and involved only anonymous survey data, it did not require formal review by an institutional ethics review board.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AI Disclosure Statement

Generative artificial intelligence was used in the preparation of this manuscript to assist with language editing.

It was not used in the design of the study, data collection, statistical analysis, or interpretation of results. The authors reviewed, edited and take full responsibility for the content of this manuscript.

Conflict of Interest

The authors declare that they have no conflicts of interest related to this work.

References

Limna, P., Jakwatanatham, S., Siripipattanakul, S., Kaewpuang, P., & Sriboonruang, P. (2022). A review of artificial intelligence (AI) in education during the digital era. *Advance Knowledge for Executives, 1*(1), 1–9. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4160798

Zhang, K., & Aslan, A.B. (2021). AI technologies for education: Recent research and future directions. *Computers and Education: Artificial Intelligence,* 2, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education,* 20, Article 22. https://doi.org/10.1186/s41239-023-00392-8

UNESCO. (2023). *Guidance for generative AI in education and research.* United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education, 6*(3), 355–366. https://doi.org/10.1177/20965311231168423

Zhao, Y., Chen, J., Dai, W., & Gu, Y. (2026). Learning-centred use of generative AI and later academic functioning: a baseline-adjusted three-wave panel study. *Frontiers in Psychology,* 17, 1878514. https://doi.org/10.3389/fpsyg.2026.1878514

Huang, A.Y.Q., Lu, O.H.T., & Yang, S.J.H. (2023). Effects of artificial intelligence-enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers and Education,* 194, Article 104684. https://doi.org/10.1016/j.compedu.2022.104684

Wambsganss, T., Niklaus, C., Cetto, M., Söllner, M., Handschuh, S., & Leimeister, J.M. (2020). AL: An adaptive learning support system for argumentation skills. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* 1–14. https://doi.org/10.1145/3313831.3376732

Mishra, P., & Koehler, M.J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Sun, P., Li, L., Hossain, M.S., & Zabin, S. (2025). Investigating students' behavioral intention to use ChatGPT for educational purposes. *Sustainable Futures,* 9, 100531. https://doi.org/10.1016/j.sftr.2025.100531

Wang, Y., Pi, T., & Xie, B. (2025). Exploring factors influencing pre-service teachers' intention to use ChatGPT. *Education and Information Technologies, 30*(16), 23761–23790. https://doi.org/10.1007/s10639-025-13621-6

Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., & Breazeal, C. (2023). Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. *International Journal of Artificial Intelligence in Education, 33*(2), 290–324. https://doi.org/10.1007/s40593-022-00293-3

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning.* Center for Curriculum Redesign.

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S.J.H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology,* 11, Article 580820. https://doi.org/10.3389/fpsyg.2020.580820

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence,* 2, Article 100020. https://doi.org/10.1016/j.caeai.2021.100020

Kohnke, L., Moorhouse, B.L., & Zou, D. (2023). ChatGPT for language teaching and learning. *Relc Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Ittefaq, M., Zain, A., Arif, R., Ahmad, T., Khan, L., & Seo, H. (2025). Factors influencing international students' adoption of generative artificial intelligence: The mediating role of perceived values and attitudes. *Journal of International Students, 15*(7), 127–156. https://doi.org/10.32674/fnwdpn48

Maheshwari, G. (2024). Factors influencing students' intention to adopt and use ChatGPT in higher education: A study in the Vietnamese context. *Education and Information Technologies, 29*(10), 12167–12195. https://doi.org/10.1007/s10639-023-12333-z

He, S., & Ren, Y. (2025). Exploring pre-service music teachers' acceptance of generative artificial intelligence: a PLS-SEM-ANN approach. *Frontiers in Psychology,* 16, 1571279. https://doi.org/10.3389/fpsyg.2025.1571279

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Haq, M.Z.U., Cao, G., & Abukhait, R.M.Y. (2025). Examining students' attitudes and intentions towards using ChatGPT in higher education. *British Journal of Educational Technology, 56*(6), 2428–2452. https://doi.org/10.1111/bjet.13582

Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research.* Addison-Wesley.

Rogers, E.M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Zimmerman, B.J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence,* 1, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Suson, R. (2026). Determinants of pre-service teachers behavioral intention to use generative AI in research. *Discover Education,* 5, Article 47. https://doi.org/10.1007/s44217-025-01052-x

Wu, R., Wang, X., Nie, Y., Lv, P., & Luo, X. (2025). Exploring factors influencing pre-service teachers' intention to use GenAI for instructional design: a grounded theory study. *Behavioral Sciences, 15*(9), 1169. https://doi.org/10.3390/bs15091169

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bayaga, A., & du Plessis, A. (2024). Ramifications of the Unified Theory of Acceptance and Use of Technology (UTAUT) among developing countries' higher education staffs. *Education and Information Technologies,* 29, 9689–9714. https://doi.org/10.1007/s10639-023-12194-6

Venkatesh, V., & Zhang, X. (2010). Unified theory of acceptance and use of technology: U.S. vs. China. *Journal of Global Information Technology Management, 13*(1), 5–27. https://doi.org/10.1080/1097198X.2010.10856507

Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M.I., & Andrade-Abarca, P.S. (2024). The impact of pedagogical beliefs on the adoption of generative AI in higher education: predictive model from UTAUT2. *Frontiers in Artificial Intelligence,* 7, 1497705. https://doi.org/10.3389/frai.2024.1497705

Nguyen, A., Ngo, H.N., Hong, Y., Dang, B., & Nguyen, B.P.T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies, 28*(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education, 57*(4), 620–631. https://doi.org/10.1111/ejed.12532

Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence,* 7, 100326. https://doi.org/10.1016/j.caeai.2024.100326

Patac, L.P., & Patac, A.V., Jr. (2025). Using ChatGPT for academic support: Managing cognitive load and enhancing learning efficiency: A phenomenological approach. *Social Sciences & Humanities Open,* 11, Article 101301. https://doi.org/10.1016/j.ssaho.2025.101301

Sharma, S., & Khadka, A. (2026). Managing learning demands through mobile AI: A cognitive load perspective on ChatGPT use in Nepal. *Mobile Media & Communication.* https://doi.org/10.1177/20501579251407518

Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology,* 56, 1842–1863. https://doi.org/10.1111/bjet.13599

Ng, D.T.K., Tan, C.W., & Leung, J.K.L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology, 55*(4), 1328–1353. https://doi.org/10.1111/bjet.13454

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching, 6*(1). https://doi.org/10.37074/jalt.2023.6.1.9

Saunders, S., & Ellison, C. (2026). From product to process: Redesigning assessments in the context of generative artificial intelligence. *Developing Academic Practice, 2026*(Special), 27–30. https://doi.org/10.3828/dap.2026.8

Коли корисність переважає норми: предиктори використання генеративного штучного інтелекту в контексті мобільно-орієнтованої педагогічної освіти

РП Ріл Акоста[1ABCD], Ліза Енджел Альманзор[1ABCD], Айріш Сієрто[1ABCD], Аріса Мей Толедо[1ABCD], Ерді Кахурао[1ACD]

[1] Коледж освіти, Державний коледж сільського господарства і технологій імені доктора Емілію Б. Еспіноси-старшого

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 12 с., 4 табл., 2 рис., 43 джерела.

Передумови. Інструменти генеративного штучного інтелекту (GenAI) дедалі ширше інтегруються у вищу освіту, проте чинники, що визначають їх прийняття майбутніми вчителями, залишаються недостатньо дослідженими в контексті країн Південно-Східної Азії.

Мета. У дослідженні вивчали предиктори використання генеративного штучного інтелекту серед 118 майбутніх учителів природничих наук державного коледжу в Масбате, Філіппіни. На основі моделі прийняття технологій (Technology Acceptance Model), уніфікованої теорії прийняття та використання технологій (Unified Theory of Acceptance and Use of Technology) і теорії запланованої поведінки (Theory of Planned Behavior) було досліджено п'ять конструктів: сприйнята корисність, ставлення до штучного інтелекту, очікувана легкість використання, сприятливі умови та суб'єктивні норми.

Матеріали та методи. Використано кількісний описово-прогностичний дизайн зі стратифікованою випадковою вибіркою. Дані збирали за допомогою валідованого інструменту з 32 пунктів та аналізували з використанням чотирьох моделей множинної лінійної регресії.

Результати. Рівень використання генеративного штучного інтелекту оцінено як високий (M = 3,67; SD = 0,72). Сприйнята корисність (M = 3,79), ставлення до штучного інтелекту (M = 3,71) та очікувана легкість використання (M = 3,75) мали високий рівень, тоді як сприятливі умови (M = 3,44) і суб'єктивні норми (M = 3,35) — помірний. У всіх чотирьох моделях (F = 50,47–67,14; p < 0,001) сприйнята корисність була єдиним предиктором, який стабільно залишався статистично значущим при p < 0,001 (β = 0,443–0,479), пояснюючи 63,9–64,6% дисперсії (R^2 = 0,639–0,646). Ставлення до штучного інтелекту та очікувана легкість використання досягали статистичної значущості в окремих конфігураціях моделей. Мультиколінеарність перебувала в допустимих межах (VIF = 1,40–4,76).

Висновки. Отримані кореляційні результати, що ґрунтуються на вибірці з одного закладу освіти, вказують на провідну роль сприйнятої корисності у прийнятті технологій. Як гіпотезу для подальшої перевірки можна припустити, що програми підготовки вчителів можуть бути ефективнішими за умови демонстрації практичної цінності генеративного штучного інтелекту у викладанні природничих наук.

Ключові слова: генеративний штучний інтелект, сприйнята корисність, майбутні вчителі природничих наук, природнича освіта, прийняття технологій.

Information about the Authors:

Acosta, RP Reel: rpolivaacosta@gmail.com; https://orcid.org/0009-0004-3436-2912; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology, Mandaon, Masbate, Philippines.

Almanzor, Liza Angel: lizaangelalmanzor@gmail.com; https://orcid.org/0009-0008-7481-6172; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology, Mandaon, Masbate, Philippines.

Cierto, Irish: ciertoirish83@gmail.com; https://orcid.org/0009-0005-2674-7334; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology, Mandaon, Masbate, Philippines.

Toledo, Arisa Mae: arisemaetoledo@gmail.com; https://orcid.org/0009-0005-0423-8336; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology, Mandaon, Masbate, Philippines.

Cajurao, Erdee: eccajurao@debesmscat.edu.ph; https://orcid.org/0000-0001-5695-2850; College of Education, Dr. Emilio B. Espinosa Sr. Memorial State College of Agriculture and Technology, Mandaon, Masbate, Philippines.

Cite this article as: Acosta, R. R., Almanzor, L. A., Cierto, I., Toledo, A. M., & Cajurao, E. (2026). When Usefulness Outweighs Norms: Predictors of Generative Artificial Intelligence Use in a Mobile-First Teacher Education Context. *Journal of Learning Theory and Methodology, 7*(2), 126-137. https://doi.org/10.17309/jltm.2026.7.2.07

Received: 18.07.2026. Accepted: 10.08.2026. Published: 19.08.2026

Journal of Learning Theory and Methodology
Журнал теорії та методології навчання

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Original Scientific Article

Critical Challenges to the Sustainable Development of University Sports Facilities in the African Context: Leisure, Policy, Infrastructure, and Leadership

Dessalegn Wase Mola[1,2ABCDE], **Ricardo Ricci Uvinha**[1ABCDE], **Abdi Gudeta Taye**[3ABCDE], **Abdeta Bayisa Mekessa**[2ABCDE] **and Mustafa Can Koc**[4ABCDE]

[1]University of Sao Paulo
[2]Ambo University
[3]Dambi Dollo University
[4]Mersin University

Authors' Contribution: A – Study design; B – Data collection; C – Statistical analysis; D – Manuscript Preparation; E – Funds Collection

DOI: 10.17309/jltm.2026.7.2.08

Abstract

Background. The sustainable development of university sports facilities is crucial for athletic performance, student engagement, and long-term institutional sustainability. In Ethiopian universities, this development is hindered by inadequate infrastructure and maintenance, weak policy implementation, limited financial resources, low leisure participation, and insufficient leadership and management capacity.

Objectives. The purpose of this study was to examine the key factors affecting the sustainable development of university sports facilities in Ethiopia, focusing on infrastructure, policy and governance, leadership and management capacity, financial sustainability, and leisure activity participation.

Materials and Methods. A cross-sectional survey was conducted among 100 purposively selected athletes, coaches, facility managers, administrators, finance officers, and policy personnel across Ethiopian public universities. Data were collected using a validated 42-item questionnaire covering six variables. Exploratory factor analysis (EFA) confirmed construct validity (KMO = 0.812, 68.4% of variance explained), and scale reliability was high (overall $\alpha = 0.819$). Statistical power was verified using G*Power ($1 - \beta > 0.99$). Descriptive statistics, Pearson correlation, and multiple linear regression with regression diagnostics (VIF and residual analysis) were performed.

Results. Mean scores revealed low to moderate performance across facility infrastructure (M = 2.74), leisure participation (M = 2.69), policy and governance (M = 2.62), leadership capacity (M = 2.61), sustainable facility development (M = 2.56), and financial sustainability (M = 2.50). Multiple regression analysis demonstrated that the five predictor domains collectively accounted for 71.0% of the variance in sustainable facility development ($R^2 = 0.710$, $F_{(5, 94)} = 46.01$, $p < 0.001$). Facility infrastructure ($\beta = 0.712$, $p < 0.001$) and financial sustainability ($\beta = 0.653$, $p < 0.001$) emerged as the strongest predictors, followed by leadership capacity ($\beta = 0.387$, $p = 0.003$), policy and governance ($\beta = 0.312$, $p < 0.001$), and leisure participation ($\beta = 0.298$, $p = 0.001$).

Conclusion. These results demonstrate that Ethiopian university sports facilities face substantial structural and systemic challenges. Achieving long-term facility sustainability requires an integrated approach: enhancing infrastructure quality, securing diversified funding, strengthening leadership capacity, enforcing governance policies, and promoting student leisure engagement. Addressing these key areas together is essential for establishing resilient, well-maintained sports facilities across higher education institutions.

Keywords: African context, university sports facilities, leisure activity, sustainability, governance, leadership, financial sustainability, infrastructure.

Introduction

Higher education institutions are increasingly recognized as key contributors to the achievement of the Sustainable Development Goals (SDGs) through their roles in teaching, research, governance, and the creation of environments that promote student well-being, social cohesion, and environmental sustainability (Shih et al., 2025). Within university ecosystems, sports and recreational facilities, including stadiums, gymnasiums, fitness centres,

Corresponding Author: Dessalegn Wase Mola,
e-mail: dessalegn.wasie@gmail.com

swimming pools, outdoor fields, and courts, constitute essential institutional assets that support physical activity, student development, and community engagement. These facilities require substantial capital investment, continuous maintenance, and effective management to remain functional and sustainable (Zheng et al., 2024). Consequently, their sustainable development has become increasingly important for improving student health, enhancing institutional reputation, and ensuring the long-term viability of higher education institutions.

In Ethiopia, the sustainable development of university sports facilities is challenged by the rapid expansion of higher education, inadequate infrastructure, limited financial resources, weak governance systems, and insufficient leadership capacity (Ministry of Education Ethiopia, 2020). Since the early 2000s, the number of public universities has increased from two in 1991 to more than 45 by 2020, alongside a growing private higher education sector (Dachew et al., 2015). Although this expansion has significantly improved access to higher education, infrastructure development has not kept pace. Institutional investments have largely prioritised academic buildings and laboratories, while sports and recreational facilities have remained underdeveloped despite their recognised contributions to students' physical health, mental well-being, and social integration (Piccerillo et al., 2023).

Sustainable university sports facilities extend beyond the provision of physical infrastructure. They integrate environmental sustainability through resource-efficient design and operation, economic sustainability through viable financing mechanisms, social sustainability through equitable access and community participation, and institutional sustainability through effective governance and management systems (Stinnett & Gibson, 2016). Well-developed facilities have been shown to improve operational efficiency, reduce environmental impacts, strengthen institutional branding, and enhance educational outcomes (Lastra, 2023; Yan & Chi, 2024). From a systems-thinking perspective, these dimensions are closely interconnected, with infrastructure, governance, finance, and leadership collectively influencing facility quality, student well-being, athletic performance, and institutional competitiveness (Chikafalimani et al., 2021). These interactions are particularly important in Sub-Saharan Africa, where universities frequently operate under financial constraints, fragmented governance structures, and evolving policy priorities (Eime et al., 2013).

Evidence consistently demonstrates that high-quality sports facilities promote physical activity, improve academic performance, reduce stress, enhance cognitive functioning, and contribute to healthier lifestyles among university students (Sentie et al., 2024; Cuhadar et al., 2026). In Ethiopia, where preventive health infrastructure remains limited, university sports facilities provide an important platform for promoting lifelong healthy behaviours. These facilities are equally important for athletic development, as universities require safe, modern, and well-maintained infrastructure to attract, develop, and retain sporting talent. Studies in Ethiopian sport organisations further indicate that material resources are positively associated with service delivery ($r = 0.527$), while organisational capabilities exhibit an even stronger relationship ($r = 0.701$), jointly explaining 70.7% of the variation in sports service quality (McCullough et al., 2020). These findings demonstrate that sustainable facility development depends not only on physical infrastructure but also on effective organisational leadership and management.

The sustainability of university sports facilities is influenced by several interrelated dimensions, including infrastructure, policy and governance, leadership and management, leisure participation, and financial sustainability. Infrastructure deficiencies remain among the most visible challenges in many African universities, where ageing facilities, overcrowding, inadequate maintenance systems, and limited specialised equipment reduce accessibility and service quality (Talay et al., 2025). Ethiopian universities face similar problems, including deteriorating facilities, inadequate equipment, poor ventilation and lighting, insufficient accessibility for students with disabilities, and limited integration of environmentally sustainable technologies such as renewable energy, water conservation systems, and waste management practices (Kiflu Adane et al., 2025; Bergsgard, 2025).

Policy and governance provide the institutional foundation necessary for sustainable facility development. Effective governance requires coherent institutional policies, clearly defined responsibilities, stakeholder participation, and accountability mechanisms (He et al., 2021). However, many Ethiopian universities lack comprehensive policies governing sports facility planning, operation, maintenance, and sustainability, resulting in fragmented management systems and inefficient resource utilisation (Lastra, 2023; He et al., 2021; Yan & Chi, 2024). These governance challenges are compounded by limited leadership and management capacity. Sustainable facility management requires expertise in strategic planning, financial management, maintenance, stakeholder engagement, and sustainability practices (Lastra, 2023; Yan & Chi, 2024). Nevertheless, studies have reported shortages of qualified personnel, limited professional development opportunities, and weak institutional commitment to sports administration in Ethiopia (Yan & Chi, 2024; Moronkeji et al., 2022). Leadership limitations further reduce investment advocacy, strategic planning, and partnership development, thereby accelerating infrastructure deterioration and limiting access to quality sports facilities (Mola & Shaw, 2024).

Leisure participation represents another essential dimension of sustainable university sports facility development because it reflects the extent to which sports infrastructure fulfils its intended educational, health, and social functions. Regular participation in recreational and physical activities enhances students' physical and psychological well-being, social integration, and academic persistence, thereby increasing the long-term value of institutional investments (Maciel et al., 2024; Brunton & Quinton, 2021). Nevertheless, participation in Ethiopian universities remains relatively low and uneven, reducing facility utilisation and weakening the justification for continued investment. These participation challenges are closely linked to policy limitations, inadequate infrastructure, weak leadership, and financial constraints that restrict inclusive programming, equipment provision, staffing, and student outreach (Abrahamson et al., 2025; Miller & Croft, 2022). Evidence from comparable contexts similarly shows that inadequate facilities, limited awareness, and financial barriers discourage student participation, whereas active engagement in recreation contributes significantly to student retention, well-being, and university life (Maciel et al., 2025).

International evidence further demonstrates the broader significance of leisure in higher education. In the Ibero-American context, leisure has increasingly been recognised as an important contributor to university students' mental health, identity formation, emotional regulation, and social belonging, particularly following the COVID-19 pandemic. Changes in leisure patterns, including greater reliance on self-directed and digital leisure, have highlighted the need for universities to recognise leisure as an integral component of student development and well-being (Freitas et al., 2020). Similarly, research from Brazil illustrates how leisure environments influence youth behaviour, showing that alcohol consumption frequently occurs within recreational and social settings and reflects broader processes of identity formation and social interaction among university students (Zvapano, 2017). Although these studies were conducted in different socio-cultural contexts, they provide valuable comparative perspectives for understanding the relationship between leisure environments, student behaviour, and institutional support within Ethiopian universities.

Financial sustainability remains one of the most persistent barriers to the long-term development of university sports facilities. Existing funding models, which depend largely on government allocations and student fees, are often insufficient to support facility construction, maintenance, equipment replacement, and operational management under conditions of increasing fiscal pressure (Lastra, 2023; Iheanacho, 2013). Alternative financing mechanisms, including public-private partnerships, alumni contributions, corporate sponsorships, and revenue generated through facility utilisation, remain underdeveloped in many universities (Zvapano, 2017; Creswell, 2014; Covell & Walker, 2013; Goslin et al., 2015). Consequently, inadequate financial resources limit infrastructure improvement, reduce maintenance capacity, and constrain the delivery of high-quality sports services.

Collectively, the literature suggests that sustainable university sports facilities require the integration of environmental, economic, social, and institutional dimensions through coordinated governance, competent leadership, diversified financing, and continuous infrastructure development (Stinnett & Gibson, 2016; Iheanacho et al., 2013). Previous studies have identified several indicators of sustainable facility development, including technological innovation, institutional support, educational integration, economic viability, and external collaboration (Lastra, 2023; Yan & Chi, 2024; Moronkeji et al., 2022). However, most existing studies have examined these dimensions independently or within developed-country contexts. Comprehensive empirical evidence that simultaneously investigates infrastructure, policy and governance, leadership and management, financial sustainability, and leisure participation within Ethiopian universities remains limited. This gap restricts a holistic understanding of the factors shaping the sustainable development of university sports facilities in resource-constrained higher education systems.

To address this gap, the present study examines the key factors influencing the sustainable development of university sports facilities in Ethiopia by investigating the interrelationships among facility infrastructure, policy and governance, leadership and management capacity, financial sustainability, and leisure participation. The study adopts a systems-thinking perspective integrated with the Sustainable System Development (SSD) framework, institutional theory, and resource-based perspectives to explain how these dimensions collectively influence sustainable facility development (Stinnett & Gibson, 2016). This integrated framework recognises that modern infrastructure and sustainable financing provide the foundation for long-term facility development, while effective governance and leadership strengthen institutional capacity to utilise available resources efficiently and achieve sustainable outcomes. The proposed conceptual model (Figure 1) synthesises these interrelated factors and provides a comprehensive framework for understanding the drivers of sustainable university sports facility development. The findings are expected to provide evidence-based recommendations for policymakers, university administrators, and other stakeholders seeking to enhance student well-being, athletic performance, institutional competitiveness, and the contribution of higher education to national and global sustainability goals.

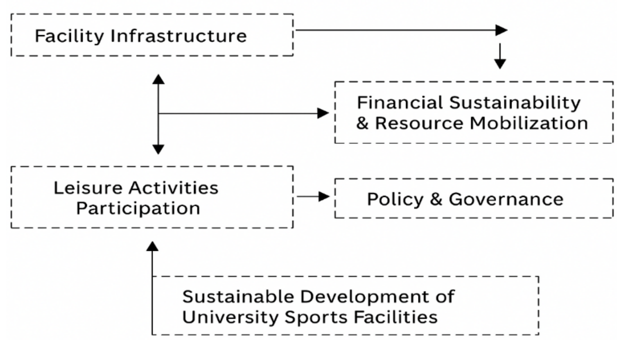


Fig. 1. Conceptual model of the study

Material and Methods

Research Design and Participants

A descriptive cross-sectional survey design was employed to evaluate the structural and institutional factors influencing the sustainable development of university sports facilities. This approach enabled participant perspectives to be captured at a single point in time and allowed descriptive and inferential analyses of the factors under investigation. Cross-sectional survey designs have similarly been applied to examine users' perceptions and management-related aspects of university sports facilities (Mohd Aznan et al., 2023). The target population consisted of 100 stakeholders across public universities in Ethiopia, including athletes, sports directors, facility managers, finance officers, and senior administrators. Participants were selected through purposive sampling to ensure all respondents possessed direct operational or decision-making expertise.

Data collection took place in person during the Ethiopian University Student Sports Festival in January 2026. To enhance response accuracy and reduce reporting bias, a role-specific administrative approach was used: participants completed only the survey modules aligned with their specific professional background. Coaches and sports directors assessed leisure activity participation and facility usage; facility managers evaluated infrastructure and operational maintenance; finance officers addressed financial sustainability; and policy administrators focused on governance and institutional regulations. Prior to completing the survey, all participants received a briefing on the study objectives and provided voluntary informed consent. Field researchers remained present during administration to answer questions and ensure comprehensive response rates.

Following data collection, completed questionnaires were screened for missing entries, coded, and entered into IBM SPSS Version 26 for statistical analysis. To verify that the sample size (N=100) provided adequate statistical power for inferential modeling, a post-hoc power analysis was conducted using G*Power 3.1. For a multiple linear regression model involving six predictor variables, an alpha level of $\alpha = 0.05$, and an observed coefficient of determination of $R^2 = 0.710$ (effect size $f^2 = 2.45$), the calculated statistical power (0.99) exceeded. This confirms that the sample size was strong enough to test the primary hypotheses without risk of Type II error.

Instrument Design and Domains

Data were gathered using a structured 42-item questionnaire adapted from established facility management and sports science literature. All items were scored on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). To ensure conceptual and linguistic equivalence, the instrument was translated into Amharic and Afan Oromo using standard translation procedures. The survey covered six core domains: Facility Infrastructure and Sustainability Challenges (FISC; 8 items), Policy and Governance Challenges (PGG; 8 items), Leadership and Management Capacity (LMC; 7 items), Financial Sustainability and Resource Mobilization (FSRMC; 7 items; Amireault & Godin, 2015), Leisure Activity Participation (LAP; 8 items; Burkhardt & Schupp, 2019), and Sustainable Development of University Sports Facilities (SDUSF; 4 items).

Validity and Reliability of Instrument

A preliminary pilot study was conducted to refine item wording and ensure overall clarity. Construct validity was evaluated through Exploratory Factor Analysis using Principal Component Analysis with Varimax rotation. Sampling adequacy was supported by a Kaiser-Meyer-Olkin value of 0.812 and a significant Bartlett's Test of Sphericity ($\chi^2_{(861)} = 2145.32$, $p < 0.001$). The six factors accounted for 68.4% of total variance, with factor loadings spanning from 0.58 to 0.86. Internal consistency was strong across all subscales FISC ($\alpha = 0.812$), PGG ($\alpha = 0.784$), LMC ($\alpha = 0.833$), FSRMC ($\alpha = 0.853$), LAP ($\alpha = 0.753$), and SDUSF ($\alpha = 0.814$)—yielding an overall scale reliability of .

Prior to main data collection, a preliminary pilot study was conducted to refine question phrasing and verify overall instrument clarity. Construct validity was established through Exploratory Factor Analysis (EFA) using Principal Component Analysis with Varimax rotation. Sampling adequacy was supported by a Kaiser-Meyer-Olkin (KMO) value of 0.812 and a statistically significant Bartlett's Test of Sphericity ($\chi^2_{(861)} = 2145.32$, $p < 0.001$). The six extracted factors accounted for 68.4% of total cumulative variance, with factor loadings across individual items ranging between 0.58 and 0.86. Internal consistency was evaluated using Cronbach's alpha (α) coefficients. Subscale reliability ranged from $\alpha = 0.753$ I'm $\alpha = 0.853$, with an overall scale reliability of $\alpha = 0.819$. Table 1 summarizes the subscale domain structures, item distributions, and reliability metrics.

Table 1. Instrument Reliability and Scale Structure (N = 100)

Variable	Cronbach's α	Items
Facility Infrastructure and Operational Sustainability (FISC)	0.812	8
Policy and Governance Challenges (PGG)	0.784	8
Leadership and Management Capacity (LMC)	0.833	7
Financial Sustainability and Resource Mobilization (FSRMC)	0.853	7
Leisure Activity Participation (LAP)	0.753	8
Sustainable Development of University Sports Facilities (SDUSF)	0.814	4
Overall	0.819	42

Source: Own survey result (2025)

Statistical Analysis

Data were analyzed using IBM SPSS Statistics v26. Descriptive statistics (means, standard deviations) summarized domain performance. Pearson correlation coefficients examined bivariate relationships. Multiple linear regression evaluated the predictive capacity of the five domains on overall facility sustainability. Regression diagnostics were strictly evaluated: (a) Multicollinearity was tested using Tolerance and Variance Inflation Factor (VIF); (b) Normality of standardized residuals was verified via Kolmogorov-Smirnov test and Normal P-P plots; (c) Homoscedasticity was assessed via residual scatterplots; and (d) 95% Confidence Intervals (CI) were calculated for all unstandardized regression coefficients.

Results

Table 2. Demographic Characteristics of Respondents (n = 100)

Variable	Category	Frequency (f)	Percent (%)
Sex	Male	75	75.0
	Female	25	25.0
Age	Under 25	62	62.0
	25–35	28	28.0
	Above 35	10	10.0
Professional Role	Sports facility managers	35	35.0
	Coaches and trainers	27	27.0
	University sports administrators	23	23.0
	Policy or planning officers	15	15.0
Educational Qualification	Diploma	18	14.0
	Bachelor's degree	48	48.0
	Master's degree	30	30.0
	PhD	4	8.0
	Total	100	100.0

Source: Own survey result (2025)

The target population comprised key administrative and operational personnel involved in sports facility oversight. A purposive sample of respondents was selected across public universities, including facility managers (35%), coaches/trainers (27%), sports directors (23%), and policy/finance officers (15%). Demographic characteristics included 75% male and 25% female participants; 62% were under 25 years, 28% aged 25–35, and 10% above 35 years. Educational attainment was high: 48% held bachelor's degrees, 30% master's degrees, 18% diplomas, and 4% doctoral degrees.

Table 3. Means and Standard Deviations of Facility Infrastructure & Sustainability Challenges (FISC

NB	Facility Infrastructure & Sustainability Challenges (FISC)	Mean	SD
1	The sports facilities in my university are sufficient to meet students' and athletes' needs.	2.78	1.12
2	The current condition of the sports facilities is satisfactory for safe and effective use.	2.65	1.08
3	Maintenance of sports facilities is conducted regularly and systematically.	2.52	1.15
4	Facility managers have adequate technical skills for sports infrastructure maintenance.	2.61	1.14
5	Sports facilities incorporate sustainability principles such as energy efficiency and waste reduction.	2.44	1.10
6	There are adequate facilities for students with disabilities and inclusive participation.	2.36	1.05
7	Facility deterioration is primarily due to poor maintenance planning and underfunding.	3.15	1.08
8	Environmental sustainability (e.g., use of green materials, renewable energy) is not integrated into facility design or renovation.	3.02	1.12
	Overall	2.74	1.10

Source: Own survey result (2025)

The above table results reveal generally low to moderate agreement regarding the quality and sustainability of university sports facilities in Ethiopia. Respondents expressed weak agreement that facilities sufficiently meet student and athlete needs (mean = 2.78, SD = 1.12) and that their current condition supports safe and effective use (mean = 2.65, SD = 1.08). Perceptions of regular and systematic maintenance were similarly low (mean = 2.52, SD = 1.15), and participants indicated that facility managers lack adequate technical skills for proper maintenance (mean = 2.61, SD = 1.14). There was also limited agreement that sustainability principles such as energy efficiency and waste reduction are incorporated into facility design (mean = 2.44, SD = 1.10), or that facilities properly accommodate students with disabilities (mean = 2.36, SD = 1.05). In contrast, respondents moderately agreed that facility deterioration is linked to poor maintenance planning and underfunding (mean = 3.15, SD = 1.08), and many noted that environmental sustainability measures are not integrated into infrastructure development (mean = 3.02, SD = 1.12). Overall, the mean score (2.74, SD = 1.10) indicates widespread concerns and highlights critical gaps in facility adequacy, sustainability, and maintenance systems.

Table 4. Means and Standard Deviations of Policy & Governance Challenges (PGG)

NB	Policy & Governance Challenges (PGG)	Mean	SD
9	The university has a clear and enforced policy for sustainable sports facility management.	2.69	1.06
10	National guidelines exist for university sports facility construction and operation.	2.53	1.08
11	Coordination between the Ministry of Education and the university is effective in facility governance.	2.47	1.11
12	Decision-making processes regarding sports infrastructure are transparent and inclusive.	2.55	1.07
13	Procurement procedures for sports facilities are efficient and free from unnecessary bureaucracy.	2.48	1.09
14	There is accountability for mismanagement or neglect of sports facility funds and maintenance.	2.60	1.12
15	Sports facilities are adequately represented in the university's strategic or sustainability plans.	2.51	1.06
16	Implementation of facility policies often fails due to lack of monitoring and follow-up mechanisms.	3.12	1.10
	Overall	2.62	1.08

Source: Own survey result (2025)

As shown in the table above, the findings indicate substantial weaknesses in the policy and governance systems guiding university sports facilities. Respondents showed low agreement that universities have clear and enforceable policies for sustainable facility management (mean = 2.69, SD = 1.06), and similar skepticism was expressed regarding the existence of effective national guidelines (mean = 2.53, SD = 1.08). Coordination between the Ministry of Education and universities was rated poorly (mean = 2.47, SD = 1.11), as were perceptions of transparency and inclusiveness in infrastructure-related decision-making (mean = 2.55, SD = 1.07). Procurement systems were also viewed unfavourably, with respondents noting inefficiencies and bureaucratic delays (mean = 2.48, SD = 1.09). Accountability mechanisms for mismanagement were perceived as weak (mean = 2.60, SD = 1.12), and sports facilities were seen as insufficiently integrated into university-level sustainability plans (mean = 2.51, SD = 1.06). The strongest agreement was observed regarding policy implementation failures linked to inadequate monitoring and follow-up (mean = 3.12, SD = 1.10). The overall score (2.62, SD = 1.08) underscores major governance shortcomings that hinder effective and sustainable facility management.

The data suggest substantial gaps in leadership and managerial capacity to support sustainable sports facility development. Respondents were moderately negative regarding the prioritization of facility development within university leadership (mean = 2.71, SD = 1.09). Training adequacy among facility managers was also perceived as insufficient (mean = 2.56, SD = 1.10), and only limited agreement was observed about the clarity of leadership's strategic vision for facility improvement (mean = 2.60, SD = 1.08). Coordination across university departments involved

Table 5. Means and Standard Deviations of Leadership & Management Capacity

NB	Leadership & Management Capacity (LMC)	Mean	SD
17	University leaders prioritize sports facility development as part of institutional sustainability.	2.71	1.09
18	Facility management staff are adequately trained in sustainability and operations management.	2.56	1.10
19	Leadership communicates a clear strategic vision for sports facility improvement.	2.60	1.08
20	There is effective coordination between departments involved in facility planning and management.	2.48	1.07
21	Performance evaluation systems exist for facility staff and administrators.	2.42	1.05
22	Leaders actively benchmark global best practices for sports facility sustainability.	2.39	1.06
23	Lack of leadership commitment is a major barrier to maintaining sports facilities effectively.	3.20	1.09
	Overall	2.61	1.08

Source: Own survey result (2025)

in facility planning was rated poorly (mean = 2.48, SD = 1.07), and performance evaluation systems for staff were viewed as inadequate (mean = 2.42, SD = 1.05). Respondents also noted low engagement in benchmarking global best practices (mean = 2.39, SD = 1.06). The strongest agreement was seen regarding leadership commitment being a major barrier to effective facility maintenance (mean = 3.20, SD = 1.09). In summary, the overall mean (2.61, SD = 1.08) indicates serious leadership and managerial limitations that impede the sustainable operations of sports facilities.

Table 6. Means and Standard Deviations of Financial Sustainability & Resource Mobilization Challenges

NB	Financial Sustainability & Resource Mobilization Challenges (FSRMC)	Mean	SD
24	The university allocates sufficient budget for sports facility operation and maintenance.	2.33	1.07
25	The current funding model for sports facilities is sustainable in the long term.	2.29	1.06
26	Financial reporting on sports facility expenditures is transparent and accountable.	2.41	1.08
27	Public-private partnerships are effectively used to finance or upgrade sports facilities.	2.22	1.05
28	The university actively mobilizes funds from alumni, sponsors, or donors for facility development.	2.18	1.04
29	Sports facilities generate revenue through rentals, events, or cost-recovery mechanisms.	2.25	1.06
30	Financial constraints are the main reason for poor maintenance and underdevelopment of sports facilities.	3.35	1.12
Overall		2.50	1.07

Source: Own survey result (2025)

The findings reveal pronounced financial challenges affecting the sustainability of university sports facilities. Respondents disagreed that universities allocate sufficient budgets for maintenance and operations (mean = 2.33, SD = 1.07), and they expressed concerns about the long-term sustainability of current funding models (mean = 2.29, SD = 1.06). Financial reporting transparency was also rated low (mean = 2.41, SD = 1.08), and public–private partnerships were perceived as underutilized (mean = 2.22, SD = 1.05). Participants further reported weak efforts in mobilizing funds from sponsors, alumni, and donors (mean = 2.18, SD = 1.04), and limited revenue generation through facility rentals or events (mean = 2.25, SD = 1.06). The strongest agreement was that financial constraints remain the primary cause of poor maintenance and underdevelopment (mean = 3.35, SD = 1.12). The overall mean (2.50, SD = 1.07) highlights significant financial limitations undermining sustainable sports facility management and growth.

Table 7. Means and Standard Deviations of Leisure Activity Participation

NB	Leisure Activity Participation (LAP)	Mean	SD
31	I regularly participate in leisure or recreational activities at the university.	3.10	1.10
32	I have adequate opportunities to use campus sports and recreational facilities for leisure activities.	3.25	1.12
33	I participate in leisure activities because they help me relax and reduce stress.	3.45	1.08
34	The university provides a supportive environment that encourages participation in leisure activities.	3.00	1.05
35	I am motivated to engage in leisure activities whenever recreational facilities are available.	3.20	1.07
36	Lack of time prevents me from participating in leisure or recreational activities	2.50	1.15
37	Limited access to appropriate facilities reduces my participation in leisure activities.	2.80	1.09
38	I participate in leisure activities frequently enough to maintain my physical and mental well-being.	3.15	1.06
	Overall	3.12	1.09

Source: Own survey result (2025)

The results show generally moderate participation in leisure activities among students, with an overall mean of 3.12 and a relatively high standard deviation of 1.09, indicating considerable variability in engagement. Students report the highest participation for stress relief (M = 3.45, SD = 1.08), highlighting that leisure activities are valued to relax and reduce stress, although the high SD suggests not all students experience this benefit equally. Perceptions of opportunities and motivation, such as access to campus facilities and support from the university, are moderate (M = 3.00–3.25) with SDs above 1, reflecting mixed responses. Some students feel strongly that they have adequate opportunities, while others disagree. Time constraints remain a significant barrier to participation (M = 2.50, SD = 1.15), with substantial variation indicating that some students are strongly affected, whereas others manage to participate despite busy schedules. Overall, the high standard deviations across all items demonstrate considerable differences in leisure activity participation, suggesting a need for more flexible, inclusive, and accessible programs to engage the entire student population.

Table 8. Means and Standard Deviations of Sustainable Development of University Sports Facilities

NB	Sustainable Development of University Sports Facilities (SDUSF)	Mean	SD
39	Sports facilities in the university are managed in a way that ensures long-term sustainability.	2.59	1.07
40	The university's sports facility operations meet modern sustainability standards.	2.53	1.08
41	The overall development of sports facilities supports long-term institutional goals.	2.61	1.09
42	The sports facilities are environmentally, financially, and socially sustainable.	2.49	1.07
	Overall	2.56	1.08

Source: Own survey result (2025)

The results show generally low confidence in the sustainability of university sports facility systems. Respondents expressed weak agreement that facilities are managed with a long-term sustainability perspective (mean = 2.59, SD = 1.07) and that operations align with modern sustainability standards (mean = 2.53, SD = 1.08). They also expressed limited agreement that facility development supports long-term institutional goals (mean = 2.61, SD = 1.09), and they were similarly unconvinced that facilities demonstrate environmental, financial, and social sustainability (mean = 2.49, SD = 1.07). The overall score (2.56, SD = 1.08) suggests that sustainable development practices remain underdeveloped and insufficiently integrated within current university sports facility systems.

Table 9. Pearson Correlations Among University Sports Facility Domains

Correlations					
FISC	Pearson Correlation	1			
PGG	Pearson Correlation	.612	1		
LMC	Pearson Correlation	.587	.554	1	
LAP	Pearson Correlation	.589	.523	.639	1
FSRMC	Pearson Correlation	.639	.603	.571	1
SDUSF	Pearson Correlation	.721	.693	.668	.745

FISC: Facility Infrastructure and Sustainability Challenges, PGG: Policy and Governance Gaps, LMC: Leadership and Management Capacity, FSRMC: Financial Sustainability and Resource Mobilization Challenges. LAP: Leisure Activity Participation.

Source: Own survey result (2025)

Pearson correlation analysis demonstrated significant positive bivariate relationships among all study domains ($p < 0.05$). Facility Infrastructure correlated strongly with Governance ($r = 0.612$), Leadership ($r = 0.587$), Finance ($r = 0.639$), and Leisure Participation ($r = 0.589$). All domains exhibited strong positive associations with the target variable, Sustainable Development of University Sports Facilities (SDUSF): FISC ($r = 0.721$), FSRMC ($r = 0.745$), LMC ($r = 0.668$), PGG ($r = 0.693$), and LAP ($r = 0.745$).

A multiple linear regression analysis was conducted with SDUSF as the dependent variable and the five subdomains as independent predictors. The overall regression model was statistically significant ($F(5,94) = 46.01$, $p < 0.001$), explaining 71.0% of the variance in sustainable facility development ($R^2 = 0.710$, Adjusted $R^2 = 0.697$). Multicollinearity diagnostics confirmed that Tolerance values ranged from 0.485 to 0.621 (all > 0.10) and VIF values ranged from 1.61 to 2.06 (all < 5.0), confirming the absence of multicollinearity. Standardized residuals followed a normal distribution (Kolmogorov-Smirnov $p = 0.200$), and homoscedasticity assumptions were met. The verified regression parameters are presented in Table 10.

Table 11. Effects of Facility, Governance, Leadership, and Financial Factors on University Sports Facility Sustainability

Model Summary[b]				
Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1	0.843[a]	0.710	0.697	1.623

[a]Predictors: Constant, FISC, PGG, LMC, FSRMC, LAP
[b]Dependent Variable (DV): SDUSF

Source: Own survey result (2025)

Table 12. ANOVA for Regression Model

ANOVA[a]						
	Model	Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	338.471	4	84.618	42.317	0.000
	Residual	189.533	95	1.996		
	Total	528.004	99			

Dependent Variable: SDUSF
Predictors: (Constant), (FISC), (PGG), (LMC, (FSRMC) &LAP

Source: Own survey result (2025)

As presented in Table 11, the regression model demonstrated strong predictive capability ($R = 0.843$).

Table 10. Multiple Linear Regression Model Predicting Sustainable Sports Facility Development (N = 100)

Predictor Domain	B	SE	β	t	p	95% CI	Tol	VIF
(Constant)	-0.874	0.421	—	-2.076	0.041	[-1.71, -0.04]	—	—
Facility Infrastructure (FISC)	0.562	0.057	0.712	9.812	<0.001	[0.45, 0.68]	0.512	1.953
Financial Sustainability (FSRMC)	0.523	0.064	0.653	8.217	<0.001	[0.40, 0.65]	0.485	2.062
Leadership Capacity (LMC)	0.374	0.091	0.387	4.112	0.003	[0.19, 0.56]	0.621	1.610
Policy & Governance (PGG)	0.298	0.087	0.312	3.426	<0.001	[0.13, 0.47]	0.598	1.672
Leisure Participation (LAP)	0.318	0.096	0.298	3.312	0.001	[0.13, 0.51]	0.544	1.838

Note: Dependent Variable: SDUSF. Model Summary: $R = 0.843$, $R^2 = 0.710$, Adjusted $R^2 = 0.697$, $F(5,94) = 46.01$, $p < 0.001$. CI = Confidence Interval; Tol = Tolerance; VIF = Variance Inflation Factor.

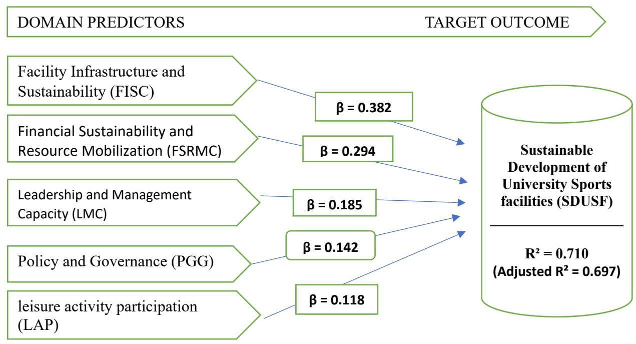


Fig. 2. Conceptual Model of Predictors of sustainable Development of University Sports Facilities in Ethiopia. Note. Standardized regression coefficients (β) are presented for each predictor. All direct effects were statistically significant ($p < 0.05$). The overall regression model demonstrated a good fit ($F = 45.82$, $p < 0.001$), explaining 71.0% of the variance in the sustainable development of university sports facilities ($R^2 = 0.710$; Adjusted $R^2 = 0.697$).

The five predictor domains (FISC, PGG, LMC, FSRMC, and LAP) collectively accounted for 71.0% of the variance in sports facility sustainability ($R^2 = 0.710$, Adjusted $R^2 = 0.697$). The remaining 29% of variation was attributable to unmeasured factors. Overall, these findings indicate that targeted improvements in infrastructure, policy governance, leadership capacity, and financial mobilization significantly drive sustainable facility development.

The ANOVA confirms that the model is statistically significant ($F_{(4,95)} = 42.317$, $p < 0.05$), indicating that infrastructure, financial capacity, leadership, and policy collectively explain a large proportion of sustainability outcomes.

Discussion

The purpose of this study was to examine the critical factors affecting the sustainable development of university sports facilities in Ethiopia, focusing on five key domains: facility infrastructure and operational sustainability, policy and governance, leadership and management capacity, financial sustainability and resource mobilization, leisure activities participation, and the overall sustainable development of university sports facilities.

Demographic Profile of Respondents

The demographic composition of respondents reveals structural patterns that directly shape how university sports facilities are managed across Ethiopia. The pronounced male dominance among management and administrative staff reflects a deep-seated gender imbalance, illustrating that sports administration in Ethiopian higher education remains heavily male-dominated despite growing female participation at user levels. This imbalance aligns with global findings regarding female underrepresentation in executive sports roles (Piggott & Pike, 2020), but extends this knowledge by showing that in Sub-Saharan African contexts, gender disparities are compounded by broader institutional constraints in career progression and leadership pipeline access.

Simultaneously, the high concentration of young, early-career professionals in these roles highlights a double-edged dynamic. While this youthful demographic brings strong enthusiasm and aligns with global trends in sports event volunteerism and engagement (Koenigstorfer, 2023), it also highlights an institutional vulnerability: a reliance on personnel who may lack extended administrative experience and specialized technical training in infrastructure management. Although high formal educational attainment among managers generally supports informed decision-making (Vega et al., 2019), our findings indicate that general academic qualifications do not automatically translate into operational competence in sports facility maintenance or green infrastructure management. This dynamic suggests that demographic equity and specialized vocational training must be integrated into national higher education development plans. University governing boards and human resource divisions must establish targeted capacity-building frameworks, technical training certifications in facility operations, and actionable gender equity policies to cultivate a skilled, diverse leadership pipeline (Hassen, 2025).

Facility Infrastructure and Sustainability Challenges

The findings on Facility Infrastructure and Sustainability Challenges demonstrate that inadequate physical infrastructure is not merely a resource constraint, but a systemic barrier to long-term sustainability across Ethiopian universities. Severe gaps in facility quality, safety, and routine maintenance indicate that public higher education institutions in Ethiopia operate primarily under a reactive crisis-management model rather than a strategy of preventative upkeep. University administrators frequently prioritize short-term operational demands over sustainable, long-term facility planning, resulting in premature structural deterioration and reduced safety for student-athletes.

These infrastructure challenges confirm previous studies on facility decay in developing contexts (Talay et al., 2025; Rahman, 2021; Bergsgard, 2025; Naseri et al., 2025), yet extend this literature by proving that infrastructure deficits do not exist in isolation, they operate within a reinforcing loop alongside managerial and financial constraints. Furthermore, the low adoption of green building standards and energy-efficient design reflects a broader African higher education reality: environmental sustainability is often treated as a secondary luxury rather than a core infrastructure requirement (Eime et al., 2013). Facility managers and campus master planners must move away from ad-hoc repairs and adopt standardized, proactive maintenance protocols alongside clear environmental compliance benchmarks for all new and renovated sports venues.

Policy and Governance Framework Gaps

Weak policy frameworks and governance gaps represent a major organizational bottleneck impeding sustainable facility management in Ethiopian universities. The absence of clear, enforceable policies, weak national guidelines, and minimal integration into broader university strategic plans indicate that sports infrastructure is treated as an auxiliary service rather than an institutional asset. Poor interdepartmental coordination and missing accountability mechanisms mean that mismanaged funds or deferred maintenance frequently go uncorrected. While previous research has highlighted broad governance gaps in facility management (Appleby, 2018; Buys & Tonono, 2007; Elmualim et al., 2009; De Aldama et al., 2024), this study advances existing knowledge by demonstrating that governance failure directly undermines financial resource mobilization and operational leadership in public universities. The disconnect between institutional leadership and national regulatory bodies, such as the Ministry of Education, creates an environment where policies exist on paper but fail completely during field implementation (Ministry of Education, Ethiopia, 2020). Higher education policymakers and ministry officials must establish transparent governance structures, enforce standardized procurement rules, and embed sports facility sustainability metrics directly into university institutional accountability frameworks.

Leadership and Management Capacity

The findings regarding Leadership and Management Capacity Deficits reveal that leadership capacity deficits severely limit the strategic direction and daily operational quality of university sports facilities. Low prioritization by senior university executives, absent performance evaluation systems, and a lack of clear strategic vision lead to administrative inertia. In many Ethiopian universities, campus sports facility management suffers from a lack of executive commitment, leaving facility managers without the administrative authority or strategic clarity required to drive institutional improvement. These results validate prior literature on municipal and public sports services that links weak leadership and lack of formal qualifications to operational inefficiency (Liang, 2023; Mola & Shaw, 2024). However, our study extends these insights into the African university context by showing that leadership deficits cripple the adoption of global best practices, leaving institutions disconnected from contemporary management approaches. Without committed leadership, even well-intentioned strategic plans fail to deliver tangible operational progress. University executive leadership must formally integrate sports facility sustainability into institutional strategic plans, establish key performance indicators (KPIs) for managers, and provide executive mentorship in strategic resource management.

Financial Sustainability and Resource

Financial limitations represent the single most acute bottleneck affecting university sports facility sustainability in Ethiopia. Reliance on static, insufficient institutional budget allocations leave facilities vulnerable to inflation, operational shortfalls, and severe maintenance backlogs. The underutilization of alternative financial mechanisms, such as public private partnerships, facility rental income, alumni sponsorship, and corporate sponsorships, highlights a widespread lack of entrepreneurial resource mobilization across Ethiopian higher education. These findings confirm existing studies identifying single-source funding dependence as the primary threat to facility survival (Ngcobo et al., 2024; Asrifan et al., 2025; Rahman, 2021). Our study enriches this debate by proving that in developing nations, financial starvation is both a cause and a consequence of poor governance and administrative weakness; institutions with weak management systems fail to attract private investment, which further deepens financial dependency (Opoku & Lee, 2022). University finance directors and facility administrators must actively diversify revenue streams by developing commercial facility leasing frameworks, pursuing corporate partnerships, and establishing dedicated facility maintenance funds.

Sustainable Development of University Sports Facilities

The underdeveloped status of overall facility sustainability across Ethiopian universities underscores a systemic failure to balance the social, environmental, and financial dimensions of facility operations. University sports facilities are rarely managed with long-term strategic lifecycle considerations, limiting their ability to support institutional goals or meet modern sustainability benchmarks. This finding aligns with broader facility management literature calling for holistic, triple-bottom-line integration (Okoro, 2023; Mejía-Manzano et al., 2023; Katsaprakakis et al., 2023). However, our analysis provides new contextual clarity for African higher education: sporadic, isolated green projects (such as simple lighting retrofits) fail to achieve lasting impact unless embedded within a comprehensive institutional strategy (Da Silva et al., 2023; Taye et al., 2025). Sustainable facility development can only occur when physical, administrative, and financial dimensions are unified into a single operational framework.

Leisure Activities Participation Challenges

The study's findings regarding Leisure Activities Participation Challenges demonstrate that the significant connection between facility, quality and student participation in leisure activities (LAP) directly influences student well-being and campus life. While students rely heavily on sports participation for stress relief and mental resilience, inconsistent facility access, safety concerns, and time constraints create substantial barriers to widespread, equitable participation. These results extend prior work on campus recreation access (Gonçalves Maciel, et al., 2026; Maciel et al., 2025; Abrahamson et al., 2025) by establishing that user participation is a primary indicator of institutional sustainability in African higher education. When facilities deteriorate, student engagement drops, diminishing the social value of the infrastructure and weakening the institutional justification for future budget allocations. Furthermore, active stakeholder collaboration, involving coaches, facility managers, and student groups, is essential for building inclusive sports environments that support health and social cohesion across diverse student bodies (Filho et al., 2020; Kaur, 2025).

Empirical Synthesis and Integrative Model

The multiple linear regression model provides robust empirical evidence that maintaining university sports facilities for long-term viability cannot be achieved through fragmented, ad-hoc interventions in Ethiopian higher education. Physical infrastructure ($\beta = 0.712$, $p < 0.001$) and financial sustainability ($\beta = 0.653$, $p < 0.001$) emerged as the primary determinants, establishing that well-maintained physical assets and diversified, reliable funding streams constitute the fundamental pillars of operational success. However, these two pillars do not function in isolation; they are supported by critical organizational drivers, including leadership capacity ($\beta = 0.387$, $p = 0.003$), enforceable policy and governance structures ($\beta = 0.312$, $p < 0.001$), and active student leisure participation ($\beta = 0.298$, $p = 0.001$).

Together, these five domains account for 71.0% of the variance in sustainable facility development (R2 = 0.710, Adjusted R2 = 0.697, F(5,94) = 46.1, $p < 0.001$). Standard diagnostic tests confirmed the statistical validity of the model, demonstrating an absence of multicollinearity (Tolerance: 0.485–0.621; VIF: 1.610–2.062), normally distributed residuals (Kolmogorov-Smirnov $p = 0.200$), and compliance with homoscedasticity assumptions.

Evaluating these empirical patterns corroborates and extends prior research (Varmus et al., 2023; Chapleo, 2010; Uvinya et al., 2025; Vega et al., 2019) by establishing an integrated framework tailored to higher education institutions in Sub-Saharan Africa. The data demonstrates that physical facilities, institutional management, and user engagement are deeply interdependent. Capital investments in sports infrastructure deteriorate rapidly in the absence of clear policy enforcement and sustainable revenue streams. Conversely, strategic leadership vision remains unfulfilled if facilities remain underfunded or physically inaccessible to the student body. Crucially, active student recreation (LAP) serves as an essential operational feedback loop for the entire system: consistent user engagement reinforces the institutional value of these facilities, maintains institutional momentum, and justifies sustained capital investment over time.

Future Research and Practical Implications

Future studies should extend the model by examining environmental efficiency, stakeholder partnerships, and technology integration in sports facility management. Since infrastructure was the strongest predictor of sustainability, research should explore cost-effective renovation strategies, innovative designs, and sustainable financing approaches, including sponsorships, public-private partnerships, commercialisation, and sport-tourism linkages. Longitudinal, cross-country, and qualitative studies involving managers, administrators, coaches, and students are recommended to improve understanding and generalisability. Universities should strengthen preventive maintenance systems, adopt green facility standards, and establish professional sports facility management units. Higher education leaders should diversify funding through partnerships, facility revenue generation, sponsorships, and alumni support while improving policy implementation and staff capacity development.

Study Limitations

This study has several limitations. First, the cross-sectional design limits causal conclusions; therefore, the findings should be interpreted as predictive rather than causal. Second, the study relied on self-reported survey data, which may not fully represent actual facility conditions, financial resources, or policy implementation. Future research should integrate objective facility audits, financial records, policy documents, longitudinal designs, and cross-country comparisons to improve the validity and generalisability of the findings.

Conclusion

The sustainable development of university sports facilities in Ethiopia depends on strong physical infrastructure, effective financial resource mobilisation, capable leadership, sound governance, and active student leisure participation. Universities should prioritise sustainable facility development, strengthen management systems, diversify funding sources, and encourage student engagement. Long-term sustainability requires an integrated approach combining infrastructure investment, governance improvement, innovative financing, and proactive participation strategies.

Acknowledgment

The authors express their gratitude to all participants who contributed to the study.

Conflicts of interest

The authors declare no conflicts of interest.

References

Shih, Y.-H., Hsu, M. C., & Chang, C. L. (2025). Sustainability transformations in higher education: Global perspectives on the challenges and solutions. *International Journal of Education & Humanities, 5*(1), 126–139. https://doi.org/10.58557/(ijeh).v5i1.290

Zheng, L.-H., Guo, S.-T., Feng, X.-W., Xu, Y.-Y., Mohd Nor, M. N., & Zainal Abidin, N. E. (2024). Sustainable development between sports facilities and ecological environment based on the dual carbon background. *Scientific Reports,* 14, 22692. https://doi.org/10.1038/s41598-024-73507-4

Ministry of Education, Ethiopia. (2020). *Annual higher education statistical report.* Addis Ababa: Ministry of Education.

Dachew, B. A., Bisetegn, T. A., & Gebremariam, L. T. (2015). Physical activity and health status among Ethiopian university students. *Ethiopian Journal of Health Sciences, 25*(3), 217–224.

Piccerillo, L., Misiti, F., & Digennaro, S. (2023). Assessing the environmental impact of a university sport event: The case of the 75th Italian National University Championships. *Sustainability, 15*(3), 2267. https://doi.org/10.3390/su15032267

Stinnett, B., & Gibson, F. (2016). Sustainable facility development: Perceived benefits and challenges. *International Journal of Sustainability in Higher Education, 17*(5), 601–615. https://doi.org/10.1108/IJSHE-09-2014-0133

Lastra, P. (2023). *Improving energy stewardship at Michigan Technological University's athletic complexes (*Doctoral dissertation, Michigan Technological University). https://doi.org/10.37099/mtu.dc.etdr/1477

Yan, M.-R., & Chi, H.-L. (2024). *Sustainable system development and strategic framework for ESGI enterprise transformation in large multi-purpose sports venues.* Preprint. https://doi.org/10.20944/preprints202403.1132.v1

Chikafalimani, S. H. P., Kibwami, N., & Moyo, S. (2021). Management of facilities at public universities in Africa: Current challenges and the way forward. *Real Estate Management and Valuation, 29*(1), 21–29. https://doi.org/10.2478/remav-2021-0003

Eime, R. M., Young, J. A., Harvey, J. T., Charity, M. J., & Payne, W. R. (2013). A systematic review of the psychological and social benefits of participation in sport for adults: Informing development of a conceptual model of health through sport. *International Journal of Behavioral Nutrition and Physical Activity, 10*(1), 135. https://doi.org/10.1186/1479-5868-10-135

Sentie, W. A., Fadoju, A. O., & Kebede, D. N. (2024). Examining the impact of organizational and management variables on sport service delivery in Ethiopia. *Retos,* 61, 927–935. https://doi.org/10.47197/retos.v61.107638

Cuhadar, A., Karakullukcu, O. F., Koc, M. C., Ozbey, O., Sahinler, Y., Gumus, H., & Mola, D. W. (2026). The mediating role of academic perfectionism in explaining the relationship between self-esteem and life satisfaction in sport sciences students. *BMC Psychology, 14*(1), 390. https://doi.org/10.1186/s40359-026-04188-z

McCullough, B., Orr, R., & Kellison, T. (2020). Sustainability and NCAA athletics: The status of STARS reporting. *Journal of Intercollegiate Sport, 13*(1), 1–21.

Talay, B., Bostan, K., Pekel, A., & Çetin, A. (2025). Sustainability practices in sports facilities and managers' perceptions. *KOSALB International Journal of Human Movements Science, 4*(1), 49–64. https://doi.org/10.70736/2958.8332.kosalb.46

Kiflu Adane, A., Wase Mola, D., & Adili, D. (2025). Challenges of inclusive physical education for students with disabilities in Ethiopian secondary schools: A mixed-methods study. *International Journal of Kinesiology and Sports Science, 13*(4), 24–31. https://doi.org/10.7575/aiac.ijkss.v.13n.4p.24

Bergsgard, N. A. (2025). Does the power structure in local facility processes hinder sustainable solutions? Case studies from Norway (and Denmark). *Annals of Leisure Research, 28*(1), 1–18. https://doi.org/10.1080/11745398.2025.2449838

He, C., Huang, Z., & Ye, L. (2021). Research on the sustainable development of physical education students' exercise in colleges and universities. *Revista Brasileira de Medicina do Esporte, 27*(4), 421–424. https://doi.org/10.1590/1517-8692202127042021_0089

Moronkeji, D. B., Ayeni, D., & Adewunmi, A. E. (2022). Participation and hindrances to students' recreation at selected universities in Ondo State, Nigeria. *Journal of Social and Development Sciences, 13*(1), 26–33. https://doi.org/10.22610/jsds.v13i1(s).3285

Mola, D. W., & Shaw, D. (2024). Significance of managerial skills for talent development in Ethiopian athletes. *International Journal of Kinesiology and Sports Science, 12*(2), 21–29. https://doi.org/10.7575/aiac.ijkss.v.12n.2p.21

Maciel, M. G., Madariaga Ortuzar, A., Martins, J. C. de O., Gutierrez, L. A. Á., Bendikova, E., Assis, K. L., Reis, R. F., Mola, D. W., & Uvinha, R. R. (2024). Leisure styles in adults: Changes throughout the COVID-19 pandemic. *Social Sciences, 13*(12), 702. https://doi.org/10.3390/socsci13120702

Brunton, J., & St Quinton, T. (2021). Applying stage-based theory to engage female students in university sport. *Journal of Human Sport and Exercise, 16*(1), 16–30. https://doi.org/10.14198/JHSE.2021.161.02

Abrahamson, K., Ross-Cypcar, S. M., Healy, S., Erwin, H., & Shapiro, D. R. (2025). An inquiry into the accessibility of campus recreation for college students with disabilities. *Recreational Sports Journal, 49*(1), 1–12. https://doi.org/10.1177/15588661251372108

Miller, J. J., & Croft, J. (2022). The influence of university recreation centers on student return and retention during COVID-19. *Recreational Sports Journal, 46*(1), 72–83. https://doi.org/10.1177/15588661221097701

Maciel, M. G., Uvinha, R. R., & Madariaga Ortuzar, A. (2025). *Estilos de ocio y salud mental de estudiantes universitarios iberoamericanos* [Leisure styles and mental health among Ibero-American university students]. In S. González Morcillo & I. Lazcano Quintana (Eds.), El papel transformador del ocio: Desde la autogestión en la juventud hasta la producción del espacio (pp. 27–39). Dykinson.

Freitas, H., Henriques, S., Uvinha, R. R., Lusby, C., & Romera, L. (2020). Alcohol consumption and night-time leisure among Brazilian university students. *International Journal of the Sociology of Leisure, 3*(4), 389–399. https://doi.org/10.1007/s41978-020-00069-w

Zvapano, K. (2017). An assessment of the state of sport infrastructural facilities in universities in Zimbabwe. *International Journal of Sport, Exercise and Health Research, 1*(2), 70–75. https://doi.org/10.31254/sportmed.1205

Iheanacho, S. B. C., Ikpeme, E. E., & Saba, I. A. (2013). An assessment on provision of recreational facilities in Nigerian universities in the 21st century. *Journal of Public Administration and Governance, 3*(1), 95–101. https://doi.org/10.5296/jpag.v3i1.3380

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approach* (4th ed.). Sage Publications.

Covell, D., & Walker, S. (2013). *Managing sport organizations: Responsibility for performance* (3rd ed.). Routledge. https://doi.org/10.4324/9780203550281

Goslin, A. E., Sere, M. G. O., & Kluka, A. (2015). Management capacity at sport and recreation facilities at the local government level. *African Journal for Physical, Health Education, Recreation and Dance, 21*(4:1), 1290–1303.

Mohd Aznan, E. A., Kamaruddin, H. K., Abu Bakar, A. H., Ahmad, M. F., & Safwan, N. S. Z. (2023). Assessing perceived risk management and satisfaction in university student's sports facility experience: A cross-sectional study. *Jurnal Sains Sukan & Pendidikan Jasmani, 12*(2), 39–47. https://doi.org/10.37134/jsspj.vol12.2.6.2023

Amireault, S., & Godin, G. (2015). The Godin-Shephard Leisure-Time Physical Activity Questionnaire: Validity evidence supporting its use for classifying healthy adults into active and insufficiently active categories. *Perceptual and Motor Skills, 120*(2), 604–622. https://doi.org/10.2466/03.27.PMS.120v19x7

Burkhardt, L., & Schupp, J. (2019). Wachsendes ehrenamtliches Engagement: Generation der 68er häufiger auch nach dem Renteneintritt aktiv. *DIW Wochenbericht, 86*(42), 765–773. https://doi.org/10.18723/diw_wb:2019-42-1

Piggott, L. V., & Pike, E. C. J. (2020). 'CEO equals man': Gender and informal organisational practices in English sport governance. *International Review for the Sociology of Sport, 55*(7), 1009–1025. https://doi.org/10.1177/1012690219865980

Koenigstorfer, J. (2023). Dealing with great challenges via rigorous and relevant empirical sport management research. *Frontiers in Sports and Active Living,* 5. https://doi.org/10.3389/fspor.2023.1180710

Vega, P., Pérez-González, B., Fernández-Luna, Á., & Burillo, P. (2019). Gender diversity on the steering committees of Spanish sports federations: Implications for economic and operational performance. *Apunts: Educación Física y Deportes,* 137, 115–128. https://doi.org/10.5672/APUNTS.2014-0983.ES.(2019/3).137.09

Hassen, M. (2025). Rethinking the physical environment and pedagogical approaches in Ethiopian higher education for real-world readiness. *Science Research, 13*(3), 32–42. https://doi.org/10.11648/j.sr.20251303.11

Rahman, M. M. (2021). BIM enabled sustainable facility management. *International Journal of Integrated Engineering, 13*(7), 101–107. https://doi.org/10.30880/ijie.2021.13.07.013

Naseri, A. S., Saffari, L., & Majedi, N. (2025). Green operations in sports facilities: A review of technological implementation. *KMAN Journal, 3*(1), 74–91. https://doi.org/10.61838/kman.aitech.3.1.8

Appleby, P. (2018). *Sustainable facilities management* (pp. 253–265). Routledge Handbooks Online. https://doi.org/10.1201/9781315622750-16

Buys, F., & Tonono, E. (2007). Reasons for the transformation of facilities management in the public sector. *Acta Structilia, 14*(2), 76–90. https://journals.co.za/content/struct/14/2/EJC110014

Elmualim, A., Czwakiel, A., Valle, R., Ludlow, G., & Shah, S. (2009). The practice of sustainable facilities management: Design sentiments and the knowledge chasm. *Gestão & Tecnologia de Projetos, 4*(2), 143. https://doi.org/10.4237/GTP. V4I2.137

De Aldama Ortuzar, I. M., Cayero, R., Terreras, B. T., Bravo, J. G., & Irigoyen, J. Y. (2024). Situación actual y retos de futuro de los servicios deportivos municipales: Percepción de los responsables deportivos. *Retos,* 55, 490–498. https://doi.org/10.47197/retos.v55.102025

Liang, W. (2023). Financial sustainability among higher education institutions in the United States. *Journal of Contemporary Educational Research, 7*(12), 277–281. https://doi.org/10.26689/jcer.v7i12.5822

Ngcobo, X. M., Marimuthu, F., & Stainbank, L. J. (2024). Revenue sourcing for the financial sustainability of a university of technology: An exploratory study. *Cogent Education, 11*(1), 2295173. https://doi.org/10.1080/2331186X.2023.2295173

Asrifan, A., Sudirman, R., & Karundeng, D. R. (2025). *Building financial resilience: Strategic partnerships between higher education and industry.* In Financial survival in higher education (pp. 211–244). IGI Global. https://doi.org/10.4018/979-8-3373-0407-6.ch008

Opoku, A., & Lee, J. (2022). The future of facilities management: Managing facilities for sustainable development. *Sustainability, 14*(3), 1705. https://doi.org/10.3390/su14031705

Okoro, C. S. (2023). Sustainable facilities management in the built environment: A mixed-method review. *Sustainability, 15*(4), 3174. https://doi.org/10.3390/su15043174

Mejía-Manzano, L. A., Vázquez-Villegas, P., Smith, A. R., Soeiro, A., Kalman, A., Atabarut, T., Otaduy-Rivera, N., Membrillo-Hernández, J., & Caratozzolo, P. (2023). An exploratory study examining the key aspects and actions for universities to achieve high sustainability rankings. *Sustainability, 15*(5), 4165. https://doi.org/10.3390/su15054165

Katsaprakakis, D. A., Papadakis, N., Giannopoulou, E., Yiannakoudakis, Y., Zidianakis, G., Dakanali, E., Stavrakakis, G. S., & Kartalidis, A. (2023). Rational use of energy in sports centres to achieve net zero: The SAVE Project (Part A). *Energies, 16*(10), 4040. https://doi.org/10.3390/en16104040

Da Silva, L. A., De Aguiar Dutra, A. R., & De Andrade Guerra, J. B. S. O. (2023). Decarbonization in higher education institutions as a way to achieve a green campus: A literature review. *Sustainability, 15*(5), 4043. https://doi.org/10.3390/su15054043

Taye, A. G., Bouchh, H., Rahman, M. H., Uvinha, R. R., Legesse, A. K., Adili, D., & Mola, D. W. (2025). Determining the influence of coaches, parents, and stakeholders on the development of youth football: A comprehensive analysis. *Physical Education Theory and Methodology, 25*(3), 556–565. https://doi.org/10.17309/tmfv.2025.3.11

Gonçalves Maciel, M., Ricci Uvinha, R., Roque Andrade, D., Bendíková, E., Mola, D. W., & Madariaga Ortuzar, A. (2026). Physical activity as leisure: legal/human rights, public health, and phenomenological perspectives. *Annals of Leisure Research,* 1–17. https://doi.org/10.1080/11745398.2026.2691038

Leal Filho, W., Salvia, A. L., Frankenberger, F., Akib, N. A. M., Sen, S. K., Sivapalan, S., Novo-Corti, I., Venkatesan, M., & Emblen-Perry, K. (2020). Governance and sustainable development at higher education institutions. *Environment, Development and Sustainability, 23*(4), 6002–6020. https://doi.org/10.1007/s10668-020-00859-y

Kaur, M. P. (2025). Policy, governance, and pedagogical strategies for integrating sports into sustainable development: A framework for equity, diversity, and gender inclusivity. *Communications on Applied Nonlinear Analysis, 32*(ICMASD-2025), 2194–2204. https://doi.org/10.52783/cana.v32.5615

Chapleo, C. (2010). What defines "successful" university brands? *International Journal of Public Sector Management, 23*(2), 169–183. https://doi.org/10.1108/0951355101102251

Uvinha, R. R., de Ridder, J. H., & Gushchina, Y. S. (2025). Advancing health through physical activity and leisure in BRICS Plus nations: A report from the BRICSCESS Plus Association. *Sports Medicine: Research and Practice, 15*(2), 5–11. https://doi.org/10.47529/2223-2524.2025.2.7

Критичні проблеми сталого розвитку спортивної інфраструктури університетів в африканському контексті: дозвілля, політика, інфраструктура та лідерство

Дессалегн Васе Мола[1,2ABCDE], Рікардо Річчі Увінья[1ABCDE], Абді Гудета Тайє[3ABCDE], Абдета Байіса Мекесса[2ABCDE], Мустафа Джан Коч[4ABCDE]

[1]Університет Сан-Паулу
[2]Університет Амбо
[3]Університет Дамбі-Долло
[4]Університет Мерсіна

Авторський вклад: A – дизайн дослідження; B – збір даних; C – статаналіз; D – підготовка рукопису; E – збір коштів

Реферат. Стаття: 13 с., 12 табл., 2 рис., 57 джерел.

Передумови. Сталий розвиток спортивної інфраструктури університетів має вирішальне значення для спортивних результатів, залучення студентів і довгострокової інституційної сталості. В університетах Ефіопії цей розвиток стримується неналежною інфраструктурою та її технічним обслуговуванням, слабкою реалізацією політики, обмеженими фінансовими ресурсами, низьким рівнем участі в дозвіллєвій діяльності та недостатнім потенціалом керівництва й управління.

Мета дослідження. Метою цього дослідження було вивчити ключові чинники, що впливають на сталий розвиток спортивної інфраструктури університетів Ефіопії, зосередивши увагу на інфраструктурі, політиці та врядуванні, потенціалі керівництва й управління, фінансовій стійкості та участі в дозвіллєвій діяльності.

Матеріали та методи. Проведено крос-секційне опитування 100 цілеспрямовано відібраних спортсменів, тренерів, менеджерів спортивних об'єктів, адміністраторів, фінансових працівників і фахівців із питань політики державних університетів Ефіопії. Дані збирали за допомогою валідованого опитувальника з 42 пунктів, що охоплював шість змінних. Експлораторний факторний аналіз (EFA) підтвердив конструктну валідність (KMO = 0,812; пояснено 68,4% дисперсії), а надійність шкали була високою (загальний $\alpha = 0{,}819$). Статистичну потужність перевірено за допомогою G*Power ($1 - \beta > 0{,}99$). Застосовано описову статистику, кореляцію Пірсона та множинний лінійний регресійний аналіз із діагностикою регресійної моделі (VIF та аналіз залишків).

Результати. Середні значення засвідчили низький або помірний рівень за показниками інфраструктури спортивних об'єктів ($M = 2{,}74$), участі в дозвіллєвій діяльності ($M = 2{,}69$), політики та врядування ($M = 2{,}62$), управлінського потенціалу ($M = 2{,}61$), сталого розвитку спортивних об'єктів ($M = 2{,}56$) та фінансової стійкості ($M = 2{,}50$). Множинний регресійний аналіз показав, що п'ять предикторних доменів сукупно пояснювали 71,0% дисперсії сталого розвитку спортивних об'єктів ($R^2 = 0{,}710$, $F(5, 94) = 46{,}01$, $p < 0{,}001$). Інфраструктура спортивних об'єктів ($\beta = 0{,}712$, $p < 0{,}001$) та фінансова стійкість ($\beta = 0{,}653$, $p < 0{,}001$) були найсильнішими предикторами, за ними йшли управлінський потенціал ($\beta = 0{,}387$, $p = 0{,}003$), політика та врядування ($\beta = 0{,}312$, $p < 0{,}001$) і участь у дозвіллєвій діяльності ($\beta = 0{,}298$, $p = 0{,}001$).

Висновки. Результати свідчать, що спортивна інфраструктура університетів Ефіопії стикається із суттєвими структурними та системними проблемами. Досягнення її довгострокової сталості потребує комплексного підходу: підвищення якості інфраструктури, забезпечення диверсифікованого фінансування, посилення управлінського потенціалу, ефективної реалізації політики врядування та сприяння залученню студентів до дозвіллєвої діяльності. Комплексне вирішення цих ключових питань є необхідним для створення стійкої та належним чином утримуваної спортивної інфраструктури в закладах вищої освіти.

Ключові слова: африканський контекст, спортивна інфраструктура університетів, дозвіллєва діяльність, сталість, врядування, лідерство, фінансова стійкість, інфраструктура.

Information about the authors:

Mola, Dessalegn Wase: dessalegn.wasie@gmail.com; https://orcid.org/0000-0001-9688-0775; 1Postdoctoral Research Fellow, Graduate Program in Physical Activity Sciences, School of Arts, Sciences and Humanities, University of Sao Paulo, Sao Paulo, Brazil; & 2Department of Sports Science, College of Natural and Computational Sciences, Ambo University, Ambo, Ethiopia.

Uvinha, Ricardo Ricci: uvinha@usp.br; https://orcid.org/0000-0003-2936-9453; Graduate Program in Physical Activity Sciences, School of Arts, Sciences and Humanities, University of Sao Paulo, Sao Paulo, Brazil.

Taye, Abdi Gudeta: abdigudeta48@gmail.com; https://orcid.org/0009-0000-7980-7074; Department of Sports Science, College of Natural and Computational Sciences, Dambi Dollo University, Dambi Dollo, Ethiopia.

Mekesa, Abdeta Bayisa: amekessa@gmail.com; https://orcid.org/0000-0003-1729-0836; Department of Sports Science, College of Natural and Computational Sciences, Ambo University, Ambo, Ethiopia.

Koc, Mustafa Can: cankoc_01@hotmail.com; https://orcid.org/0000-0002-3007-5367; Faculty of Sports Sciences, Mersin University, Mersin, Turkiye.

Cite this article as: Mola, D. W., Uvinha, R. R., Taye, A. G., Mekesa, A. B., & Koc, M. C. (2026). Critical Challenges to the Sustainable Development of University Sports Facilities in the African Context: Leisure, Policy, Infrastructure, and Leadership. *Journal of Learning Theory and Methodology, 7*(2), 138-150. https://doi.org/10.17309/jltm.2026.7.2.08

Received: 24.07.2026. Accepted: 13.08.2026. Published: 20.08.2026

ISSN 2708-7581 (Online)
ISSN-L 2708-7573

Journal of Learning Theory and Methodology

Scientific journal

August 2026
Volume 7
Number 2

Відповідальний за випуск О. М. Худолій
Комп'ютерна верстка М. О. Худолій
Коректор Є. Б. Бланк

Зареєстровано в Національній раді України з питань телебачення і мовлення:
рішення No 1751 від 23.05.2024 р., протокол No 16, оприлюднено 23 травня 2024 р., ідентифікатор медіа – R40-04690

https://webportal.nrada.gov.ua/wp-content/uploads/2024/05/R-2024-01751.pdf

Засновник і видавець — ТОВ «ОВС».
Адреса редакції: https://www.ltmjournal.com.Тел.: (067) 578-40-08. E-mail: editor-in-chief@ltmjournal.com

Підписано до друку 28.08.2026. Формат 60×84 1/8. Електронне видання. Гарнітура Таймс. PDF формат.
Ум. друк. арк. 6,989. Обл.-вид. арк.7,25. Вид. № 7(1)-2026. Зам. № 56. Тираж 300 прим. Ціна договірна.

ТОВ «ОВС» Україна, 61003 Харків, пл. Конституції, 18, к. 11.
Свідоцтво Держкомінформу України Серія ДК № 331 від 08.02.2001 р.