Students' AI Dependency in 3R's: Questionnaire Construction and Validation

Andie Tangonan Capinding

College of Education, Faculty, Nueva Ecija University of Science and Technology-Gabaldon Campus, Gabaldon, Philippines
Email: andiecapinding103087@gmail.com (A.T.C.)
Manuscript received June 6, 2024; revised July 15, 2024; accepted August 21, 2024; published November 12, 2024

***Abstract*—The integration of Artificial Intelligence (AI) into education has introduced both groundbreaking opportunities and concerns. Among these concerns is the extent of students' reliance on AI in the realms of reading, writing, and numeracy/arithmetic (3Rs). While existing instruments delve into the broader impact of AI, they exhibit certain limitations. Consequently, this research endeavors to develop and validate a specialized questionnaire tailored to assess students' dependency on AI in the 3Rs. The process includes interviews with student groups, consultations with professionals in the education sector, face validation, content validation, exploratory factor analysis, confirmatory factor analysis, Rasch analysis, and reliability testing to navigate the construction and validation of the instruments. Initial item identification involved a 45-item questionnaire distributed across three constructs, derived from qualitative interviews with students and experts. The survey received a total of 727 responses. Post EFA, nine items were eliminated due to their failure to achieve a loading factor of 0.5, and certain items exhibited cross-loadings. Subsequent Rasch analysis affirmed the construct validity of the instruments, prompting the removal of three additional items. The resulting 33-item questionnaire, divided into three constructs—Reading (10 items), Writing (11 items), and Numeracy/Arithmetic (12 items)—emerges as a validated and reliable tool for measuring students' dependency in the 3Rs. The author confirms the validity and reliability of the questionnaire. Future research should focus on longitudinal studies to assess how AI dependency evolves over time and impacts educational outcomes.**

***Keywords*—Artificial Intelligence (AI) dependency, exploratory factor analysis, Rasch analysis, reliability test, validity test**

I. Introduction

The pervasive influence of Artificial Intelligence (AI) has fundamentally transformed education, introducing unprecedented possibilities and challenges. In this era of rapid technological progression, educational environments are increasingly leveraging AI to augment learning experiences. Chen *et al.* [1] report the widespread adoption and utilization of AI in education, especially within educational institutions and in various manifestations. Moreover, according to Almusaed *et al*. [2], the utilization of AI has the potential to transform hybrid education by boosting autonomy for both students and instructors, thereby cultivating a more dynamic and participatory learning atmosphere. While the integration of AI offers advantages such as tailored instruction, adaptive learning modules, and improved accessibility, it also raises concerns about potential drawbacks, particularly regarding students' reliance on these technologies. Overreliance on AI and online tools, including the use of chatbots for assignments, is criticized for potentially hindering students' critical thinking, problem-solving abilities, and overall creativity [3].

On the other hand, the 3Rs—Reading, Writing, and Numeracy/Arithmetic—have long been regarded as the cornerstones of education. These fundamental skills not only serve as academic building blocks but also play a pivotal role in shaping cognitive abilities and preparing individuals for various aspects of life. Furthermore, according to Millacci [4], the significance of reading, writing, and mathematics lies in the essential development of cognitive skills, which are imperative for lifelong learning. The advent of AI introduces both promise and challenges to the cultivation of these skills.

In the context of Reading, AI-powered tools offer personalized reading materials, language translation, and comprehension aids. As an illustration, Srinivasan and Murthy [5] demonstrated that the incorporation of AI into students' reading resulted in a 20-40% improvement in learning. Moreover, Hsiao and Chang [6] revealed in their research that the utilization of three AI-driven tools, namely Linggle Write, Linggle Read, and Linggle Search, resulted in an enhanced optimal experience during student-centered presentations compared to teacher-centered lectures. Additionally, language proficiency was identified as a predictor for both semester grades and assignment quantity. While these applications enhance accessibility and individualized learning experiences, questions arise about potential detriments to critical reading skills and a reliance on automated comprehension. Ahmad *et al*. [7] demonstrated that artificial intelligence has a substantial effect on diminishing human decision-making, fostering laziness, and influencing security and privacy, with the study revealing percentages of 68.9% for laziness, 68.6% for personal privacy and security concerns, and 27.7% for the erosion of decision-making in both Pakistani and Chinese societies. Moreover, the study by Buçinca *et al*. [8] emphasized that individuals relying on AI-powered decision support tools tend to excessively trust the AI's suggestions, even in cases where the suggestions are incorrect. This raises important questions about the balance between leveraging AI for support and maintaining critical thinking skills in the face of technological advancements.

AI's impact on Writing is evident in grammar correction, content generation, and style suggestions. According to Malik *et al*. [9], students express a positive reception of AI-powered writing tools, acknowledging their benefits in grammar checks, plagiarism detection, language translation, and essay outlines. Furthermore, the research highlights that the integration of AI enhances students' writing abilities, self-efficacy, and understanding of academic integrity. While these tools can enhance writing efficiency, concerns surface regarding their influence on creativity, originality, and the development of a distinct voice [10]. Furthermore, Huang and Tan [11] contended that an overreliance on AI technologies could diminish one's capacity for creative and critical thinking, while also impeding the ability to autonomously evaluate the quality of writing. Likewise, students develop a reluctance towards writing, leading them to rely on AI text-generation tools such as autocomplete and predictive texting, which seamlessly integrate into writing, providing suggestions for online searches and text messages; although students acknowledge the support, surveys express apprehensions about potential drawbacks, including the potential loss of spelling skills and the allure of opting for an easier path through spellcheck and AI software [12].

In the field of Numeracy/Arithmetic, AI presents opportunities for enhancing problem-solving, mathematical modeling, and adaptive learning platforms. According to Qawaqneh *et al.* [13], artificial intelligence-driven virtual laboratories can not only facilitate these aspects but also have the potential to elevate students' motivation in learning mathematics. Furthermore, Cunska [14] adds to this perspective by asserting that AI, with its capacity to supplement or replace existing learning approaches, has the potential to transform mathematics into a more accessible and meaningful subject in the future. However, the convenience offered by automated calculations raises concerns regarding the potential decline in students' manual computation skills and their understanding of mathematical concepts. Rane [15] underscores the risk of diminishing critical thinking and problem-solving skills due to an overreliance on technology, exemplified by relying solely on ChatGPT for mathematical solutions. This reliance may lead to a surface-level understanding, hindering the development of deep mathematical insights by circumventing active engagement in the learning process. Moreover, Wardat *et al.* [16] contribute to the discussion by suggesting that artificial intelligence tools, like ChatGPT, have limitations in comprehending specific mathematical concepts, such as geometry, and may not proficiently rectify misconceptions. This emphasizes the importance of balancing the advantages of AI in mathematics education with the need for maintaining essential manual skills and a comprehensive understanding of mathematical principles.

This discussion aims to highlight the intricate dynamics between AI and the foundational 3Rs, emphasizing the need for a nuanced examination of students' dependency on AI within each domain. By exploring these intersections, the research endeavors to contribute insights that inform educational practices and policies, ensuring a balanced integration of AI in the pursuit of holistic learning. Thus, this study aims to construct and validate a comprehensive questionnaire designed to assess students' AI dependency specifically within the realms of the 3Rs. The 3Rs, long considered the cornerstone of education, play a foundational role in academic development and are crucial for cognitive growth. As AI tools become more integrated into the learning process, it is crucial to gauge the impact of this integration on students' learning behaviors, preferences, and perceptions.

Presently, numerous research endeavors aim to assess the extent of students' reliance on technology and Artificial Intelligence. Choliz [17] and Chóliz *et al.* [18] have devised a mobile phone dependence questionnaire. Similarly, Emelin *et al.* [19] introduced a questionnaire addressing psychological consequences related to technology, encompassing internet usage, mobile phones, and computers. Furthermore, Capinding [20] designed and validated a tool for gauging the influence of artificial intelligence (AI) in education. One of the study's facets involved assessing AI dependency, but the questionnaire lacked specificity regarding the individual factors contributing to AI dependency. Moreover, Martínez-Córcoles *et al.* [21] have created and validated the Technophobia and Technophilia Questionnaire (TTQ). In the realm of medicine, Muñoz-Neira *et al.* [22] have developed and validated an updated version of the Activities of Daily Living Questionnaire (ADLQ), which incorporates an Information and Communication Technology (ICT) subscale. Additionally, Kieslich *et al.* [23] conducted an assessment using the Treats of AI (TAI) scale across three distinct AI domains—medical treatment, job recruitment, and loan origination.

The current research landscape, as evidenced by studies conducted by Choliz [17], Chóliz *et al.* [18], Emelin *et al.* [19], Martínez-Córcoles *et al.* [21], Muñoz-Neira *et al.* [22], and Kieslich [23], underscores a concerted effort to assess students' reliance on technology in various domains. However, a conspicuous gap exists in the literature concerning the measurement of students' dependency on Artificial Intelligence (AI) specifically in fundamental academic areas such as reading, writing, and numeracy/arithmetic. In light of the pervasive influence of AI in education, the absence of a dedicated questionnaire for evaluating students' dependence on AI in these core skills is noteworthy. This gap is not just an academic oversight; it has practical implications for understanding and guiding students' interactions with educational technology.

The development of such an instrument is imperative, considering the rapid integration of AI tools into educational settings. This specialized questionnaire would not only offer insights into the evolving nature of education but also provide a nuanced understanding of students' interactions with AI in acquiring essential academic competencies. Thus, creating a tailored questionnaire for measuring students' dependency on AI in reading, writing, and arithmetic is not only justified but essential for comprehensively exploring the dynamic relationship between students and educational technology. This research specifically focused on developing items for three constructs of AI dependency, validating their face and content validity, and establishing their construct validity. Subsequently, it measured the psychometric properties of these items. Specifically, the research aimed to: (a) develop initial questionnaire items through qualitative interviews with students and teachers; (b) conduct face validity assessments; (c) perform content validation; (d) carry out Exploratory and Confirmatory Factor Analyses; (e) conduct Rasch Analysis; and (f) test the reliability of the instrument.

II. Methodology

A. Research Design

The objective of this research initiative was to develop and authenticate a research questionnaire designed to evaluate students' reliance on Artificial Intelligence (AI) in the realms of Reading, Writing, and Numeracy/Arithmetic (3Rs). The procedure involved a series of meticulous steps for the construction and validation of the instruments, including interviews with student groups, consultations with professionals in the education sector, face validation, content validation, exploratory factor analysis, confirmatory factor analysis, Rasch analysis and reliability testing.

B. Interviews with Group of Students

In the initial phase of this investigation into students' reliance on Artificial Intelligence (AI) in the domains of reading, writing, and numeracy/arithmetic, an unstructured approach was adopted to explore the qualitative dimensions shared by a carefully selected group of students. The unstructured interview questions were reviewed and validated by two English professors. They possess expertise in qualitative research and the formulation of questionnaires. The primary aim was to grasp the diverse perspectives of students regarding the integration of AI tools into their academic pursuits. Upholding ethical considerations as a priority, necessary permissions were obtained from school authorities and parents or guardians to ensure voluntary participation and maintain confidentiality. A diverse group of 4-6 students within the targeted age range was chosen, taking into account factors such as academic performance, socio-economic background, and gender for a holistic viewpoint. The student group was purposefully chosen, and it includes those who have utilized or tried to employ AI in their studies. The group of students represented various programs within the university.

To encourage open and insightful conversations, discussions were initiated with a probing: "Can you share your experiences and thoughts on how Artificial Intelligence is involved in your learning process?" According to Birt [24], probing questions are crafted with the intention of enhancing the knowledge and comprehension of both the individual posing the question and the individual responding. This open-ended question aimed to elicit a comprehensive exploration of their reliance on AI tools. The absence of a predetermined set of questions allowed the natural flow of the conversation, enabling students to express their experiences freely.

Before the interviews, a pilot test was conducted to refine the approach. Subsequently, interviews were scheduled, providing a comfortable and distraction-free environment. Sessions commenced with an introduction, outlining the study's purpose, and obtaining informed consent from both students and their parents or guardians.

During the interviews, a focus was maintained on the central question, allowing students to provide unstructured responses about their dependence on AI in learning. Responses were accurately documented through note-taking and audio recording, with explicit consent. Post-interview, data analysis was conducted, emphasizing the identification of common themes, patterns, and insights associated with AI dependency in reading, writing, and numeracy/arithmetic.

C. Consultations with Professionals in The Education Sector

Following student interviews, a critical phase of this research involved engaging education professionals, including professors, teachers, and instructors, to delve into the perspective and experiences of their students regarding the use of Artificial Intelligence (AI) in studying. As Ugalde [25] suggests, engaging with an expert in the relevant field or another individual possessing knowledge about your topic can provide access to distinct information that may not be readily accessible elsewhere. The aim was to gather nuanced insights from seasoned educators about the depth of students' reliance on AI tools in the domains of reading, writing, and arithmetic. Embracing an unstructured approach, the interviews allowed for an organic exploration of their observations and experiences.

Two professors, two instructors from NEUST, and two Grade 12 teachers were selected purposively. Professionals with expertise in educational technology, curriculum development, and AI in the classroom were carefully chosen. A diverse group provided varied perspectives across different educational contexts. Ethical considerations were paramount, with explicit permissions obtained for voluntary participation.

Unstructured interviews were conducted, fostering a natural and free-flowing conversation. The unstructured interview questions were reviewed and validated by two English professors. They possess expertise in qualitative research and the formulation of questionnaires. Discussions centered on the professionals' observations of their students' practical experiences with AI, the challenges encountered, the perceived advantages, and insights into how deeply students rely on AI in studying. The absence of a predetermined set of questions allowed for a more comprehensive exploration of their perspectives.

Engaging with education professionals provided valuable insights beyond the student viewpoint, offering a broader understanding of the implications of AI integration in educational practices. The unstructured nature of the interviews allowed for a rich exploration of professionals' observations and experiences regarding the depth of students' reliance on AI. The data collected significantly contributed to constructing the research instrument, enriching the study with diverse perspectives and fostering a comprehensive analysis of the intricate relationship between AI and the foundational aspects of reading, writing, and arithmetic education.

D. Face Validation Process

Subsequent to the insightful interviews with students and education professionals, the research progressed to a pivotal phase: face validation, conducted by the researcher. Face validity refers to the extent to which a test seems to evaluate the specific construct it is intended to measure directly [26]. In this step, the aim was to personally assess and confirm the relevance and clarity of the constructed research instrument. The process involved meticulous scrutiny of the compiled questions to ensure they aligned effectively with the research objectives.

The researcher, equipped with expertise in both the subject matter and research methodology, conducted a thorough examination of the instrument. Attention was given to the clarity, relevance, and appropriateness of each question in relation to the study's goals. Any potential ambiguities or areas for improvement were identified and addressed through refinements to enhance the overall quality of the instrument.

As the face validation was carried out by the researcher, this personalized assessment added a layer of methodological rigor to the validation process. The careful consideration of the instrument's appropriateness and clarity further strengthened its reliability and suitability for gathering pertinent data on AI dependency in the domains of reading, writing, and arithmetic from both students and education professionals. This meticulous validation process lays a robust foundation for the subsequent phases of the research study.

E. Content Validation

In the content validation phase, Aiken's V technique was applied, incorporating the insights of 15 education professionals, including instructors and professors. Each item in the research instrument underwent scrutiny, with the panel assigning scores based on relevance to the study's objectives. Aiken's V formula, $V = s / [n (c - 1)]$, was then employed, where s represents the difference between the given score (r) and the lowest possible score (lo), c is the maximum possible score, and n is the number of raters. A minimum Aiken's V value of 0.73 was set for validation, indicating substantial agreement among the raters for the researcher to deem the items valid [27].

This meticulous quantitative approach, guided by Aiken's V, fortified the content validation process. The calculated Aiken's V values, exceeding the established threshold, affirmed the validity of the research instrument. The collaborative expertise of the education professionals, coupled with statistical rigor, enhanced the credibility of the instrument, establishing it as a reliable tool for probing AI dependency in the specific domains of reading, writing, and arithmetic among students and education professionals.

F. Sample Size

The survey was administered to 1635 students at Nueva Ecija University of Science and Technology, Gabaldon campus, on December 5, 2023. Subsequently, 727 responses were gathered after a period of three months. The survey instrument utilized for data collection was distributed among the students through Google Forms. According to Mundfrom *et al*. [28], numerous recommendations exist for determining the appropriate sample size when conducting a factor analysis, with suggested minimums ranging from three to twenty times the number of variables and absolute ranges spanning from 100 to well beyond 1,000. Moreover, White [29] proposed that the optimal sample size for a study should be contingent upon the demographic characteristics of the participants. Specifically, for student populations, a sample size ranging from 500 to 600 individuals is deemed suitable. Hence, the inclusion of a sample size of 727 participants in this study is deemed sufficient for conducting factorial analysis.

G. Data Analysis

Exploratory Factor Analysis (EFA) was employed to analyze the data, aiming to reveal patterns and structures within it and thereby obtain a deeper understanding of the fundamental constructs at play. Exploratory Factor Analysis (EFA) represents a conventional and formal measurement model employed when it is presumed that both observed and latent variables are measured at the interval level [30]. The Kaiser-Meyer-Olkin (KMO) values falling within the range of 0.6 to 0.69 are considered mediocre, those within 0.7 to 0.79 are deemed middling, and values from 0.8 to 1.0 indicate adequate sampling [31]. Additionally, Kaiser [32] proposed that KMO values exceeding .9 are excellent, those in the .80s are meritorious, those in the .70s are middling, those in the .60s are mediocre, in the .50s are deemed miserable, and values below .5 are considered unacceptable. Bartlett's Test of Sphericity must yield a significant result, as this test assesses the probability of significant correlations within the correlation matrix, a prerequisite for effective factor analysis [33]. The p-value of .05 indicates that the matrix of relationships among the items significantly differs from the unit matrix with no relations [34]. Varimax rotation was employed in this research. Factors possessing eigenvalues of 1.00 or above should be retained [35]. Moreover, Field [36] advises that a minimum loading factor of 0.3 is appropriate, and items with loadings below this threshold should be excluded. Additionally, Guadagnoli and Velicer [37] contend that values exceeding 0.4 are considered stable. Items exhibiting cross-loading factors should be eliminated, followed by a repetition of the analysis, as proposed by Güvendir and Özkan [38].

After conducting an Exploratory Factor Analysis, the Rasch Analysis was used to treat the data. The goal is to refine and validate the measurement properties of a scale, ensuring that the items are functioning well in measuring the intended construct. Rasch analysis was used to examine the internal construct validity of the questionnaires [39]. Moreover, the following criteria were used: Person reliability >0.80, Item reliability >0.90, Person Separation >2, and Item Separation >3 (Linacre, 2023); standard error Infit and Outfit Mean Square (MNSQ) is between 0.60 and 1.4 [40].

It is important to note that the reliability coefficients of Cronbach's alpha that are 0.5 or below are considered unacceptable. Coefficients between 0.50 to 0.59 are classified as poor, 0.60 to 0.69 as questionable, 0.70 to 0.79 as acceptable, 0.80 to 0.89 as good, and 0.90 to 1.00 as excellent [41].

III. RESULTS

A. Constructed Items from Interviews and Face Validity

Table 1 presents the outcomes derived from interviews conducted with both students and experts, along with the assessment of face validity. Within the domain of reading, 15 items were formulated, while another 15 items were crafted for writing, and an additional 15 for numeracy/arithmetic.

The validation of all 45 items was established through a meticulous review by the authors, who scrutinized the interview data to affirm their validity.

Table 1. Constructed items from interviews and face validity

	Dependence on artificial intelligence in the domain of reading	Code
1	I use AI-generated summaries to understand lengthy articles or documents.	RAI1
2	I rely on AI-powered tools for language translation when reading content in a foreign language.	RAI2
3	I depend on AI-driven recommendations to discover new books, articles, or other reading materials.	RAI3
4	I use AI-driven audiobook or text-to-speech applications to consume written content	RAI4
5	AI-driven news aggregators help me stay updated on current events.	RAI5
6	I rely on AI-enhanced tools for extracting key information from research papers or academic articles.	RAI6
7	AI-driven language comprehension tools assist me in understanding complex texts.	RAI7
8	I use AI-powered educational platforms for personalized learning experiences.	RAI8
9	AI-driven accessibility tools help me consume written content more comfortably (e.g., screen readers).	RAI9
10	AI-generated transcripts or subtitles assist me when watching videos or listening to podcasts.	RAI10
11	I use AI-driven language learning apps to improve my reading skills in different languages.	RAI11
12	AI-driven search engine suggestions help me find relevant reading materials more efficiently.	RAI12
13	AI-powered content curation tools assist me in organizing and managing my reading list.	RAI13
14	I depend on AI-generated book reviews or ratings when deciding what to read.	RAI14
15	I use AI-based tools for extracting information from online forums or discussion boards.	RAI15
	Dependence on artificial intelligence in the domain of writing	**Code**
1	I utilize AI-based spelling and grammar checkers when composing written content.	WAI
2	I rely on AI tools for generating ideas or suggestions when drafting written documents.	WAI1
3	I use AI-driven writing assistants to enhance the overall quality of my written work.	WAI2
4	AI-generated templates or frameworks assist me in structuring my written content.	WAI3
5	I depend on AI-powered paraphrasing tools to rephrase sentences in my written work.	WAI4
6	I use AI-driven auto-complete features when composing emails or other written correspondence.	WAI5
7	AI-driven content creation tools help me generate creative or promotional written materials.	WAI6
8	I rely on AI-enhanced proofreading tools to identify errors in my written work.	WAI7
9	AI-based sentiment analysis tools assist me in gauging the tone of my written communication.	WAI8
10	I use AI-powered social media management tools to generate or schedule written posts.	WAI9
11	AI-generated suggestions for email subject lines improve the effectiveness of my written communication.	WAI10
12	I depend on AI-based content summarization tools for condensing lengthy written materials.	WAI11
13	I use AI-driven brainstorming tools to generate ideas for written projects.	WAI12
14	AI-powered writing prompts or exercises help me overcome writer's block.	WAI13
15	I rely on AI-generated suggestions for enhancing the clarity and coherence of my written content.	WAI14
	Dependence on artificial intelligence in the domain of numeracy/arithmetic	**Code**
1	I employ AI-powered calculators or apps for basic arithmetic calculations.	NAI1
2	I rely on AI algorithms to solve complex mathematical problems beyond basic arithmetic.	NAI2
3	I use AI-based tools to check my calculations for accuracy.	NAI3
4	I depend on AI-driven educational platforms for learning and practicing mathematical concepts.	NAI4
5	I integrate AI-powered features, such as smart suggestions, when performing estimation.	NAI5
6	I utilize AI algorithms to analyze and interpret data in mathematical contexts.	NAI6
7	I trust AI-generated solutions for problem-solving.	NAI7
8	I incorporate AI-based tutoring systems to enhance my spatial numeracy skills (geometry and spatial reasoning).	NAI8
9	I rely on AI-assisted methods for generating mathematical models and equations.	NAI9
10	I use AI-driven platforms to collaborate on problem-solving tasks.	NAI10
11	I depend on AI algorithms to automate repetitive tasks involving various numeracy skills in my work or studies.	NAI11
12	I am open to adopting new AI tools and technologies for improving my numeracy skills.	NAI12
13	I prefer AI-generated explanations and step-by-step solutions for learning new mathematical concepts.	NAI13
14	I trust AI systems to assist in decision-making processes involving numeracy tasks.	NAI14
15	I regularly use AI-enhanced platforms to generate and explore mathematical patterns and sequences.	NAI15

B. Content Validity Using Aiken's V

Table 2 shows that the content validity of the items in each of the three constructs is demonstrated. The content validity ratio of the items in reading, writing, and numeracy/arithmetic is beyond the critical value of 0.73; thus, the items are deemed valid.

Table 2. Content validity ratio and content validity index

	Dependence on artificial intelligence in the domain of reading	Aiken's V
1	I use AI-generated summaries to understand lengthy articles or documents.	0.97
2	I rely on AI-powered tools for language translation when reading content in a foreign language.	0.95
3	I depend on AI-driven recommendations to discover new books, articles, or other reading materials.	0.95
4	I use AI-driven audiobook or text-to-speech applications to consume written content	0.95
5	AI-driven news aggregators help me stay updated on current events.	0.95
6	I rely on AI-enhanced tools for extracting key information from research papers or academic articles.	0.95
7	AI-driven language comprehension tools assist me in understanding complex texts.	0.97
8	I use AI-powered educational platforms for personalized learning experiences.	0.97
9	AI-driven accessibility tools help me consume written content more comfortably (e.g., screen readers).	0.97
10	AI-generated transcripts or subtitles assist me when watching videos or listening to podcasts.	0.97
11	I use AI-driven language learning apps to improve my reading skills in different languages.	0.93
12	AI-driven search engine suggestions help me find relevant reading materials more efficiently.	0.91
13	AI-powered content curation tools assist me in organizing and managing my reading list.	0.91
14	I depend on AI-generated book reviews or ratings when deciding what to read.	0.93
15	I use AI-based tools for extracting information from online forums or discussion boards.	0.95
	Dependence on artificial intelligence in the domain of writing	**Aiken's V**
1	I utilize AI-based spelling and grammar checkers when composing written content.	0.95
2	I rely on AI tools for generating ideas or suggestions when drafting written documents.	0.95

3	I use AI-driven writing assistants to enhance the overall quality of my written work.	0.95
4	AI-generated templates or frameworks assist me in structuring my written content.	0.93
5	I depend on AI-powered paraphrasing tools to rephrase sentences in my written work.	0.93
6	I use AI-driven auto-complete features when composing emails or other written correspondence.	0.93
7	AI-driven content creation tools help me generate creative or promotional written materials.	0.95
8	I rely on AI-enhanced proofreading tools to identify errors in my written work.	0.95
9	AI-based sentiment analysis tools assist me in gauging the tone of my written communication.	0.95
10	I use AI-powered social media management tools to generate or schedule written posts.	0.95
11	AI-generated suggestions for email subject lines improve the effectiveness of my written communication.	0.97
12	I depend on AI-based content summarization tools for condensing lengthy written materials.	0.95
13	I use AI-driven brainstorming tools to generate ideas for written projects.	0.93
14	AI-powered writing prompts or exercises help me overcome writer's block.	0.88
15	I rely on AI-generated suggestions for enhancing the clarity and coherence of my written content.	0.88
	Dependence on artificial intelligence in the domain of numeracy/arithmetic	**Aiken's V**
1	I employ AI-powered calculators or apps for basic arithmetic calculations.	0.93
2	I rely on AI algorithms to solve complex mathematical problems beyond basic arithmetic.	
3	I use AI-based tools to check my calculations for accuracy.	0.97
4	I depend on AI-driven educational platforms for learning and practicing mathematical concepts.	0.97
5	I integrate AI-powered features, such as smart suggestions, when performing estimation.	0.95
6	I utilize AI algorithms to analyze and interpret data in mathematical contexts.	0.93
7	I trust AI-generated solutions for problem-solving.	0.93
8	I incorporate AI-based tutoring systems to enhance my spatial numeracy skills (geometry and spatial reasoning).	0.95
9	I rely on AI-assisted methods for generating mathematical models and equations.	0.97
10	I use AI-driven platforms to collaborate on problem-solving tasks.	0.97
11	I depend on AI algorithms to automate repetitive tasks involving various numeracy skills in my work or studies.	0.95
12	I am open to adopting new AI tools and technologies for improving my numeracy skills.	
13	I prefer AI-generated explanations and step-by-step solutions for learning new mathematical concepts.	0.93
14	I trust AI systems to assist in decision-making processes involving numeracy tasks.	0.95
15	I regularly use AI-enhanced platforms to generate and explore mathematical patterns and sequences.	0.97

C. *Exploratory Factor Analysis (EFA)*

Table 3 presents the results of the Exploratory Factor Analysis. The KMO measure assesses the adequacy of the data for conducting a factor analysis. It ranges from 0 to 1, where higher values indicate better suitability for factor analysis [32]. A KMO value of 0.896 is generally considered excellent, suggesting that the variables in the dataset are well-suited for factor analysis. In addition, Bartlett's test checks the null hypothesis that the correlation matrix is an identity matrix (no correlation between variables) [33]. A small p-value (in this case, $p < 0.001$) indicates that the correlation matrix is significantly different from an identity matrix. This suggests that there are significant correlations between variables, making the data appropriate for factor analysis. The EFA was utilized to validate the construct identified through qualitative analysis as well as to remove items that did not add to the target construct's testing power.

Table 3. KMO and Bartlett's test of Sphericity

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.896
Bartlett's Test of Sphericity	Approx. Chi-Square	30242.768
	df	630
	Sig.	0.000

D. *Scree Plot*

A Scree plot represents a line graph illustrating the eigenvalues of factors [42]. In the scree plot shown in Fig. 1, the elbow point appears to occur around the third factor. Beyond this point, the eigenvalues level off, suggesting that additional factors contribute less to the overall variance. Therefore, it seems reasonable to consider retaining the first three factors as they capture the most significant amount of variance in the data.

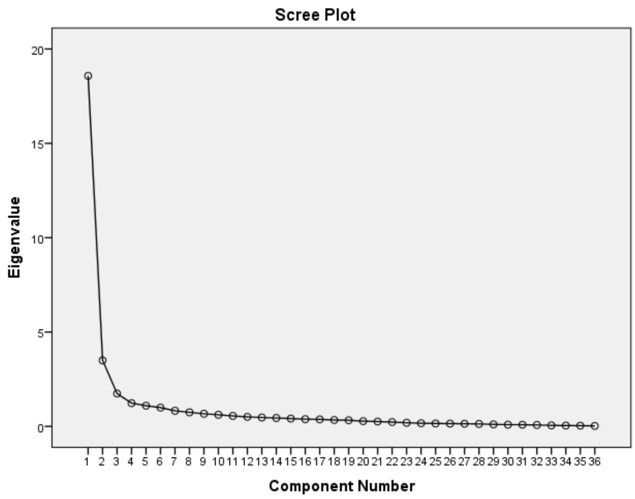

Fig. 1. Scree plot result of EFA.

E. *Loading Factors*

Table 4 illustrates that all loading factors associated with the retained items surpass the 0.5 threshold. Following the recommendation by Hair *et al*. [43], who propose a cutoff for factor loadings relative to sample size, a factor loading of 0.35 is deemed significant for a sample size of 250. Consequently, the loading factors indicate a robust positive relationship between the observed variables and their respective underlying factors. Conversely, certain items failed to meet the loading factor criterion of 0.5 and exhibited cross-loadings. Consequently, these items were excluded from the analysis. Specifically, the items WAI: 6, 10, 13, 16, and RAI: 2, 4, 12, 13, and 15 were omitted due to their failure to meet the specified loading factor threshold

Table 4. Rotated component matrix

	Component		
	1	2	3
NAI2	0.849		
NAI14	0.845		
NAI3	0.828		
NAI7	0.825		
NAI15	0.806		
NAI6	0.796		

NAI11	0.794		
NAI9	0.786		
NAI1	0.772		
NAI13	0.749		
NAI10	0.747		
NAI5	0.724		
NAI8	0.706		
NAI12	0.687		
NAI4	0.669		
WAI3		0.829	
WAI15		0.781	
WAI1		0.760	
WAI4		0.729	
WAI2		0.707	
WAI8		0.697	
WAI9		0.689	
WAI5		0.681	
WAI11		0.647	
WAI7		0.631	
WAI12		0.568	
RA13			0.784
RAI5			0.726
RAI9			0.666
RAI11			0.608
RAI6			0.603
RAI8			0.603
RAI10			0.583
RAI7			0.582
RAI14			0.559
RAI1			0.553

F. Rasch Analysis

Table 5 presents the results of the Rasch Analysis. Person Separation Reliability (PSR) serves as a metric assessing the instrument's efficacy in differentiating individuals with varying levels of the latent trait under consideration [44]. The person reliability scores for the constructs of reading, writing, and numeracy/arithmetic are .901, .915, and .948, respectively, surpassing the established threshold of .80. Similarly, the person separation values for reading, writing, and numeracy/arithmetic are 2.12, 2.53, and 2.45, respectively, exceeding the threshold of 2. These reliability and separation values collectively indicate the instrument's sensitivity in distinguishing between individuals with high and low-performance levels [45]. Likewise, the values for item reliability in the domains of reading, writing, and numeracy/arithmetic are .98, .99, and 1.00, respectively, all surpassing the threshold of .90. Additionally, the item separation values for reading, writing, and numeracy/arithmetic are 5.41, 3.92, and 4.32, respectively, exceeding the threshold of 3. These item reliability and separation metrics signify that the sample size is substantial enough to validate the item difficulty hierarchy and ensure the construct validity of the instrument concerning reading, writing, and numeracy/arithmetic [46].

Table 5. Person and item separation reliability

Construct	Person Reliability >0.80	Person Separation >2	Item Reliability >0.9	Item Separation >3
Reading	0.901	2.12	0.98	5.41
Writing	0.915	2.53	0.99	3.92
Numeracy/ Arithmetic	0.948	2.45	1.00	4.32

G. Wright Map

The Wright Map was utilized to ascertain the spectrum of candidates' measured abilities, ranging from highest to lowest, as well as to assess the difficulty levels of the items [47]. Figs. 2–4 show that the items range from moderately easy to moderately difficult, clustering close to zero for effective measurement. Additionally, the wide dispersion of respondents' latent traits from the average is acceptable for the rating scale.

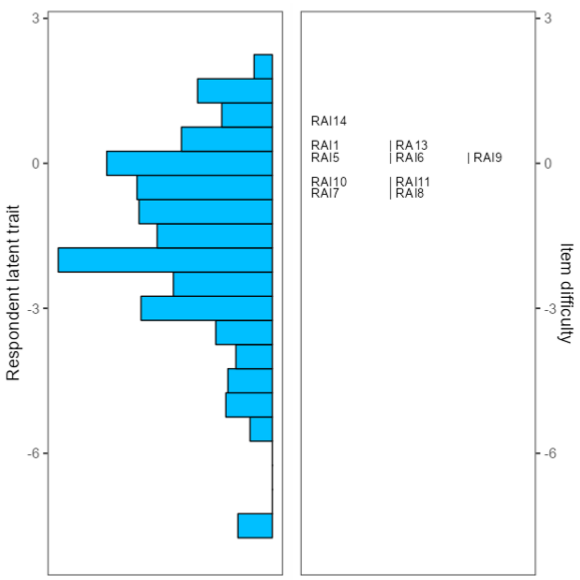

Fig. 2. Wright map for reading.

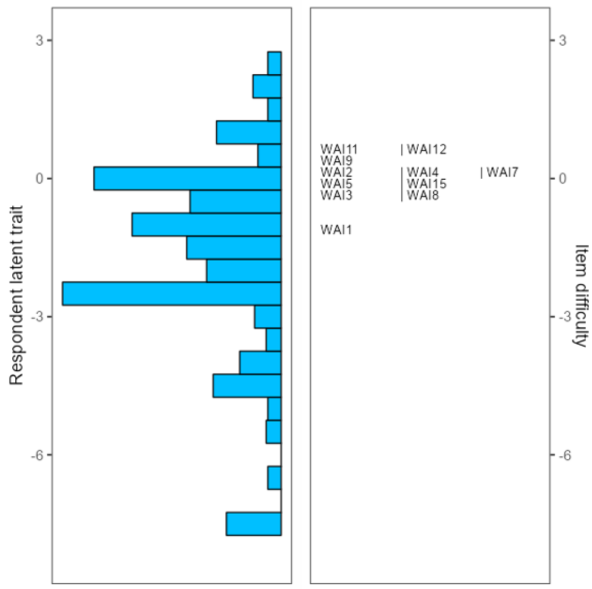

Fig. 3. Wright map for writing.

H. Fit Analysis and Difficulty Indices

As seen in Table 6, the measures of item difficulty range from -3 logits to +3 logits [48]. The difficulty index ranges from -0.9899 to 0.8548, indicating that the items vary from moderately easy to moderately difficult. The standard error measures, indicated by S.E. Measures, reflect the precision of the Rasch estimate [46]. According to Tham *et al.* [49], reliable measurement is generally expected when estimating individual ability and item difficulty in the context of Rasch analysis, with the standardized error for each item ideally being ≤0.30 logits. In this measurement, all item standard errors (S.E.) are below 0.30, indicating a high level of precision in the Rasch estimates. Additionally, the majority of the items' Infit and Outfit MNSQ fall within the acceptable range of 0.60 and 1.4 [40]. These items fit well with the predicted Rasch model. However, three items (NAI9, NAI10, and NAI11) did not meet the Infit and Outfit MNSQ criteria; consequently, these items were excluded.

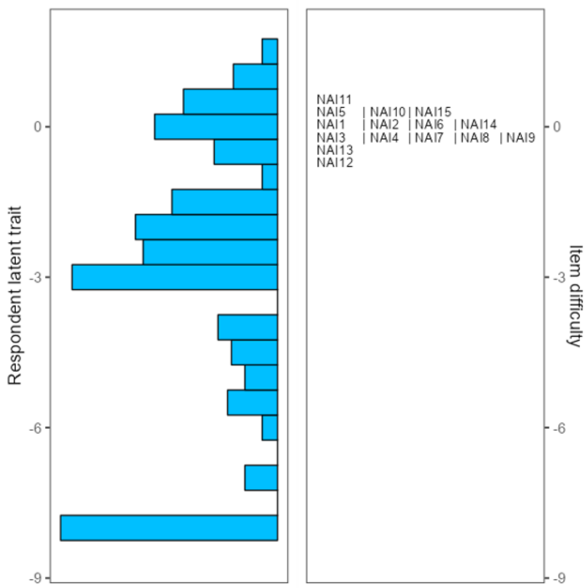

Fig. 4. Wright map for numeracy/arithmetic.

Table 6. Fit statistics of measurement items

	Item mean	Item Difficulty Measure	S.E. Measure	Infit	Outfit
RAI1	2.3	0.3795	0.063	1.162	1.181
RAI5	2.41	0.0686	0.0625	0.952	0.94
RAI6	2.42	0.0413	0.0625	0.933	0.965
RAI7	2.62	−0.526	0.0619	0.924	0.919
RAI8	2.66	−0.6254	0.0618	0.984	0.99
RAI9	2.39	0.1234	0.0626	0.84	0.854
RAI10	2.57	−0.3957	0.062	1.216	1.201
RAI11	2.58	−0.3996	0.062	0.78	0.79
RA13	2.27	0.479	0.0632	0.609	0.593
RAI14	2.14	0.8548	0.064	0.937	0.917
WAI1	2.7	−0.9899	0.0636	0.8	0.801
WAI2	2.34	0.0971	0.0651	0.69	0.719
WAI3	2.51	−0.4133	0.0643	0.623	0.62
WAI4	2.32	0.1564	0.0651	0.744	0.742
WAI5	2.42	−0.1553	0.0647	1.019	1.008
WAI7	2.31	0.1819	0.0652	0.984	1.004
WAI8	2.5	-0.3967	0.0643	1.049	1.036
WAI9	2.24	0.4041	0.0656	0.779	0.768
WAI11	2.18	0.5942	0.0659	0.835	0.827
WAI12	2.16	0.6726	0.0661	1.078	1.287
WAI15	2.42	−0.1512	0.0647	0.695	0.681
NAI1	2.03	0.0822	0.0721	0.99	0.953
NAI2	2.02	0.1239	0.0722	0.755	0.754
NAI3	2.08	−0.1192	0.0716	0.748	0.73
NAI4	2.08	−0.1192	0.0716	1.097	1.352
NAI5	1.98	0.2497	0.0726	0.885	0.837
NAI6	2.02	0.1134	0.0722	0.615	0.59
NAI7	2.09	−0.1601	0.0715	0.613	0.626
NAI8	2.1	−0.1908	0.0714	0.883	0.851
NAI9	2.09	−0.1653	0.0715	***0.486***	***0.465***
NAI10	1.99	0.2286	0.0725	***0.58***	***0.561***
NAI11	1.88	0.6499	0.0735	***0.493***	***0.468***
NAI12	2.22	−0.612	0.0702	1.035	1.103
NAI13	2.17	−0.4231	0.0707	1.065	1.101
NAI14	2.02	0.1291	0.0722	0.616	0.618
NAI15	1.99	0.2129	0.0725	0.613	0.608

I. Reliability Testing Using Cronbach's Alpha

Table 7 displays the Cronbach alpha value for reading, writing, and numeracy/arithmetic. The table indicates that reading (0.89) has good reliability, whereas writing (0.91) and numeracy/arithmetic (0.90) both have an excellent reliability coefficient. This data suggests a strong level of consistency within each factor, indicating that the items in each group are adequately correlated [50].

Table 7. Cronbach's alpha coefficients

	Cronbach's Alpha	Interpretation
Reading	.89	Good
Writing	.91	Excellent
Numeracy/arithmetic	.90	Excellent

IV. Discussion

The 45-item assessment framework presented in Table 1, targeting reading, writing, and numeracy/arithmetic, leverages data triangulation by incorporating student and expert interviews. This initial validation via meticulous author review establishes face validity, a positive first impression of alignment with the intended measurement. After establishing face validity, the instrument was administered to 15 raters to assess its content validity. According to Almenara and Cejudo [51], utilizing expert judgment as an evaluation approach offers advantages such as obtaining high-quality responses from judges and acquiring extensive information on the subject matter. All items in the reading, writing, and numeracy/arithmetic domains exceeded the critical Aiken value of 0.73, indicating strong content validity. As described by Telenius *et al.* [52], Aiken values below 0.4 signify low validity, while values between 0.4 and 0.8 represent moderate validity, and values above 0.8 indicate high validity. Therefore, the instrument demonstrates high content validity across all assessed domains.

Following the implementation of Aiken's V for content validity assessment, Exploratory Factor Analysis was undertaken. This analytical process aims to ascertain the existence of correlations among the items within each factor represented by the questionnaire [53]. EFA also serves to eliminate redundant items, resulting in a more concise and efficient questionnaire while enhancing its reliability, as a reduced number of items minimizes the likelihood of measurement error. The data demonstrated excellent suitability for conducting factor analysis to explore student dependency in the 3Rs. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy yielded a value of 0.896, significantly exceeding the recommended threshold of 0.85. This indicates strong underlying relationships between the measured variables, suggesting their effectiveness in capturing student dependency dimensions. Shrestha [31] asserted that a Kaiser-Meyer-Olkin (KMO) value ranging between 0.80 and 1.00 is considered satisfactory. Furthermore, Bartlett's test of sphericity resulted in a highly significant p-value of 0.001, definitively rejecting the null hypothesis of no correlation between variables [54]. This confirms the presence of substantial inter-variable relationships, further solidifying the appropriateness of the data for factor analysis. These positive outcomes provide strong evidence for the validity of proceeding with factor analysis to uncover the latent factors associated with student dependency in the 3Rs, potentially leading to valuable insights for educational practices.

Table 4 reveals promising findings concerning the robustness of associations between observed variables and the latent factors extracted through factor analysis. All retained items demonstrate loading factors surpassing the recommended threshold of 0.5. Hair *et al*. [43] suggest that for a sample size of 250, a factor loading of 0.35 is sufficient for significance. Furthermore, Tabachnick and Fidell [55] propose the exclusion of factor loadings with an absolute value of less than 0.32. In this instance, the observed loading factors ranging from 0.553 to 0.849 signify a notably robust and positive association between the measured variables and their respective underlying factors, thereby enhancing the instrument's construct validity.

However, it's important to acknowledge that certain items were excluded due to their failure to meet the 0.5 loading factor criterion and exhibiting cross-loadings, potentially indicating ambiguity or redundancy. Items WAI: 6, 10, 13, 16, and RAI: 2, 4, 12, 13, 15 were subsequently omitted to ensure the instrument's overall clarity and interpretability. Following the removal of items exhibiting cross-loadings, the Exploratory Factor Analysis (EFA) was subsequently re-executed, as outlined by Samuels [56]. This refinement strengthens the internal consistency and construct validity of the instrument, ultimately improving its ability to accurately measure student dependency in the 3Rs. Overall, the analysis presented in Table 4 reveals promising results, highlighting the instrument's potential to effectively capture the intended constructs. The strong loading factors coupled with the exclusion of problematic items contribute to a more robust and reliable measure of student dependency, paving the way for further exploration and insights into this crucial educational domain.

The analysis of person and item reliability provides compelling evidence for the instrument's strong capacity to accurately measure student dependency in the 3Rs. Person Reliability for reading (.901), writing (.915), and numeracy/arithmetic (.948) all comfortably exceed the established threshold of .80, indicating the instrument's effectiveness in differentiating individuals with varying levels of dependency within each domain. According to Tesio *et al*. [57], a persons reliability exceeding 0.7 is deemed satisfactory when assessing groups of individuals. Furthermore, the corresponding person separation values (2.12, 2.53, 2.45) surpass the recommended threshold of 2, solidifying the instrument's sensitivity in distinguishing between high and low performers. These results align with the findings of Hrnjičić & Alihodžić [44], highlighting the instrument's ability to reliably capture individual differences in student dependencies for each of the 3Rs. Additionally, the person reliability and person separation metrics suggest that a sufficient number of items are available for each factor [45].

On the item level, the analysis unveils equally positive outcomes. Item reliability values in reading (0.98), writing (0.99), and numeracy/arithmetic (1.00) significantly surpass the threshold of 0.90, demonstrating exceptional internal consistency for each item group. This observation aligns with the research conducted by Linacre [45], underscoring the instrument's capacity to consistently and reliably gauge the intended constructs. Moreover, the item separation values (5.41, 3.92, 4.32) all exceed the suggested threshold of 3, indicating a large enough sample size to validate the item difficulty hierarchy and further supporting the instrument's construct validity across all three domains.

Both person and item reliability analyses provide robust evidence for the instrument's strong psychometric properties and its ability to accurately assess student dependency in reading, writing, and numeracy/arithmetic. These positive findings pave the way for further confidence in the instrument's utility for research and educational practice, allowing researchers and educators to gain valuable insights into student dependencies and implement effective interventions to address them.

The analysis of item difficulty and fit statistics reveals encouraging results with some minor exclusions. The item difficulty measures span a healthy range of -3 to +3 logits, as suggested by Boone [48], capturing a diverse spectrum of dependency levels. Additionally, the standard error measures remain low, signifying precise estimations of item difficulty [46]. Notably, the majority of items exhibit Infit and Outfit MNSQ values within the acceptable range of 0.60-1.4 [40], confirming their overall good fit to the Rasch model.

However, it's worth noting the exclusion of three items (NAI9, NAI10, and NAI11) due to their exceeding the Infit and Outfit MNSQ criteria. This indicates a potential misfit with the Rasch model, suggesting these items might not function consistently in measuring the intended construct. Removing these items ensures the overall coherence and accuracy of the instrument.

The positive item difficulty distribution, low standard errors, and good fit for most items indicate the instrument's strong ability to reliably assess student dependency across its range of difficulty levels. The exclusion of a few problematic items further strengthens the overall reliability and validity of the instrument, leading to more accurate measurements and insightful interpretations.

According to the provided data, the Cronbach alpha value for reading is 0.89, indicating a commendable level of reliability. This suggests that the items assessing AI dependency within the realm of reading consistently measure the same construct. Moreover, both writing and numeracy/arithmetic exhibit even higher Cronbach alpha values, standing at 0.91 and 0.90, respectively, signifying excellent reliability. These results imply not only internal consistency within the items measuring AI dependency in writing and numeracy/arithmetic but also a consistent measurement of the intended constructs. The elevated reliability coefficients denote a robust level of internal consistency within each factor. This finding is noteworthy, emphasizing that the items within each category are sufficiently correlated [58]. This correlation indicates that individual items within each factor effectively measure the same underlying construct and operate cohesively as a unified set.

The findings indicate that the questionnaire effectively measures students' reliance on AI and demonstrates validity and reliability. AI appears to positively impact the educational process [59] and enhances academic enthusiasm [60, 61]. However, the use of AI tools like ChatGPT may inadvertently lead to plagiarism, presenting challenges to academic integrity [62]. Therefore, this newly developed instrument can be instrumental in identifying and addressing potential misuse of AI in student learning.

V. Conclusion and Recommendations

In conclusion, the comprehensive assessment framework developed in this study demonstrates strong psychometric properties, establishing itself as a reliable and valid tool for measuring student dependency in reading, writing, and numeracy/arithmetic. The meticulous approach, incorporating data triangulation, face validity, and content validity assessment through Aiken's V, ensures that the instrument effectively captures the targeted constructs. Exploratory Factor Analysis further validated the robustness of the associations between observed variables and latent factors, while refinement processes and reliability analyses confirmed the instrument's internal consistency and ability to differentiate between varying levels of student dependency.

Despite the exclusion of a few misfitting items, the overall outcomes affirm the instrument's utility in both research and educational practice. Its strong psychometric properties make it a valuable resource for understanding student dependencies and implementing targeted interventions. Researchers and educators can confidently use this framework to enhance educational practices and improve student outcomes in the fundamental areas of reading, writing, and numeracy/arithmetic.

To maximize the benefits of this framework, it is recommended that students actively engage in assessments and seek support when dependencies are identified. Teachers should use the diagnostic insights to tailor their teaching strategies, continuously monitoring and adapting instruction to meet evolving student needs. Curriculum developers are encouraged to regularly review and align curricula with the identified dependencies, ensuring a comprehensive and flexible educational experience. School leaders should support these efforts by providing professional development opportunities and allocating resources for targeted interventions.

Future researchers are advised to conduct validation studies across diverse populations and educational settings, and to explore the long-term impact of interventions informed by the assessment framework. By implementing these recommendations, all stakeholders can contribute to a more effective and personalized educational environment, fostering student growth and success in the essential domains of reading, writing, and numeracy/arithmetic.

The respondents in this study were drawn from a single school, which may constrain the generalizability of the findings. In addition, the author has no control over the respondents' perceptions of how or why they integrate AI technology into their learning process.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

The author would like to thank the Campus Director of NEUST-Gabaldon campus, Dr. Franklin T. Dumayas, for allowing the study to take place on the campus. Special thanks also go to the faculty raters who evaluated the items in the questionnaire. Additionally, special thanks are extended to the students who served as respondents for the study.

References

[1] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," *IEEE Access*, vol. 8, pp. 75264–75278, Apr. 2020. doi: https://doi.org/10.1109/ACCESS.2020.2988510

[2] A. Almusaed, A. Almssad, I. Yitmen, and R. Z. Homod, "Enhancing student engagement: Harnessing 'AIED's power in hybrid education—a review analysis," *Education Sciences*, vol. 13, no. 7, p. 632, Jul. 2023. doi: https://doi.org/10.3390/educsci13070632

[3] S. RIAZ, "Should students be using online tools and AI to study?" *eLearning Industry*, Sep. 22, 2023.

[4] T. S. Millacci, "How to promote cognitive development: 23 activities & games," PositivePsychology.com, Jan. 11, 2022.

[5] V. Srinivasan and H. Murthy, "Improving reading and comprehension in K-12: Evidence from a large-scale AI technology intervention in India," *Computers and Education: Artificial Intelligence*, vol. 2, 100019, 2021. doi: https://doi.org/10.1016/j.caeai.2021.100019

[6] J.-C. Hsiao and J. S. Chang, "Enhancing EFL reading and writing through AI-powered tools: design, implementation, and evaluation of an online course," *Interactive Learning Environments*, pp. 1–16, May 2023. doi: https://doi.org/10.1080/10494820.2023.2207187

[7] S. F. Ahmad *et al*., "Impact of artificial intelligence on human loss in decision making, laziness and safety in education," *Humanities and Social Sciences Communications*, vol. 10, no. 1, pp. 1–14, Jun. 2023. doi: https://doi.org/10.1057/s41599-023-01787-8

[8] Z. Buçinca, M. B. Malaya, and K. Z. Gajos, "To trust or to think," in *Proc. the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–21, Apr. 2021. doi: https://doi.org/10.1145/3449287

[9] A. R. Malik *et al*., "Exploring artificial intelligence in academic essay: Higher education student's perspective," *International Journal of Educational Research Open*, vol. 5, 100296, pp. 100296–100296, Dec. 2023. doi: https://doi.org/10.1016/j.ijedro.2023.100296

[10] (Jul. 12, 2023). The synergy of human expertise and AI: Maximizing academic writing efficiency and impact. *Editage Insights*. [Online]. Available: https://www.editage.com/insights/the-synergy-of-human-expertise-and-ai-maximizing-academic-writing-efficiency-and-impact?refer=scroll-to-1-article&refer-type=article

[11] J. Huang and M. Tan, "The role of ChatGPT in scientific communication: Writing better scientific review articles," *American Journal of Cancer Research*, vol. 13, no. 4, pp. 1148–1154, 2023.

[12] N. Baron, "Even kids are worried ChatGPT will make them lazy plagiarists, says a linguist who studies tech's effect on reading, writing and thinking," *Fortune*, Jan. 19, 2023.

[13] H. Qawaqneh, F. B. Ahmad, and A. R. Alawamreh, "The impact of artificial intelligence-based virtual laboratories on developing students' motivation towards learning mathematics," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 18, no. 14, pp. 105–121, Jul. 2023. doi: https://doi.org/10.3991/ijet.v18i14.39873

[14] A. Cunska, "Effective learning strategies and Artificial Intelligence (AI) Support for accelerated math acquisition," in *ICEEPSY 2020: Education and Educational Psychology, European Proceedings of International Conference on Education and Educational Psychology*, European Publisher, 2020, pp. 118–129.

[15] N. Rane, "Enhancing mathematical capabilities through ChatGPT and similar generative artificial intelligence: Roles and challenges in solving mathematical problems," *Social Science Research Network*, Aug. 20, 2023.

[16] W. Yousef, M. A. Tashtoush, R. AlAli, and A. M. Jarrah, "ChatGPT: A revolutionary tool for teaching and learning mathematics," *EURASIA Journal of Mathematics, Science and Technology Education*, vol. 19, no. 7, pp. em2286–em2286, Jul. 2023. doi: https://doi.org/10.29333/ejmste/13272

[17] M. Choliz, "Mobile-phone addiction in adolescence: The test of mobile phone dependence (TMD)," *Progress in Health Sciences*, vol. 2, no. 1, pp. 33–34, 2012.

[18] M. Chóliz *et al*., "Development of a brief multicultural version of the test of mobile phone dependence (TMDbrief) questionnaire," *Frontiers in Psychology*, vol. 7, May 2016. doi: https://doi.org/10.3389/fpsyg.2016.00650

[19] V. A. Emelin, E. I. Rasskazova, and A. S. Tkhostov, "Development and approbation the methodic assessing changed of psychological boundaries in technological gadgets use," *Psychological Research*, vol. 2, no. 22, p. 5, 2012.

[20] A. T. Capinding, "Development and validation of instruments for assessing the impact of artificial intelligence on students in higher education," *International Journal of Educational Methodology*, vol. 10, issue 2, pp. 197–211, May 2024. doi: https://doi.org/10.12973/ijem.10.2.997

[21] M. Martínez-Córcoles, M. Teichmann, and M. Murdvee, "Assessing technophobia and technophilia: Development and validation of a questionnaire," *Technology in Society*, vol. 51, pp. 183–188, Nov. 2017. doi: https://doi.org/10.1016/j.techsoc.2017.09.007

[22] C. Muñoz-Neira, O. L. López, R. Riveros, J. Núñez-Huasaf, P. Flores, and A. Slachevsky, "The technology—activities of daily living questionnaire: A version with a technology-related subscale," *Dementia and Geriatric Cognitive Disorders*, vol. 33, no. 6, pp. 361–371, 2012. doi: https://doi.org/10.1159/000338606

[23] K. Kieslich, M. Lünich, and F. Marcinkowski, "The threats of Artificial Intelligence Scale (TAI)," *International Journal of Social Robotics*, Jan. 2021. doi: https://doi.org/10.1007/s12369-020-00734-w

[24] J. Birt, "Probing questions: Definition, comparisons and examples," *Indeed Career Guide*, Mar. 11, 2023.

[25] V. Ugalde, "Using an interview in a research paper | writing & speaking center," University of Nevada, Reno, 2023.

[26] L. Neumann, "Face validity—a simple guide with examples," Bachelor Print, 2022.

[27] L. R. Aiken, "Three coefficients for analyzing the reliability and validity of ratings," *Educational and Psychological Measurement*, vol. 45, no. 1, pp. 131–142, Mar. 1985. doi: https://doi.org/10.1177/0013164485451012

[28] D. J. Mundfrom, D. G. Shaw, and T. L. Ke, "Minimum sample size recommendations for conducting factor analyses," *International Journal of Testing*, vol. 5, no. 2, pp. 159–168, Jun. 2005. doi: https://doi.org/10.1207/s15327574ijt0502_4

[29] M. White, "Sample size in quantitative instrument validation studies: A systematic review of articles published in Scopus, 2021," *Heliyon*, vol. 8, no. 12, p. e12223, Dec. 2022. doi: https://doi.org/10.1016/j.heliyon.2022.e12223

[30] J. R. J. Fontaine, "Equivalence," *Encyclopedia of Social Measurement*, pp. 803–813, 2005. doi: https://doi.org/10.1016/b0-12-369398-5/00116-x

[31] N. Shrestha, "Factor analysis as a tool for survey analysis," *American Journal of Applied Mathematics and Statistics*, vol. 9, no. 1, pp. 4–11, Jan. 2021. doi: https://doi.org/10.12691/ajams-9-1-2

[32] H. F. Kaiser, "An index of factorial simplicity," *Psychometrika*, vol. 39, no. 1, pp. 31–36, Mar. 1974. doi: https://doi.org/10.1007/bf02291575

[33] M. S. Bartlett, "The effect of standardization on a χ 2 approximation in factor analysis," *Biometrika*, vol. 38, no. 3/4, p. 337, Dec. 1951. doi: https://doi.org/10.2307/2332580

[34] A. Can, *Quantitative Data Analysis in the Process of Scientific Research with SPSS*, Ankara: Pegem Akademi, 2016.

[35] J. Brown, *What Is an Eigenvalue*, JALT Testing & Evaluation SIG Newsletter, 2001.

[36] A. Field, *Discovering Statistics Using IBM SPSS*, 4th ed. London: Sage, 2013.

[37] E. Guadagnoli and W. F. Velicer, "Relation of sample size to the stability of component patterns," *Psychological Bulletin*, vol. 103, no. 2, pp. 265–275, Mar. 1988. doi: https://doi.org/10.1037/0033-2909.103.2.265

[38] M. Acar Güvendir and Y. Özer Özkan, "Item removal strategies conducted in exploratory factor analysis: A comparative study," *International Journal of Assessment Tools in Education*, vol. 9, no. 1, pp. 165–180, Feb. 2022. doi: https://doi.org/10.21449/ijate.827950

[39] E. Ningrum, S. Evans, and S.-E. Soh, "Validation of the Indonesian version of the safety attitudes questionnaire: A Rasch analysis," *PLOS ONE*, vol. 14, no. 4, p. e0215128, Apr. 2019. doi: https://doi.org/10.1371/journal.pone.0215128

[40] T. G. Bond and C. M. Fox, *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (2nd ed.), 2nd ed. Routledge, 2007.

[41] N. F. Habidin, A. F. M. Zubir, N. M. Fuzi, N. A. M. Latip, and M. N. A. Azman, "Sustainable performance measures for Malaysian automotive industry," *World Applied Sciences Journal*, vol. 33, no. 6, pp. 1017–1024, 2015. doi: https://doi.org/10.5829/idosi.wasj.2015.33.06.257

[42] R. B. Cattell, "The scree test for the number of factors," *Multivariate Behavioral Research*, vol. 1, no. 2, pp. 245–276, Apr. 1966. doi: https://doi.org/10.1207/s15327906mbr0102_10

[43] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 7th edition. Prentice Hall, 2009.

[44] A. Hrnjičić and A. Alihodžić, "Measuring students' conceptual understanding of real functions: A Rasch model analysis," *International Electronic Journal of Mathematics Education*, vol. 19, no. 1, pp. em0761–em0761, Jan. 2024. doi: https://doi.org/10.29333/iejme/13942

[45] J. M. Linacre. (2023). Reliability and separation of measures. [Online]. Available: https://www.winsteps.com/winman/reliability.htm

[46] J. M. Linacre. (2023). Standard errors: model and real. [Online]. Available: https://www.winsteps.com/winman/standarderrors.htm#:~:text=of%20the%20parameter

[47] M. E. Lunz. Using the very useful wright map. [Online]. Available: https://www.rasch.org/mra/mra-01-10.htm

[48] W. J. Boone, "Rasch analysis for instrument development: Why, when, and how?" CBE—life sciences education, vol. 15, no. 4, p. rm4, Dec. 2016. doi: https://doi.org/10.1187/cbe.16-04-0148

[49] T. Kerstin, B. Bernspång, and G. F. Anne, "Development of the assessment of awareness of disability," *Scandinavian Journal of Occupational Therapy*, vol. 6, no. 4, pp. 184–190, Jan. 1999. doi: https://doi.org/10.1080/110381299443663

[50] J. Bruin. (2006). Stata programs for data analysis. [Online]. Available: https://stats.oarc.ucla.edu/stata/ado/analysis/

[51] J. C. Almenara and M. D. C. L. Cejudo, "La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información y comunicación (TIC)," *Revista Eduweb*, vol. 7, no. 2, 2013.

[52] E. W. Telenius, K. Engedal, and A. Bergland, "Inter-rater reliability of the Berg Balance Scale, 30 s chair stand test and 6 m walking test, and construct validity of the Berg Balance Scale in nursing home residents with mild-to-moderate dementia," *BMJ Open*, vol. 5, no. 9, e008321, Sep. 2015. doi: https://doi.org/10.1136/bmjopen-2015-008321

[53] A. Dabbagh, H. Seens, J. Fraser, and J. C. MacDermid, "Construct validity and internal consistency of the Home and family work roles questionnaires: A cross-sectional study with exploratory factor analysis," *BMC Women's Health*, vol. 23, no. 1, p. 56, 2023.

[54] D. Strengejacke. (June 5, 2024). Check suitability of data for Factor Analysis (FA) with Bartlett's Test of Sphericity and KMO—check_factorstructure. [Online]. Available: http://tinyurl.com/538uvk89

[55] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 6th edition. Harlow: Pearson, 2014.

[56] P. Samuels. (June 9, 2017). Advice on exploratory factor analysis. [Online]. Available: https://www.open-access.bcu.ac.uk/6076/

[57] L. Tesio, A. Caronni, A. Simone, Dinesh Kumbhare, and S. Scarano, "Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment," *Disability and Rehabilitation*, vol. 46, no. 3, pp. 1–14, Feb. 2023. doi: https://doi.org/10.1080/09638288.2023.2169772

[58] E. C. Fein, J. Gilmour, T. Machin, and L. Hendry, *Statistics for Research Students,* University of Southern Queensland, 2022.

[59] W. Yousef, M. Tashtoush, R. Alali, and S. Saleh, "Artificial intelligence in education: mathematics teachers' perspectives, practices and challenges," *Iraqi Journal for Computer Science and Mathematics*, vol. 5, no. 1, pp. 60–77, Jan. 2024. doi: https://doi.org/10.52866/ijcsm.2024.05.01.004

[60] M. Fannakhosrow, S. Nourabadi, D. T. Ngoc Huy, N. Dinh Trung, and M. A. Tashtoush, "A comparative study of Information and Communication Technology (ICT)-based and conventional methods of instruction on learners' academic enthusiasm for L2 learning," *Education Research International*, vol. 2022, pp. 1–8, Mar. 2022. doi: https://doi.org/10.1155/2022/5478088

[61] M. A. Tashtoush *et al*., "The impact of Information and Communication Technologies (ICT)-based education on the mathematics academic enthusiasm," *Journal of Educational and Social Research*, vol. 13, no. 3, pp. 284–284, May 2023. doi: https://doi.org/10.36941/jesr-2023-0077

[62] A. M. Jarrah, Yousef Wardat, and P. Fidalgo, "Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature

say?” *Online Journal of Communication and Media Technologies*, vol. 13, no. 4, pp. e202346–e202346, Oct. 2023. doi: https://doi.org/10.30935/ojcmt/13572