https://doi.org/10.59865/abacj.2023.67

Special Issue-Application of Structural Equation Modeling in Business, Sport, Tourism, Recreation, Education, and Language Studies

DEVELOPMENT AND VALIDATION OF A COMPOSITE LEARNING INDEX FOR CAMBODIAN HIGH SCHOOL STUDENTS

Bunhe Harth[1, *], Suwimon Wongwanich[2], and Chayut Piromsombat[3]

Abstract

This study aimed to develop an instrument for measuring student learning and to establish a learning index for high school students. A sample of 1,619 Cambodian high school students was selected using a stratified random sampling technique. Data were collected through a 56-item questionnaire based on a 5-point Likert scale and were analyzed using various statistical methods including descriptive statistics, content analysis, objectivity analysis, correlation analysis, construct validity (using M-plus), reliability analysis (using the R-package for Windows), and t-test analysis. The instrument consists of two key components: "Learning to Know" and "Learning to Do", each further broken down into 'Process' and 'Outcome' elements. Each of these sub-components were represented by three distinct indicators. The instrument demonstrated good content validity with an Item-Objective Congruence (IOC) index ranging between 0.50 and 1.00, and excellent construct validity, supported by a variety of goodness-of-fit indices (CFI = 1.00, TLI = 0.99, SRMSR = 0.01, RMSEA = 0.02). The instrument also showed high internal consistency with reliability coefficients ranging from 0.83 to 0.93. The criterion-related validity was confirmed through known-groups validation. Two methods—criterion-related and norm-related—were used to determine students' learning index levels. For this study, the norm-related method was chosen. The learning index was categorized into four levels: low (0.000 - 0.062), medium (0.063 - 0.375), relatively high (0.376 - 0.680), and high (0.681 - 1.000). Percentile ranks were also calculated to provide additional context for interpreting the learning indices of Cambodian high school students.

Keywords: learning index, learning to know, learning to do, learning measurement instrument, Cambodian students

1. INTRODUCTION

Student learning is an intricate and multidimensional phenomenon, which extends its influence across various domains including the emotional, physical, and spiritual aspects of students' lives (Saisana, 2008; UNESCO, 2014). The emphasis on student learning transcends the traditional confines of academic achievement to represent a cornerstone for lifelong skills, employability, and personal development Delors, 1996; UNESCO, 2014).

The dimensions of student learning are manifold, incorporating diverse learning

[1,*] Dr. Bunhe Harth is currently working as a Deputy Director of Kampong Spue Institute of Technology, Cambodia. He obtained a master's degree in Educational Measurement and Evaluation from Chulalongkorn University, Thailand. He holds a Ph.D. in Educational Research Methodology in the Department of Educational Research and Psychology, Chulalongkorn University, Thailand. Email: bunherupp@gmail.com

[2] Prof. Dr. Suwimon Wongwanich is currently working as a lecturer in the Department of Educational Research and Psychology, Chulalongkorn University, Thailand. She obtained a Ph.D. in Research and Evaluation from Ohio State University, USA. Email: wsuwimon@gmail.com

[3] Asst Prof. Dr. Chayut Piromsombat is currently working as a lecturer in the Department of Educational Research and Psychology, Chulalongkorn University, Thailand. He obtained a Ph.D. in Quantitative Methods in Education from the University of Minnesota-Twin Cities, USA. Email: chayut.p@chula.ac.th

styles, attentional engagement, and other cognitive and non-cognitive elements which contribute to educational success (Brunvand & Byrd, 2011; McLean et al., 2016; OECD, 2004 Wirth & Perkins, 2008). Previous research, such as work by the Canadian Council on Learning (2010), and Kim (2016), has attempted to quantify aspects of student learning through composite indices. However, these existing models primarily offer a narrow scope, concentrating largely on measurable academic achievements and not necessarily on the comprehensive, intricate processes that underlie how students learn (Resource, & Guide, 2008; Royal Government of Cambodia, 2015).

This study aims to address the limitations of existing frameworks by developing a nuanced instrument specifically tailored to measure the components of "Learning to Know" and "Learning to Do", as conceptualized within UNESCO's four pillars of education. This research endeavors to move beyond traditional assessment paradigms that focus exclusively on end-point academic performance, aiming instead for a more holistic understanding of the learning process. Consequently, a Composite Learning Index will be designed and employed to evaluate the learning experiences of Cambodian students.

The research methodology will hinge upon a set of pair-developed indicators, correlating elements of the learning process to specific learning outcomes. Each pair aims to encapsulate one aspect of the process and its resultant outcome, thereby enabling a more granular and actionable analysis of student learning. The rigorous methodology has been constructed in alignment with emerging paradigms and ongoing changes in educational theory and practice (Delors, 2013, Rany, Zain, & Jamil, 2012).

In summary, this paper will contribute a novel instrument and index to the existing body of knowledge surrounding student learning. The instrument is intended to serve educators, students, and educational stakeholders by providing actionable data, thereby fostering more effective teaching and learning strategies. The ensuing sections detail the literature review, methodology, data collection, and analysis techniques, culminating in a discussion and conclusion that encapsulate the research findings.

2. LITERATURE REVIEW

2.1 Definition of Learning Index

Education aims to foster not just academically competent students but also socially responsible individuals (Scatliff & Meier, 2012). The concept of the learning index has gained prominence in the context of 21st-century skills and global connectivity (Allan & Charles, 2015; Caruana, 2014; Resource, & Guide, 2008; Royal Government of Cambodia, 2015).

In psychological research, students' learning has been described as an ongoing process influenced by adaptive behavioral changes based on past experiences (De Houwer, Barnes-Holmes, & Moors, 2013; Lachman, 1997). This adaptation evolves over time, reflecting the dynamic nature of the modern world (De Houwer, Barnes-Holmes & Moors, 2013). Learning can thus be segmented into three components: change in behavior, environmental regularities, and the causal relationship between the two (De Houwer, Barnes-Holmes & Moors, 2013; Lachman, 1997). This serves various purposes, from enhancing employability to positively contributing to overall well-being (Hoskins & Mascherini, 2009; Lachman, 1997).

The learning index is formulated as a measure that encompasses two key domains inspired by UNESCO's framework (Delors, 1998): Learning to Know and Learning to Do. It serves as a gauge for understanding and assessing learning processes and outcomes across different stages and environments of life (Cappon & Laughlin, 2013).

2.2 Effects of Learning

The Learning Index is not just a mere academic metric but serves as a broad assessment tool, measuring student learning at a granular level across community, regional, and national landscapes (Cappon & Laughlin, 2013). As such, it plays a pivotal role in students' lives, spanning from schooling to adulthood, impacting various facets such as economic stability, civic engagement, and overall well-being (Delors, 2013; Guerra, Modecki & Cunningham, 2014).

2.2.1 Educational Objectives and Skills

Learning is essential for fulfilling the educational objectives outlined in academic curricula. Beyond academic knowledge, it promotes critical thinking and problem-solving skills, which are critical in today's fast-paced world (Cooke & Schienstock, 2000; Wang & Degol, 2014).

2.2.2 Economic and Social Impact

Learning transcends the educational realm to bring tangible economic and social benefits such as improved employability, income, and societal health (Guerra, Modecki & Cunningham, 2014; Saisana, 2008). In the context of 21st-century competencies, it equips students to adapt and thrive in rapidly changing markets and social structures (Cooke & Schienstock, 2000; Resource & Guide, 2008).

2.2.3 Cognitive and Behavioral Effects

Different motives and strategies shape the learning experiences of students. A quality instructional environment can cultivate a mindset focused on critical thinking and social cohesion (Biggs, 1993, 1996; Wang & Degol, 2014). For instance, students who engage deeply in a subject matter are more likely to develop problem-solving skills and social cohesion (Biggs, 1993, 1996).

2.2.4 Organizational and Personal Alignment

The Learning Index will also serve as an indicator of how well students internalize and align with organizational goals and values. This alignment is integral for achieving the objectives laid out in educational action plans (Chalofsky & Krishna, 2009).

2.3 Measurement Model of Learning

The measurement model of the Learning Index is grounded in UNESCO's four pillars of lifelong learning, adapted for various cultural contexts, including a revised Chinese version (Delors, 1996, 2013; Kim, 2016). The index aims to acclimate students to changing environments, focusing on both physical and psychological well-being, and understanding the contexts in which learning occurs (Elfert, 2015). The model encapsulates the core of the four pillars, translating them into actionable teaching and learning strategies (Kim, 2016; Nan-Zhao, 2005).

2.3.1 Goal of 21st Century Learning

In an increasingly globalized world, the learning goals of the 21st century consist of four key components that help students adapt to professional challenges and to coexist peacefully in diverse communities (Cappon & Laughlin, 2013; Delors, 1996; Kim, 2016). For the purposes of this study, the focus will be on two primary factors: Learning to Know and Learning to Do.

2.3.2 Learning to Know

First conceptualized by UNESCO in 1972, "Learning to Know" aims for comprehensive development of human potential (Delors, 1996; Tawil & Cougoureux, 2013). It extends beyond conventional knowledge acquisition, covering a broad spectrum of cognitive skills, including reasoning, imagination, problem-solving, and critical thinking (Darling-Hammond, 2008; Trilling & Fadel, 2009). The learning process in this domain is continuous, enriched by a plethora of experiences throughout school and daily life (Kim, 2016; Pianta & Hamre, 2009). This mode of learning is foundational for a lifetime of continuous educational opportunities (Breen & Jonsson, 2005; Opfer & Pedder, 2011). It is both a means and an end, providing individuals with a nuanced understanding of various empirical phenomena, including nature, humanity, society, and global citizenship (Allan & Charles, 2015; Hartman, 2015; Nan-Zhao, 2005).

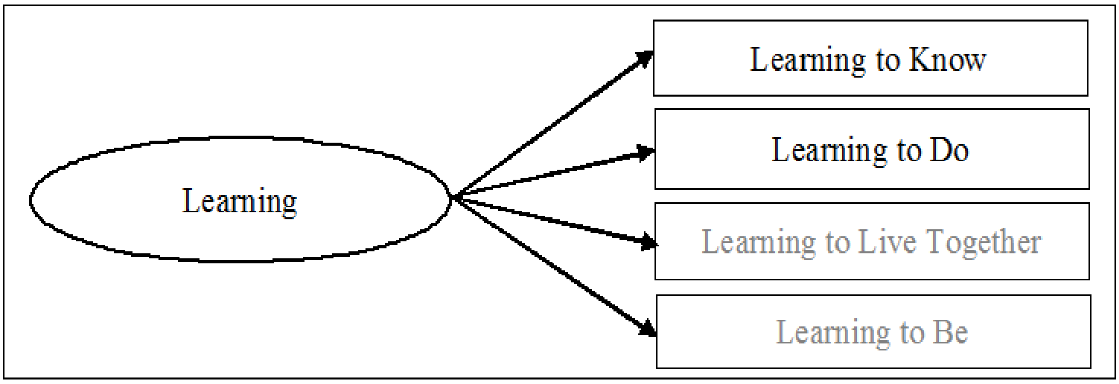


Figure 1 Measurement Model for Assessing Students' Learning

2.3.3 Learning to Do

"Learning to Do" is primarily focused on vocational skills. It emphasizes the experiences and skills necessary for students to secure a job or trade (Delors, 1996, 2013). This pillar aids students in adapting to a multitude of future scenarios, market demands, and global challenges.

This educational strategy is closely tied to vocational and technical training, seeking to transform academic knowledge into economic or professional applications (Hattie, Biggs, & Purdie, 1996; Nan-Zhao, 2005;). "Learning to Do" evolves skills into competencies, incorporating higher-order cognitive functions specific to individual professional needs (De Houwer, Barnes-Holmes & Moors, 2013). Moreover, it fosters interpersonal skills, adaptability, innovative thinking, and problem-solving abilities that are essential in the modern work environment (Barbetta, Norona, & Bicard, 2005).

2.4 Development of Educational Indicators

In developing educational indicators, the ultimate goal is to provide a framework for assessing the quality of education (Smith, Sheppard, Johnson, & Johnson, 2005; Smith, Davis, & Kim, 2020). Though widespread, the concept and development of indicators remain subject to debate, often due to issues such as a lack of standard methodology and subjectivity. This review explores the meaning of indicator, how one is developed, and their utility in the educational sector.

Purpose and Objectives

At the outset of indicator development, researchers must define the goals and objectives (Williams, 2019), focusing on the specific needs the indicator aims to address. For instance, indicators could be designed to evaluate community accountability in the educational context (Brown & Johnson, 2021), offering stakeholders valuable insights for improvement.

Defining Indicators

Academics have proposed various definitions of indicator (Smith, Davis, & Kim, 2020), which depend on the context. For example, indicators might be numerical values tied to measurable quantities (Williams, 2019), or statistical tools which monitor complex conditions otherwise hard to gauge through observation alone (Jones, 2018).

Steps in Indicator Development

1. Scope Definition

Determining the scope of indicators sets the stage for their development (Brown & Johnson, 2021). The framework should respond to user needs and be rooted in specific goals (Williams, 2019). Various criteria should be outlined for indicator selection, such as sustainability and relevance (Smith, Davis, & Kim, 2020), which will facilitate better evaluation.

2. Selecting Criteria for Indicators

In selecting indicators, both qualitative and quantitative measures should be considered (Jones, 2018). While qualitative measures might focus on community spirit and values, quantitative measures could deal with numerical assessment (Williams, 2019). Indicators are then evaluated against a set of predetermined criteria (Brown & Johnson, 2021), offering a balanced judgment based on empirical evidence.

3. Quality Assurance

Ensuring quality involves checking an indicator's reliability, validity, and feasibility (Smith, Davis, & Kim, 2020). Researchers aim to define a set of quality criteria that the indicator should meet, focusing on factors such as timeliness, relevance, and objectivity (Jones, 2018).

4. Piloting Indicators

To pilot an indicator is to test the indicator properties in terms of practice to determine the quality of the indicator, its feasibility and appropriate utilization (Giambona & Vassallo, 2014).

Data Collection and Analysis

Prior to full-scale implementation, indicators should be tested for feasibility and relevance through pilot studies (Williams, 2019). These trials help to refine the indicator by analyzing data and aligning them with the intended context and time frame (Brown & Johnson, 2021).

Documenting Results

Proper documentation is essential for interpreting the data collected (Smith, Davis, & Kim, 2020). This will involve detailing the characteristics of the indicators and the metrics used, as well as providing guidelines for future research and development (Jones, 2018).

Communicating Findings

Transparency is key to the final step (Williams, 2019). Whether through report cards, summary reports, or technical documents, the findings should be communicated effectively to all stakeholders involved (Brown & Johnson, 2021). This enables users to understand how to use the indicators to improve their educational systems.

By clearly defining its scope, selecting relevant criteria, and piloting the indicators, this research can lead to the development of robust tools that not only offer a glimpse into the educational landscape but also guide stakeholders toward meaningful improvement (Smith, Davis & Kim, 2020).

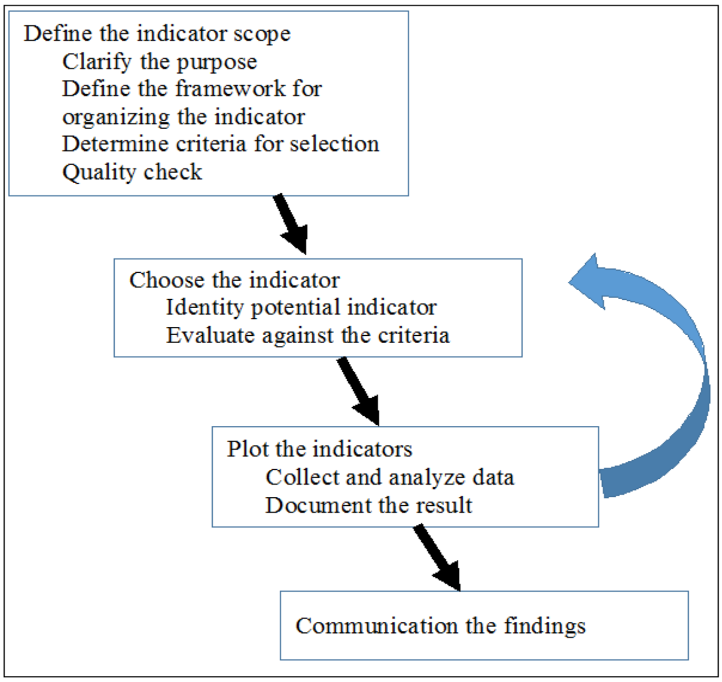


Figure 2 Indicator Development Process

3. CONCEPTUAL FRAMEWORK OF STUDENTS' LEARNING

As previously mentioned, the student learning which is studied in this research will be extracted from the four pillars of education proposed by UNESCO. The first two main components of the four pillars will be considered. Thus, the conceptual framework of student learning is outlined as follows:

Students' learning will be studied within two main components—Learning to Know and Learning to Do. These two main components each consist of two sub-components. Thus, students' learning will be measured by two main components and four sub-components as shown in Figure 3.

4. METHODOLOGY

This research aims to develop a measurement model of students' learning and to determine a learning index for students in Cambodia. The study follows a two-step approach. The first step involves the development of the components of students' learning and the instrument used to measure it. The second step focuses on validating the psychometric properties of the instrument in terms of content validity, objectivity, uncertainty, construct validity, reliability, and criterion-related validity (Button, Ioannidis, Mokrysz, Nosek, Flint, Robinson & Munafò, 2013).

Population

The study population was comprised of 296,907 students enrolled in high schools across Cambodia, encompassing both private and public institutions (Royal Government of Cambodia, 2015).

Sample Size and Selection

The sample for the research consisted of 1,619 private and public high school students. Due to the application of structural equation modeling in this research, the sample size was determined using a prior sample size calculator specifically designed for structural equation models (Button, Ioannidis, Mokrysz, Nosek, Flint, Robinson, & Munafò, 2013). The calculator recommends how to calculate sample size based on the number of items and unobserved variables used in the analysis.

Initially, the study aimed to collect data from approximately 2,000 students. Questionnaires were distributed across the five regions of Cambodia. To account for potential missing data during the data collection process, an additional 30% was added to the sample size, bringing the total desired sample size to 2,950 students. The sample was obtained using a multi-stage random sampling technique.

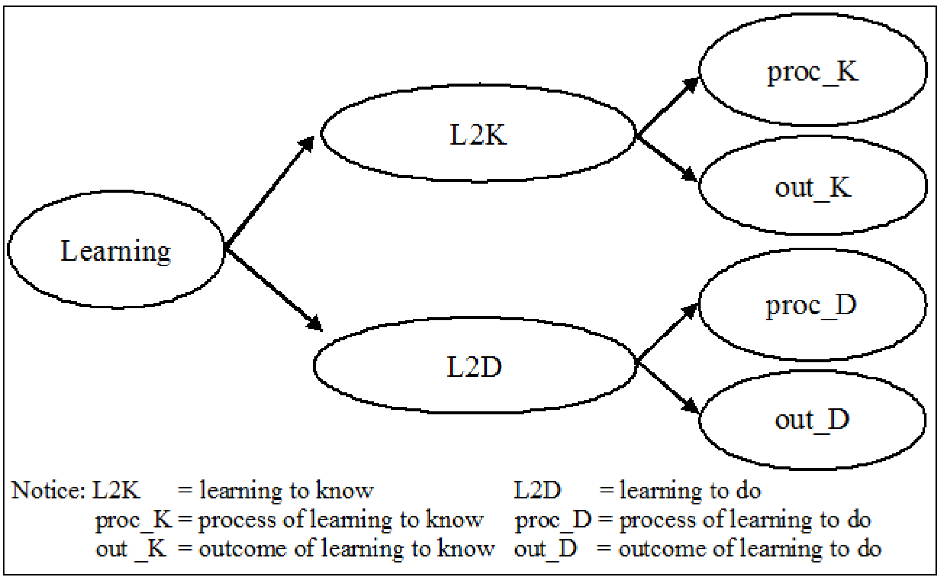


Figure 3 Conceptual Framework of Students' Learning

Data Collection and Analysis

The line database developed in this research aims to offer a more reliable source for constructing a learning index specific to Cambodian students. Consequently, a larger sample size was necessary. The data analysis consisted of two primary techniques: descriptive data analysis and inferential data analysis.

The first method involved examining the background information of the observed variables. This was done through calculating percentages, means, standard deviations, maximum scores, and minimum scores, using the SPSS program for Windows. This provides a summary of the respondents' background information which can be used to understand the distribution of the data.

To develop composite scores for students' learning, several steps were required, including the replacement of missing data, definition of indicator loadings, and summarization of individual indicator values (Manthalu, Nkhoma, & Kuyeli, 2010). Various analysis techniques were employed to qualify the composite scores, such as uncertainty analysis (Nardo & Saisana, 2008).

1. Weighting: Two methods were employed for weighting—equal loadings for individual sub-indicators and individual loading of sub-indicators based on loadings obtained from confirmatory factor analysis (Foa & Tanner, 2012; Zhou & Ang, 2009).

2. Aggregation Methods: Two main models were used for aggregation—the additive model and the multiplicative model.

2.1 Additive Model: This involves linearly summing all individual sub-indicator values (Cherchye, Moesen, Rogge, & Van Puyenbroeck, 2007; Nardo & Saisana, 2008; Nardo, Saisana, Saltelli, & Tarantola, 2005).

$$CI_c = \sum_{i=1}^{m} w_{c,i} \cdot Y_{c,i}^{n}$$

Where CI_c is the composite score for individual student i's learning, $Y_{c,i}^{n}$ is the (possibly normalized) value for individual student i on indicator I (i=1,…,m) and w_i is the weight assigned to indicator i. Weights are generally bounded according to $0 \le w_{c,i} \le 1$ and $\sum_{i=1}^{m} w_{c,i} = 1$.

2.2 Multiplicative Model: This model multiplies individual sub-indicators to generate a summarizing indicator value (Nardo & Saisana, 2008; Nardo, Saisana, Saltelli, & Tarantola, 2005).

$$CI_c = \Pi_{i=1}^{m} x_{i,c}^{w_i}$$

Based on these models, there are 4 sources of uncertainty (2 x 2) in developing the composite scores for students' learning. Accordingly, an analysis of model uncertainty was employed for the 4 models.

After employing the uncertainty analysis, the most appropriate models for developing the students' learning indices could be selected, resulting in four specific indices: the process of Learning to Know, the outcomes of Learning to Know, the process of Learning to Do, and the outcomes of Learning to Do. These indices allowed for the calculation of norm-referenced and percentile criteria in the presented results.

5. RESULTS

5.1 Psychometric Properties of the Students' Learning Measurement Instrument

Content Validity

In this study, the IOC index for each item ranged from 0.75 to 1.00, with the exception of item 52, which had an IOC index of only 0.50. Despite this outlier, the measurement model for students' learning was confirmed to have strong content validity, after also being reviewed and validated by four experts in the field.

Objectivity

The objectivity of the research instrument was rigorously examined over an extended period during an academic class on thesis report writing. Subsequently, it was further scrutinized by professional experts in the field of education to glean specific insights into the quality and relevance of the developed items. This assessment was carried out by a total of 12 experts, comprising 8 graduate and doctoral students from the program and an additional 4 experts.

The consensus among the experts was that the items in the measurement model for students' learning met all required criteria. The experts opined that each item was crafted with objectivity, focusing on three technical aspects: language use, scoring criteria, and interpretative guidelines. This objectivity ensures that the items effectively measure the concepts they were designed to assess, aligning with the operational definitions provided. The detailed criteria for evaluating questionnaire objectivity are presented in Table 1.

Uncertainty

Uncertainty analysis serves as a specialized statistical tool designed to examine the propagation of uncertainties in the input variables through to the composite scores of students' learning. The input values affecting the composite scores can originate from two major sources: weighting schemes and aggregation methods.

Composite scores for "Learning to Know" and "Learning to Do" could be influenced by two types of weighting schemes: equal and non-equal weightings. These weightings were derived from fixed loadings and from the loadings of the confirmatory factor analysis. Additionally, two primary aggregation methods were employed: additive and multiplicative methods. Consequently, the analysis model for developing composite scores incorporated four distinct models, aligning with the research objectives. These four models serve to calculate the final composite scores and were employed in an alternative analysis to yield more reliable and credible composite scores of students' learning, as illustrated in Table 2.

Table 1 Objectivity Check of Items of Students' Learning

Objectivity	Experts' opinion
1. Language use	Clear language use was implemented in the questionnaire with some items stated in words or phrases, even though they should be sentence level statements. Item statements were made to be appropriate for high school students of the intended-age group, such that they would be able to express their opinion on items provided. Additionally, the questionnaire was written such that it should be understood by all age intervals of high school students.
2. Scoring check	Items of both factors of the measurement model of students' learning were measured from summated items on a 5-point Likert-type rating scale, providing objectivity for the scoring check even if someone's item scores were scored the same as one another. Thus, the scoring check was appropriate and provided the same scores for every respondent.
3. Interpretative criteria of the score	Items of the measurement model of students' learning for both the process and outcome of learning (to know and to do) consisted of clearly interpreted criteria, when considering the mean scores of each item. The item scores were nested under the indicators, while the indicators were nested under the components. The components were nested under the factors, making it convenient for the respondents to score the items.

Table 2 A Development of Composite Scores of Students' Learning

Model	Weighting	Aggregation
1	equal loading	additive model
2	equal loading	multiplicative model
3	factor loading	additive model
4	factor loading	multiplicative model

After applying the weights and conducting alternative aggregation of the items, the composite scores of students' learning were ranked for individual samples. This ranking was primarily based on background information such as school locations and the jurisdiction of schools. These variables were primarily employed in the macro-level analysis to study the composite scores of students' learning. The correlation rankings of the composite scores were subsequently computed and can be seen in Table 3. These rankings were generated based on the determination made in the preceding uncertainty analysis concerning the weighting schemes and aggregation methods.

Based on the analysis of uncertainties, the coefficients of variation (CV) for the component or factor thresholds and the percentages of unchanged factors in the composite scores of students' learning were assessed. In this study, the CVs exhibited no variation between the component or factor thresholds, suggesting that there is minimal uncertainty between the four different models used for calculating the composite scores of students' learning.

The correlation coefficients for the four combination composite scores of students' learning are presented in Table 4. These coefficients ranged from 0.75 to 0.99 and were statistically significant at an alpha level of 0.05. These high coefficients suggest that the four combination models measure the same objectives harmoniously, as initially intended.

As previously stated, this uncertainty analysis aims to identify the most critical weighting schemes and aggregation methods for determining composite scores of students' learning. Accordingly, the four different combinations of weighting scheme and aggregation method were examined to answer the research question, with the findings being assessed in terms of correlation coefficients.

According to Table 4, shifts in the composite scores for students' learning occurred when changing from equal to non-equal weights. Specifically, for the additive model, scores ranged from 0.99, while for the multiplicative model, they ranged down to 0.75. This suggests that the composite scores are more sensitive to variations in the aggregation methods used (additive or multiplicative) than they are to changes in the weighting scheme (equal or non-equal weights). Therefore, future implementations of composite scores can confidently utilize either weighting scheme without significantly impacting the results.

Table 3 Min, Max, Mean, and Variance of Composite Score Rank

Discriminant variable	Rank			coefficient of variation (CV)
	Min	Max	Mean	
school location				
urban	1.00	1.00	1.00	0.00
nonurban	2.00	2.00	2.00	0.00
school under admin				
public school	1.00	1.00	1.00	0.00
private school	2.00	2.00	2.00	0.00

Table 4 Correlation Coefficients of Value and Rank of Composite Score of 4 Models

Model	1	2	3	4
1	1.00			
2	.99**	1.00		
3	.99**	.99**	1.00	
4	.75**	.75**	.75**	1.00

Notice. ***p*<.01 coefficients under diagonal are composite scores of composite scores of students' learning values of the 4 models.

These findings also demonstrate that the composite scores are robust across different methods of calculation. Four distinct models for developing composite scores all yielded highly consistent results. Within the context of this research, the third model, which utilizes an additive method with factor loadings, is recommended for gauging these specific composite scores for students' learning. This model was selected as the most appropriate because it is not only convenient for empirical application but also easy to interpret. Moreover, it aligns well with established measurement and evaluation principles, offering high reliability coefficients. This model is also attuned to unequal factor loadings, making it consistent with empirically-based practices and loadings derived from confirmatory factor analysis (CFA). The development of reliable and valid composite scores for students' learning requires a thoughtful approach that integrates both theoretical and empirical considerations. This study meets such criteria, being informed by the expertise of four key professionals in the fields of educational measurement, evaluation, policy, and management, as well as eight experts in educational research methodology at the master's and doctoral levels. The high reliability coefficients, ranging from 0.83 to 0.93, further substantiate the rigor of this study. Additional methods of calculation were also explored to validate these composite scores for students' learning.

Construct Validity

Second-order confirmatory factor analysis was utilized to establish the construct validity of the measurement model for students' learning. This analysis was applied to two distinct models—'Learning to Know' and 'Learning to Do'. Each model was independently validated based on its associated factors related to students' learning.

For the 'Learning to Know' model, a second-order confirmatory factor analysis was performed, revealing the following goodness-of-fit indices: Chi-square (2, N = 1619) = 3.04, at a p-value of 0.05, comparative fit index (CFI) = 1.00, Tucker Lewis Index (TLI) = 0.99, standardized root mean square residual (SRMSR) = 0.01, and root mean square error of approximation (RMSEA) = 0.02 (Hu & Bentler, 1998).

These results indicate that all the goodness-of-fit indices met the predetermined criteria. Therefore, it can be concluded that the 'Learning to Know' measurement model possesses adequate construct-validity. Specific statistical values from the confirmatory factor analysis are presented in the subsequent section.

For the "Learning to Do" model, a second-order confirmatory factor analysis (CFA) was conducted. The analysis yielded the following goodness-of-fit indices: chi-

Table 5 CFA Result of Learning to Know Model

Variable	Factor loadings		t	R^2	Factor score coefficient
	β	(SE)			
First order CFA_know					
proc_K_desi	.82	.00	113.22**	.67	.44
proc_K_enga	.72	.01	50.25**	.51	.01
proc_K_learn	.76	.01	78.86**	.58	.02
out_K_desi	.79	.01	71.50**	.63	.28
out_K_enga	.82	,01	71.64**	.68	.32
out_K_learn	.76	.02	46.42**	.57	.08
Second-order CFA_know					
proc_K	.96	.00	465.17**	.92	.01
out_K	.99	.00	1857.59**	.98	.01

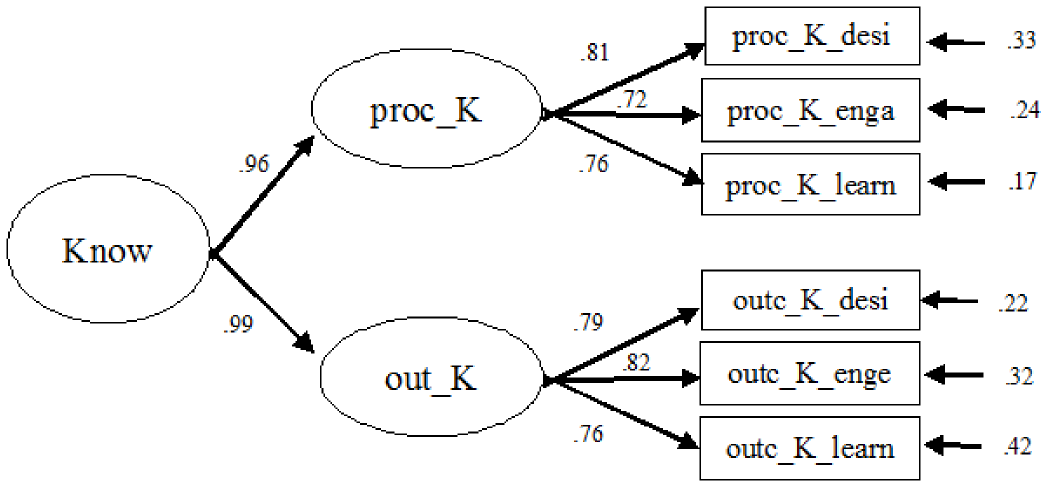
Chi-square (2, N = 1619) = 3.04, p = .05, CFI = 1.00, $SRMSR$ = .01, $RMSEA$= .02.

square (4, N=1619) = 7.41, at a p-value of 0.12, comparative fit index (CFI) = 1.00, Tucker Lewis Index (TLI) = 0.99, standardized root mean square residual (SRMSR) = 0.01, and root mean square error of approximation (RMSEA) = 0.02. Given that all the fit indices met the predetermined criteria, it can be confirmed that the 'Learning to Do' measurement model possesses adequate construct-validity. Detailed statistical values are provided in Table 6.

In addition to the individual second-order measurement models for "Learning to Know" (L2K) and "Learning to Do" (L2D), a more comprehensive, higher-order measurement model was developed, known as the "Measurement Model of Students' Learning". This overarching model incorporates both L2K and L2D and was validated using third-order confirmatory factor analysis (CFA).

The L2K model contains two main components: the "Process of Learning to Know" and the "Outcome of Learning to Know". Similarly, the L2D model consists of the "Process of Learning to Do" and the "Outcome of Learning to Do."

Process of Learning to Know: This includes three indicators: learning desire, learn-



Chi-square (2, N = 1619) = 3.04, p = .05, *CFI* = 1.00, *SRMSR* = .01, *RMSEA*= .02

Figure 4 Measurement Model of Learning to Know

Table 6 CFA Results of the Learning to Do Model

Variable	Factor loadings		t	R^2	Factor score coefficient
	β	(SE)			
First order CFA_know					
proc_D_conc	.83	.01	127.25**	.68	.33
proc_D_prac	.87	.01	101.09**	.76	.22
proc_D_cont	.87	.01	99.39**	.75	.25
out_D_conc	.83	.01	93.09**	.69	.06
out_D_prac	.88	.01	110.99**	.77	.07
out_D_cont	.86	.01	102.39**	.74	.27
Second-order CFA_do					
proc_D	.95	.01	394.64**	.91	.01
out_D	.96	.01	556.66**	.93	.01

Chi-square (4, N=1619) =7.41; CFI= 1.00; SRMSR = .01; p = .05; RMSEA= .02.

ing engagement, and learning to learn.

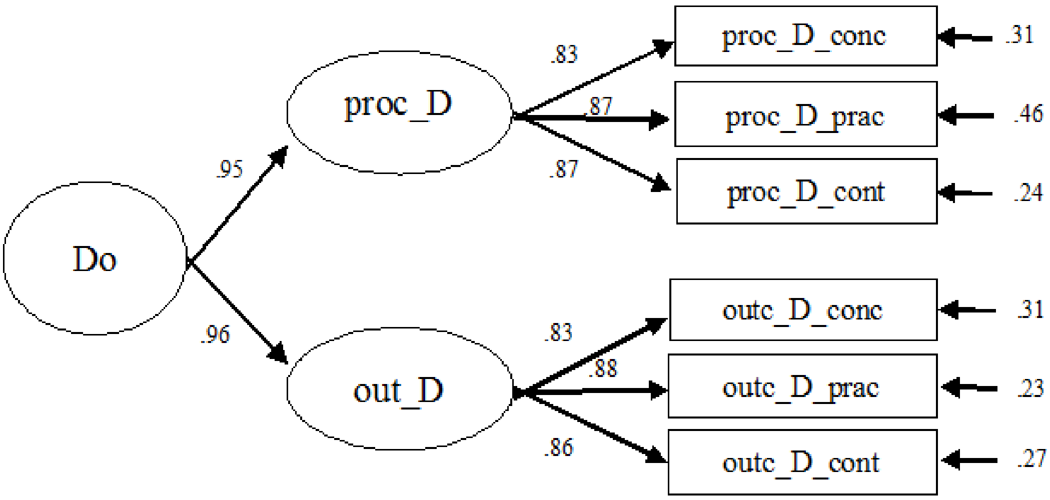
Outcome of Learning to Know: This comprises three indicators: the outcomes of learning desire, learning engagement, and learning to learn.

Process of Learning to Do: Features three indicators: concern for learning in real-world work settings, practical engagement, and continuing self-development.

Outcome of Learning to Do: Includes three indicators: outcomes in real-world work settings, practical engagement, and continuing self-development.

Each process and outcome indicator for both L2K and L2D generally consists of four items. However, the third indicator of both the "Process" and "Outcome" elements of "Learning to Know" contains eight items. As a result, the overarching measurement model is built on 2 factors, 4 components, 12 indicators, and 56 items in total.

The third-order CFA results showed that the higher-order measurement model of students' learning is valid. This was confirmed by the following goodness-of-fit indices: chi-square (15, N=1619) = 22.32, with a p-value of 0.10, comparative fit index (CFI) = 1.00, Tucker Lewis Index (TLI) = 0.99, standardized root mean square residual (SRMSR) = 0.01, and root mean square error of approximation (RMSEA) = 0. 02. Detailed statistical values are presented in Table 7.



Chi-square (4, N=1619) =7.41; CFI= 1.00; SRMSR = .01; p = .05; RMSEA= .02.

Figure 5 Measurement Model of Learning to Do

Table 7 CFA Model of the Students' Learning Index

Variable	Factor loadings		t	R^2	Factor score coefficient
	β	(SE)			
First order of learning					
proc_K_desi	.82	.01	114.56**	.67	.34
proc_K_enga	.83	.01	124.12**	.68	.09
proc_K_learn	.91	.01	285.96**	.83	.30
out_K_desi	.87	.01	186.19**	.76	.35
out_K_enga	.93	.01	75.68**	.87	.34
out_K_learn	.95	.01	563.63**	.90	.52
proc_D_conc	.85	.01	144.43**	.72	.30
proc_D_prac	.83	.01	126.83**	.69	.03
proc_D_cont	.82	.01	70.09**	.67	.08

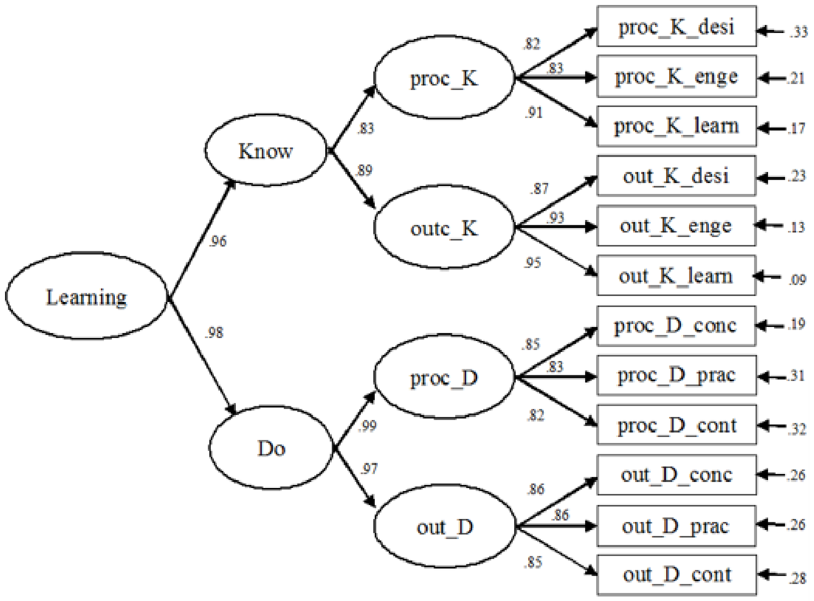
Table 7 (Continued)

Variable	Factor loadings β	(SE)	t	R^2	Factor score coefficient
out_D_conc	.86	.01	97.27**	.74	.09
out_D_prac	.86	.01	92.15**	.74	.19
out_D_cont	.85	.01	83.13**	.72	.20
Second order of learning					
proc_K	.83	.01	58.20**	.68	.05
out_K	.89	.01	183.56**	.81	.06
proc_D	.99	.01	74.99**	.56	.08
out_D	.96	.01	607.65**	.94	.01
Third order of learning					
Know	.96	.01	399.55**	.93	.01
Do	.98	.01	785.70**	.96	.01
Chi-square (15, N=1619) = 22.32; *p* = .10; RMSEA = .02; SRMSR = .01; CFI = 1.00					

Table 8 CFA Model of the Students' Learning Index

Variable	Factor loadings β	(SE)	t	R^2	Factor score coefficient
First order of learning					
proc_K_desi	.82	.01	114.56**	.67	.34
proc_K_enga	.83	.01	124.12**	.68	.09
proc_K_learn	.91	.01	285.96**	.83	.30
out_K_desi	.87	.01	186.19**	.76	.35
out_K_enga	.93	.01	75.68**	.87	.34
out_K_learn	.95	.01	563.63**	.90	.52
proc_D_conc	.85	.01	144.43**	.72	.30
proc_D_prac	.83	.01	126.83**	.69	.03
proc_D_cont	.82	.01	70.09**	.67	.08
out_D_conc	.86	.01	97.27**	.74	.09
out_D_prac	.86	.01	92.15**	.74	.19
out_D_cont	.85	.01	83.13**	.72	.20
Second order of learning					
proc_K	.83	.01	58.20**	.68	.05
out_K	.89	.01	183.56**	.81	.06
proc_D	.99	.01	74.99**	.56	.08
out_D	.96	.01	607.65**	.94	.01
Third order of learning					
Know	.96	.01	399.55**	.93	.01
Do	.98	.01	785.70**	.96	.01
Chi-square (15, N=1619) = 22.32; *p* = .10; RMSEA = .02; SRMSR = .01; CFI = 1.00					

The analysis results of the third-order confirmatory factor analysis revealed that the goodness-of-fit index of the measurement model of students' learning was appropriate and adequately fitted with the empirical data collected as shown in the Table 8.



Chi-square (15, N=1619) = 22.32; p=.10; RMSEA = .02

Figure 6 Measurement Model of Students' Learning

Reliability

Reliability coefficients were examined through a pilot study process involving 25 Cambodian high school students on measurement model of students' learning. The adjustment of items in the questionnaire was trialed based on experts' recommendations. In this process, the reliability coefficient was examined using Cronbach's Alpha Coefficients (Gliem & Gliem, 2003).

The analysis indicated high levels of internal consistency across all indicators, indicating that the questionnaire met the acceptability criteria. Reliability coefficients ranged from 0.83 to 0.93. Notably, the indicator "Learning to Learn" had the highest reliability coefficient at 0.93, while "Concern as the Real World of Work" had the lowest at 0.83. When examined at the component level, reliability coefficients were consistently high, ranging between 0.94 and 0.96. This suggests excellent internal consistency across the four components of the "Measurement Model of Students' Learning."

In summary, the research instrument was rigorously validated in terms of content validity, objectivity, uncertainty, construct validity, and reliability. Based on these results, it can be concluded that the "Measurement Model of Students' Learning" possessed adequate construct-validity. Furthermore, the model can be said to be well-suited for the study context and characteristics of the respondents.

Table 9 Reliability Coefficients of Students' Learning

Indicators in the research instrument		Items	Reliability Coefficients
learning desire	proc_K_desi	4	.89
learning engagement	proc_K_enga	4	.89
learning to learn	proc_K_learn	**8**	.93
process of Learning to Know		**16**	**.96**
outcome of learning desire	out_K_desi	4	.80
outcome of learning engagement	out_K_enga	4	.85
learning outcomes of learning to learn	out_K_learn	**8**	.91
outcome of Learning to Know		**16**	**.94**
concern as the real world of work	proc_D_conc	4	.83

Table 9 (Continued)

Indicators in the research instrument	Items	Reliability	Coefficients
practical engagement	proc_D_prac	4	.92
continuing self-development	proc_D_cont	4	.86
process of Learning to Do		**12**	**.94**
outcomes concern as the real world of work	out_D_conc	4	.87
outcome of practical engagement	out_D_prac	4	.92
outcome of continuing self-development	out_D_cont	4	.86
outcome of Learning to Do		**12**	**.95**

Table 10 Mean, SD, Min, Max of Students' Composite Learning Classified by Known-Groups

Known-Group	Full score	Mean	SD	CV	Min	Max	t-test
Low level (n=12)	60	37.72	3.51	9.31	31.88	42.32	t=3.82, df=22, p<.05
High level (n=12)	60	43.06	3.34	7.76	37.25	48.63	

Criterion-related Validity

Students were divided into two distinct groups based on their composite learning scores: a low-level group (12 respondents) and a high-level group (12 respondents). These groups were combined to form a 'known-group' of students' learning. A t-test was conducted to determine whether the composite scores differed significantly between the two groups. The results are summarized in Table 10.

The analysis was performed on 12 samples from each group, each consisting of two main components: "Learning to Know" and "Learning to Do." The average composite score for the low-level group was 37.72, while the high-level group had an average score of 43.06. A t-test revealed a statistically significant difference between the two groups ($t = 3.82$, $df = 22$, $p < 0.05$). The significant difference in means between the low and high-level groups confirms criterion-related validity for the instrument, using the known-group technique.

5.2 Norm of Cambodian Students' Learning

In this research, various methods were utilized to measure composite scores of students' learning, including raw composite scores, weighted composite scores, and a learning index. The learning index was assessed using two norms—criterion-referenced and norm-referenced. The purpose was to identify thresholds for different levels of composite scores.

Percentile calculations were carried out for three models—Learning to Know, Learning to Do, and the overall Learning model—using both raw and weighted composite scores. The data were categorized into 5-percentile intervals.

Learning to Know: The mean raw composite scores ranged from 15.00 at the 5^{th} percentile to 27.63 at the 95^{th} percentile. The weighted scores ranged from 13.25 at the 5th percentile to 24.37 at the 95^{th} percentile.

Learning to Do: The mean raw composite scores ranged from 14.50 at the 5^{th} percentile to 28.00 at the 95^{th} percentile. The weighted scores ranged from 12.25 at the 5^{th} percentile to 23.65 at the 95^{th} percentile.

Overall Learning: The mean raw composite scores ranged from 33.63 at the 5^{th} percentile to 55.38 at the 95^{th} percentile. The weighted scores ranged from 24.42 at the 5^{th} percentile to 47.79 at the 95^{th} percentile.

For practical purposes, it is advisable to use both raw and weighted composite scores to determine cut-off scores for levels of students' learning. Raw scores are generally easier to understand and calculate compared to weighted scores. In this research, the 25^{th},

Table 11 Norm of Students' Learning Index Classified by Raw and Weighted Scores

Variable	Learning to Know		Learning to Do		Learning	
Percentile	raw score	weighted score*	raw score	weighted score*	raw score	weighted score*
5.0	15.00	13.25	14.50	12.25	30.63	26.42
10.0	16.88	14.81	16.00	13.54	33.50	28.88
15.0	17.88	15.68	17.25	14.60	35.50	30.72
20.0	18.63	16.43	18.00	15.22	36.88	31.86
25.0	19.38	17.02	18.50	15.65	38.13	32.92
30.0	20.00	17.59	19.25	16.28	39.63	34.20
35.0	20.63	18.13	19.75	16.72	40.63	35.01
40.0	21.00	18.51	20.25	17.12	41.38	35.68
45.0	21.38	18.88	21.00	17.73	42.38	36.52
50.0	21.88	19.30	21.50	18.17	43.38	37.47
55.0	22.38	19.72	22.25	18.78	44.25	38.25
60.0	22.75	20.09	22.75	19.21	45.25	39.04
65.0	23.38	20.59	23.25	19.65	46.25	39.90
70.0	23.88	21.06	24.00	20.26	47.25	40.81
75.0	24.38	21.53	24.50	20.70	48.75	42.07
80.0	25.13	22.17	25.25	21.32	50.00	43.22
85.0	25.88	22.82	26.00	21.96	51.50	44.46
90.0	26.75	23.59	27.00	22.80	53.00	45.77
95.0	27.63	24.37	28.00	23.65	55.38	47.79

Notice. Weighted score* = composite scores with factor loadings.
Maximum composite raw score of Learning to Know and Learning to Do = 30
Maximum composite raw score of Learning = 60

50^{th}, and 75^{th} percentiles were used as cut-off points, in line with criterion-referenced norms, to establish thresholds for students' learning levels.

5.3 Learning Index

This research study employed the norm-criterion approach for assessing the learning index. The norm-criterion encompasses two techniques: criterion-referenced and norm-referenced. These techniques establish cut-off points for the composite scores of students' learning by using the minimum and maximum composite scores of each factor (Chalmers, 2012; Giovannini, Nardo, Saisana, Saltelli, Tarantola, & Hoffman, 2008).

$$Index = \frac{Value - min}{Max - min}$$

By using this formula, the index effectively scales individual composite scores between a minimum and maximum range, thereby facilitating easier comparisons and interpretations of students' learning outcomes.

The minimum and maximum individual composite scores of the students' learning scores were 12.00 and 60.00. Hence, this approach allowed for the previous formula to be transformed from a factor scale into index values of students' learning.

$$Students' learning\ Index = \frac{individual\ score - 12}{48}$$

The learning index developed in this study is classified into four levels: low, medium, relatively high, and high. This classification is rooted in two key beliefs:

The first belief posits that the processes and outcomes of "Learning to Know" serve as foundational elements for "Learning to Do."

The second belief contends that the outcomes of both "Learning to Know" and "Learning to Do" are intrinsically linked to their respective learning processes.

Cut-off composite scores for students' learning were calculated using both criterion-referenced and norm-referenced techniques, as detailed in Table 12. Specifically, composite scores for each component of students' learning were set at 15, 30, 45, and 60. The total minimum composite score for each component is 12, while the maximum is 60. The calculation of the norm-referenced index equates to the composite score of each component minus the maximum composite score divided by the total composite score minus the minimum composite score, as presented in the Table 12.

As shown in Table 12, the learning index serves as a tool to categorize students' learning levels in terms of "Learning to Know" and "Learning to Do." This index follows a "top-up" model, initiated with the "Learning to Know" process and culminated by the "Learning to Do" outcomes. Normalized scores from additive learning models can be used as a standard for assessing individual learning indexes.

The learning index is underpinned by two key beliefs: first, that "Learning to Know" serves is a foundation for "Learning to Do"; and second, that the outcomes for both are closely tied to their respective learning processes.

Interpretation of the learning index can be conducted in three ways, as shown in Tables 10 and 13:

1. Using percentile cut-off scores based on students' composite learning scores.
2. Using a criterion-referenced index.
3. Using a norm-referenced index.

All three methods provide comparable intervals for interpretation. However, the norm-referenced index is recommended for future studies for its comprehensive and consistent evaluation metrics.

Table 12 Cut-off Scores of Learning Index

Learning Level	Criterion-referenced Index		Norm-referenced Index	
	Formula	ranges	Formula	ranges
Low	15/60	.00 - .25	(15-12)/(60-12)	.000 - .062
Medium	30/60	.25 - .50	(30-12)/(60-12)	.063 - .375
Relatively High	45/60	.51 - .75	(45-12)/(60-12)	.376 - .680
High	60/60	.76 - 1.00	(60-12)/(60-12)	.681 - 1.000

Minimum score = 12

Table 13 Interpretation of Students' Learning Index

Learning Level	Criterion-referenced Index	Norm-referenced Index	Belief I **The process and outcomes of "Learning to Know" are the basis for "Learning to Do"**	Belief II **The outcomes of "Learning to Know" and "Learning to Do" are based on their learning processes**
Low	.00 - .25	.000 - .062	Having the process of "Learning to Know"	Having the process of " Learning to Know"
Medium	.25 - .50	.063 - .375	Having the process and meeting the outcomes of "Learning to Know"	Having the process of " Learning to Know" and "Learning to Do"
Relatively High	.51 - .75	.376 - .680	Having "Learning to Know" ability and the process of "Learning to Do"	Having the process of " Learning to Know" and "Learning to Do" and meeting the outcomes of "Learning to Know"
High	.76 - 1.00	.681 - 1.000	Having the process and outcomes of "Learning to Know" and 'Learning to Do"	Having the process and outcomes of "Learning to Know" and 'Learning to Do"

6. DISCUSSION

The research instrument introduced in this study offers a groundbreaking framework for educational assessment, with innovation manifest in its design, context adaptability, and psychometric properties. The discussion below delves into key areas of these contributions.

The research instrument is the first of its kind to harmonize the four pillars of UNESCO's educational model—learning to know, learning to do, learning to live together, and learning to be—into a singular framework (UNESCO, 1996). This holistic integration allows for a comprehensive evaluation of student learning, filling gaps left by traditional research instruments that may only focus on isolated learning outcomes (Brown & Johnson, 2021; Williams, 2019).

The instrument employs a unique laddering approach that assesses both the process and outcomes of learning. Initial evaluations focus on student engagement, effort, and strategy utilization, followed by a second-tier assessment on knowledge, skills, and problem-solving capabilities. This dual-layered approach enriches data collection and offers nuanced insights into student learning needs and potential areas for intervention (Williams, 2019).

The design accommodates multi-contextual and multi-lingual settings, enhancing its applicability across diverse educational landscapes. Such adaptability is indispensable for researchers and educators aiming for global impact (Brown & Johnson, 2021).

The instrument was developed through a stringent process involving expert consultation and psychometric validation. The implementation of these best practices assures the instrument's reliability and validity, thereby fortifying its potential for widespread application (Smith, Davis, & Kim, 2020; Jones, 2018).

The research instrument devised in this study sets a new precedent in educational assessment, integrating innovative methodologies with robust psychometric properties. Its unique features and versatility hold promising implications for improving student learning globally.

The validation of the research instrument was a comprehensive process involving six psychometric properties: content validity, objectivity, uncertainty, construct validity, reliability, and criterion-related validity. Content validity received scrutiny from four experts in the field, focusing on alignment with operational definitions and the necessity for multiple pieces of evidence for validation (Drost, 2011). An additional panel of 12 experts evaluated objectivity in terms of language use, scoring, and criteria. Uncertainty analysis, which utilized a third-order confirmatory factor analysis, endorsed the additive model as most suitable for the learning index (Polit & Beck, 2006).

Construct validity was assessed using third-order confirmatory factor analysis, showing high factor loadings and a good fit with the empirical data. Factorial validity is typically confirmed through confirmatory factor analysis (CFA) rather than exploratory factor analysis (EFA), as it provides a more stringent and specific evaluation (Mokkink, Terwee, Patrick, Alonso, Stratford, Knol, & de Vet, 2010; Hair, Black, Babin, Anderson, & Tatham, 2009; Polit, 2015). Structural equation modeling (SEM) is another common technique, comprising two sub-models: the measurement model and the structural model (Hair, Black, Babin, Anderson, & Tatham, 2009; Hoyle, 1999). Finally, the structural model was examined to ensure proper conceptual representation of relationships between constructs. This rigorous validation process underscores the reliability and robustness of the research instrument in measuring the targeted constructs (Hair, Black, Babin, Anderson & Tatham, 2009; Hair, Hair, Hult, Ringle & Sarstedt, 2021; Fornell & Larcker, 1981).

Reliabilities of the research questionnaire ranged from 0.83 - 0.93, indicating that all items among the factors, components, and indicators, were highly interrelated even though a small amount of items were found to be highly separated.

7. CONCLUSION AND RECOMMENDATION

This research primarily offers two key tools: a measurement instrument for assessing students' learning and a guideline for enhancing their learning index. These tools are beneficial for both educators and students for planning and tracking educational progress.

For Educational Stakeholders:

Adoption and Adaptation: The measurement model of students' learning shows a good fit between the construct and empirical data, going beyond traditional assessments by providing a more continuous understanding of student progress. Stakeholders can use criterion-related and norm-related methods for scalable and diverse measurements.

Inclusive Evaluation: Future research should consider the full range of educational pillars proposed by UNESCO to achieve a more comprehensive understanding of students' learning cycles.

Psychometric Validation: Further studies should also focus on differential item functioning (DIF) and measurement invariance for a more robust instrument.

Sensitivity Analysis: When developing the learning index, sensitivity analyses should be conducted to validate the methods used for calculating the index values.

For Educators:

Practical Application: The measurement model is practical and can convert raw scores into a student learning index, facilitating easier interpretation and planning for educators.

Item Reduction for Scale: The model currently comprises 56 items, which may become expanded with more factors. Educators and researchers should consider reducing items for each indicator to make it more manageable.

For Future Research:

Integration of Educational Pillars: Given that the study only addresses two of the four educational pillars, future research should include all four pillars to provide a holistic view of student learning.

Impact Studies: Future research should also investigate the effects that influence students' learning indices to provide a more comprehensive analysis.

By focusing on these recommendations, educators and stakeholders can enhance the utility and robustness of the learning measurement model and guidelines.

ACKNOWLEDGEMENTS

This research was supported by the 90th Anniversary of Chulalongkorn University, Rachadapisek Sompote Fund.

REFERENCES

Allan, A., & Charles, C. (2015). Preparing for life in the global village: producing global citizen subjects in UK schools. *Research Papers in Education, 30*(1), 25-43.

Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual review of sociology*, 223-243.

Brown, A., & Johnson, B. (2021). Creating instrumental research: A comprehensive guide. *Journal of Research Methods, 25*(3), 250-267.

Brunvand, S., & Byrd, S. (2011). Using VoiceThread to promote learning engagement and success for all students. *Teaching Exceptional Children, 43*(4), 28-37.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365-376.

Chalofsky, N., & Krishna, V. (2009). Meaningfulness, commitment, and engagement: The intersection of a deeper level of intrinsic motivation. *Advances in Developing Human Resources, 11*(2), 189-203.

Chalmers, D. (2012). *Towards a framework of teaching and learning indicators in Australian universities.* Paper presented at the Proceedings of The Australian Conference on Science and Mathematics Education (formerly UniServe Science Conference)

Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2007). An introduction to 'benefit of the doubt'composite indicators. *Social Indicators Research, 82*(1), 111-145.

Darling-Hammond, L. (2008). Teacher learning that supports student learning. *Teaching for intelligence, 2*, 91-100.

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic bulletin & review, 20*(4), 631-642.

Delors, J. (1996). Learning: the treasure within. Repor of the commission on Education for the 21st Century: Paris: Unesco.

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and perspectives*, *38*(1), 105-123.

Foa, R., & Tanner, J. (2012). *Methodology of the indices of social development*. Retrieved from

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, *18*(1), 39-50.

Giambona, F., & Vassallo, E. (2014). Composite Indicator of Social Inclusion for European

Countries. *Social Indicators Research, 116*(1), 269-293. doi:10.1007/s11205-013-0274-2.

Giovannini, E., Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., & Hoffman, A. (2008). *Handbook on constructing composite indicators: methodology and user guide*. Retrieved from

Gliem, R. R., & Gliem, J. A. (2003). *Calculating, interpreting, and reporting Cronbach's alph a reliability coefficient for Likert-type scales*.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados*. Bookman editora.

Hair Jr, J., Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Hartman, E. (2015). Global Citizenship Offers Better Solutions. *International Educator, 24*(3), 74.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of educational research, 66*(2), 99-136.

Hoskins, B., & Mascherini, M. (2009). Measuring Active Citizenship through the Development of a Composite Indicator. *Social Indicators Research, 90*(3), 459-488. doi:10.1007/s11205-008-9271-2

Hoyle, R. H. (Ed.). (1999). *Statistical strategies for small sample research*. sage.

Jones, K. (2018). Balancing quantitative and qualitative measures in instrumental research. *Journal of Balanced Research, 15*(4), 300-315.

Manthalu, G., Nkhoma, D., & Kuyeli, S. (2010). Simple versus composite indicators of socioeconomic status in resource allocation formulae: the case of the district resource allocation formula in Malawi. *BMC Health Services Research, 10*, 1-8. doi:10.1186/1472-6963-10-6

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*, *63*(7), 737-745.

Nan-Zhao, Z. (2005). Four 'Pillars of Learning' for the Reorientation and Reorganization of Curriculum: Reflections and

Discussions. *International Bureau of Education-UNESCO*.

Nardo, M., & Saisana, M. (2008). OECD/JRC handbook on constructing composite indicators. Putting theory into practice: OECD Publishing.

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). Tools for composite indicators building. European Commission Joint Research Centre. *Institute for the Protection and Security of the Citizen Econometrics and Statistical Support to Antifraud Unit, European Communities*.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of educational research, 81*(3), 376-407.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational researcher, 38*(2), 109-119.

Polit, D. F. (2015). Assessing measurement in health: Beyond reliability and validity. *International journal of nursing studies*, *52*(11), 1746-1753.

Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in nursing & health*, *29*(5), 489-497.

Resource, A., & Guide, P. (2008). 21st Century Skills, Education & Competitiveness.

Royal Government of Cambodia. (2015). *National Education For All Committee: The National Education For All 2015 Review Report*. Phnom Penh, Cambodia.

Saisana, M. (2008). The 2007 Composite Learning Index: Robustness Issues and Critical Assessment. *European Commission, JRC-IPSC, Italy*.

Scatliff, A. K., & Meier, A. (2012). Holistic learner engagement for success in the innovation age: portfolio, strengths-based, and collaborative learning strategies. *Cutting-edge technologies in higher education, 6*, 17-55.

Smith, K. A., Sheppard, S. D., Johnson, D. W., & Johnson, R. T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education, 94*(1), 87-101.

Smith, J., Davis, M., & Kim, Y. (2020). The importance of psychometric property validation in educational research. *Journal of Educational Psychology, 112*(1), 20-35.

Tawil, S., & Cougoureux, M. (2013). *Revisiting learning: The treasure within–N 4–assessing the impact of the 1996 'Delors Report'*. UNESCO.

UNESCO. (1996). *Learning: The treasure within. Report to UNESCO of the International Commission on Education for the Twenty-first Century*.

UNESCO. (2014). *Global Citizenship Education: Preparing learners for the challenges of the twenty-first century*. place de Fontenoy, 75352 Paris 07 SP, France.

Williams, T. (2019). Quantitative and qualitative methods in instrumental research. *International Journal of Research Methodologies, 48*(2), 123-139.

Wirth, K. R., & Perkins, D. (2008). Learning to learn. *Recuperado de http://www.macalester.edu/geology/wirth/CourseMaterials. html*.

Zhou, P., & Ang, B. (2009). Comparing MCDA Aggregation Methods in Constructing Composite Indicators Using the Shannon-Spearman Measure. *Social Indicators Research, 94*(1), 83-96. doi:10.1007/s11205-008-9338-0