

# DEMOGRAPHIC CONSIDERATIONS IN INSTRUMENT VALIDATION: ENHANCING RELIABILITY AND VALIDITY IN THE CONTEXT OF BTLED STUDENTS

Joan Grace Q. Duero\*, and Sarah O. Namoco

University of Science and Technology of Southern Philippines  
C.M. Recto Ave., Lapasan, Cagayan de Oro City 9000, Misamis Oriental, Philippines

\*Corresponding Author: [joan.duero@ustp.edu.ph](mailto:joan.duero@ustp.edu.ph)

**ABSTRACT.** *This study focuses on evaluating the reliability and validity of an adapted instrument designed to assess Learning and Innovation Skills (LIS) and Social-Emotional Skills (SEL) among Bachelor of Technology and Livelihood Education (BTLED) students at a State University and Colleges (SUC) in Northern Mindanao, Philippines. The instrument underwent rigorous testing to ensure its effectiveness across diverse demographic groups, considering factors like age, gender, and academic specialization that may influence measurement properties. The research objectives include examining the reliability and validity of the instrument specifically tailored for BTLED students. The study sampled 71 third-year and fourth-year students who previously completed the course EDU 215 during their second year, focusing on their resilience and adaptability amidst challenges brought by the COVID-19 pandemic. The majority of respondents were female students specializing in Home Economics. Data analysis confirmed acceptable internal consistency reliability for most constructs. Validity assessments, including Heterotrait-Monotrait (HTMT) analyses, demonstrated the instrument's effectiveness in distinguishing between LIS and SEL constructs. Recommendations for enhancing instrument reliability and validity include diversifying the participant pool to address demographic biases and conducting additional validation studies across different academic disciplines.*

**Keywords:** BTLED students, Instrument Validation, Learning and Innovation Skills, Social-Emotional Skills, SUC

## INTRODUCTION

Key factors indicating the quality of a measuring instrument include the reliability and validity of its measures. The process of developing and validating an instrument primarily aims to minimize errors in the measurement process. [1]. Reliability evaluations assess the stability of measures, internal consistency of measurement instruments, and agreement among raters scoring the instrument [2]. Validity refers to the extent to which the interpretations of test results are justified, depending on the intended use of the test [3].

This study presents the reliability and validity of the adapted instrument for assessing Learning and Innovation Skills (LIS) and Social-Emotional Skills (SEL) in the context of Bachelor of Technology and Livelihood Education (BTLED) students who are enrolled in a Science and Technology University in Northern Mindanao, Philippines. While the instrument utilized in this study has been adapted and was subjected to reliability and validity testing [4, 5], it remains imperative to assess its validity and reliability across diverse respondents. These differences encompass various dimensions such as age, gender, educational level, cultural background and socio-economic status carry the potential to exert significant influence on the validity and reliability of instruments initially tailored for distinct demographic respondents [4]. While an instrument may exhibit robust psychometric properties within its original population, the transferability of these properties to new demographic contexts demands a meticulous examination of its consistency in performance across various demographic groups [5].

The respondents of this study were the 3rd and 4th year students who already took up the course, *EDU215: Facilitating Learner-Centered Teaching: The Learner-Centered Approaches with Emphasis on Trainer's Methodology 1* during their 2<sup>nd</sup> year level. They were characterized by unexpected challenges and adaptations in education brought by the COVID-19 pandemic. These students demonstrated resilience, adaptability, and perseverance in navigating these raging times while grappling with the social, emotional, and academic impacts of the

global health crisis. Numerous studies have documented the effect of this event on their emotion and academic performance [6, 7, 8]. Additionally, these students were default beneficiaries of the UniFAST scholarship program through RA No. 10687, to provide special education assistance to tertiary students, especially the poor but academically able to successfully pursue and complete tertiary education programs quality institutions, hence exempting them from tuition fees. Thus, the majority of the respondents come from socio-economic backgrounds close to or below the poverty line. Literatures reported about the significant contribution of SES of students towards academic performance and behavior in school [9, 10]. This means that individuals from diverse demographic backgrounds might interpret and respond to instrument items differently, posing a potential risk to the instrument's measurement properties.

Assessing the instrument's validity and reliability serve as invaluable tools for identifying biases or limitations inherent in the instrument's application across diverse demographic groups. This ensures the instrument's validity and reliability despite demographic heterogeneity. Therefore, the effort to address demographic differences emerges as an essential aspect in ensuring the integrity and generalizability of research outcomes across varied population groups. By recognizing and justifying the impact of demographic disparities on measurement instruments, researchers can enhance the robustness and applicability of the research findings.

The objective of this study is to describe the validity and reliability of the instrument when administered to BTLED students in order to provide a more robust and reliable data analysis and interpretation. Specifically, it pursued the following research question: "What is the reliability and validity of LIS and SEL instrument when administered to BTLED students?"

## REVIEW OF RELATED LITERATURE

Measurement instruments are fundamental tools in research, enabling the collection of accurate and meaningful data for

various scientific inquiries. The quality of these instruments is paramount, with reliability and validity serving as key indicators of their effectiveness. [11, 12]. However, the utilization of measurement instruments across diverse populations poses a significant challenge due to demographic differences that can influence the validity and reliability of these instruments [13].

### Demographic Differences and Instrument Validation

Demographic factors such as age, gender, education level, as well as cultural background, and socio-economic status, can exert significant influence on the interpretation and response to instrument items, potentially compromising their measurement properties [14]. For instance, cultural and linguistic differences may impact the understanding and interpretation of test items, leading to variations in responses across demographic groups [15]. Cultural values and beliefs shape individuals' understanding and expression of emotions, as well as their interpersonal relationships [18, 19]. These cultural values influence how individuals perceive and manage their emotions within social contexts. Additionally, differences in cognitive abilities or life experiences among demographic subgroups may affect the internal consistency and agreement among raters scoring the instrument [16]

To address the challenge posed by demographic differences, this study undertakes comprehensive assessments to ensure the validity and reliability of measurement instruments across the BTLED students. A methodological technique employed in this study may help uncover biases or limitations integral to the instrument's application across different demographic strata [17]. These analytical approaches help safeguard the instrument's validity and reliability amidst demographic heterogeneity, ensuring the integrity and generalizability of research outcomes [17].

## METHODOLOGY

### Research Design

A descriptive research design was employed to evaluate the validity and reliability of the instrument for LIS and SEL for BTLED students. Descriptive research design is defined as the collection of quantitative data to describe the characteristics of the population or phenomenon being studied [22]. More so, it does not investigate causality nor relationships but rather reveals distinctive characteristics of a specific population.

In this study, the objective is to describe the validity and reliability of the instrument of LIS and SEL when used to BTLED students. Therefore, descriptive design is appropriate, as it utilizes numerical data to describe the validity and reliability of the instrument as implemented to BTLED students specifically taking the course, EDUC 215.

### Respondents of the Study

In this study, the respondents were drawn from third and fourth-year students enrolled in the Bachelor of Technology and Livelihood Education (BTLED) program who had previously completed the course EDU215: Facilitating Learner-Centered Teaching: The Learner-Centered Approaches with Emphasis on Trainer's Methodology 1

during their second year. These students were selected deliberately due to their familiarity with the course content, having already completed it. As a result, they were considered the most suitable respondents to provide relevant insights aligned with the study's objectives. A total of 71 respondents actively and comprehensively engaged in the survey.

**Table 1 Demographic Profile of the Respondents (n=71)**

Profile		Frequency	Percentage
Age	20 and below	10	14.1
	21-22	45	63.3
	23-24	12	17
	25-26	2	2.8
	27-28	1	1.4
	29 and above	1	1.4
Sex	Female	64	90.1
	Male	6	8.5
	Prefer not to say	1	1.4
Specialization	HE	70	98.6
	IA	1	1.4
Year Level	3 <sup>rd</sup> Year	53	74.6
	4 <sup>th</sup> Year	18	25.4

Table 1 presents the demographic profile of the respondents. It reveals certain trends and characteristics of the participants, including their age distribution, gender representation, academic specialization, and academic year level.

### Research Instrument

The instrument used in the study was composed of three sections. The first section covered the respondent's demographic profile. The second section of the instrument which was adapted from the study of Kelley, Knowles, Han, & Sung [18], would ask the respondents to provide their responses regarding LIS. In this section, the respondents answered 30 questions across the four constructs of LIS such as collaboration, communication, critical thinking, and creativity. The third section, adapted from the work of Zhou and Ee [19], asked the respondents' responses on SEL skills. In this section, the respondents answered 25-item statements with five constructs such as self-awareness, social awareness, self-management, relationship management, and responsible decision-making. Each item is measured using a 6-point Likert scale, with 5 described as Very true and 1 as not true at all [20].

### Reliability and Validity of the Instrument

The reliability and validity of the research instrument used in the study were measured using the Convergent Validity and Discriminant Validity (HTMT) through the SmartPLS 4 version. Table 2. presents the Convergent Validity of the instrument including Factor Loading, Composite Reliability, and AVE. The measure of internal consistency will determine the reliability of the instrument which is found in Cronbach's alpha, composite reliability [21]. These measures indicate the extent to which the items within each construct are correlated and consistently measure the same underlying construct [11]. Cronbach's alpha value of 0.7 or higher is considered acceptable [21]. As presented in Table 2, the majority of constructs demonstrate acceptable levels of internal consistency reliability, as indicated by Cronbach's.

**Table 2 Convergent Validity (Factor Loading, Composite Reliability, and AVE)**

Construct	Items	Loading	Cronbach's alpha	Composite reliability	Average variance extracted (AVE)
<i>Learning and Innovation Skills</i>					
<b>Critical Thinking</b>			0.929	0.940	0.589
	CC1	0.559			
	CC2	0.782			
	CC3	0.724			
	CC4	0.855			
	CC5	0.862			
	CC6	0.827			
	CC7	0.777			
	CC8	0.748			
	CC9	0.772			
	CC10	0.685			
	CC11	0.805			
<b>Collaboration</b>			0.900	0.919	0.564
	COLL1	0.790			
	COLL2	0.677			
	COLL3	0.810			
	COLL4	0.678			
	COLL5	0.871			
	COLL6	0.746			
	COLL7	0.735			
	COLL8	0.530			
	COLL9	0.860			
<b>Communication</b>			0.904	0.928	0.721
	COMM1	0.828			
	COMM2	0.792			
	COMM3	0.889			
	COMM4	0.871			
	COMM5	0.863			
<b>Creativity</b>			0.872	0.907	0.661
	CRE1	0.830			
	CRE2	0.790			
	CRE3	0.807			
	CRE4	0.854			
	CRE5	0.783			
<i>Social-Emotional Skills</i>					
<b>Responsible Decision-making</b>			0.850	0.893	0.626
	RDM1	0.824			
	RDM2	0.809			
	RDM3	0.744			
	RDM4	0.859			
	RDM5	0.712			
<b>Relationship Management</b>			0.747	0.833	0.560
	RM1	0.830			
	RM2	0.851			
	RM3	0.588			
	RM5	0.694			
<b>Self-Awareness</b>			0.824	0.874	0.640
	SA1	0.930			
	SA2	0.890			
	SA3	0.635			
	SA4	0.708			
<b>Self-Management</b>			0.871	0.861	0.554
	SM1	0.730			
	SM2	0.736			
	SM3	0.661			
	SM4	0.767			
	SM5	0.819			
<b>Social Awareness</b>			0.848	0.877	0.590
	SOA1	0.863			
	SOA2	0.763			
	SOA3	0.660			
	SOA4	0.750			
	SOA5	0.790			

*Note: RM4 and SA5 were removed due to poor loading*

**Table 3 Discriminant Validity (HTMT)**

	CC	COLL	COMM	CRE	RDM	RM	SA	SM	SOA
CC									
COLL	0.510								
COMM	0.762	0.615							
CRE	0.821	0.7350	0.775						
RDM	0.632	0.5840	0.535	0.555					
RM	0.375	0.5900	0.249	0.445	0.632				
SA	0.385	0.4730	0.214	0.447	0.452	0.463			
SM	0.268	0.2190	0.331	0.272	0.435	0.392	0.247		
SOA	0.238	0.3250	0.205	0.284	0.563	0.404	0.308	0.434	

## ANALYSIS AND DISCUSSION

### Characteristics of the Respondents

The majority of respondents (63.3%) fall within the age range of 21-22 years. This indicates a significant proportion of young adults in the study. In the Philippines, this age bracket corresponds to the expected ages for students at a certain year level of education following the K-12 Program key stages [19]. The majority of females suggest that this population specializes in Home Economics, which happens to be a female-dominated specialization. The majority of respondents (74.6%) are in their 3rd year of study, signifying a larger representation of students at 3<sup>rd</sup> year academic level compared to the 4<sup>th</sup> year. The demographic profile reveals certain trends and characteristics of the participants, including their age distribution, gender representation, academic specialization, and academic year level. The findings suggest potential implications for the generalizability of study results, as the sample may not fully represent the broader population due to certain demographic biases, such as gender imbalance and academic specialization [20].

### Reliability of Instrument

The majority of constructs demonstrate acceptable levels of internal consistency reliability, as indicated by Cronbach's alpha values exceeding 0.7. This suggests that the items within each construct are highly correlated and consistently measure the same underlying construct, enhancing the reliability of the instrument's measurements [21]. In addition, it is also noted that items RM4 and SA5 were deleted from the instrument due to poor loading. In the instrument, RM4 is described as "I can read people's faces when they are angry", while SA5 is described as "I am tolerant of my friend's mistakes". A poor loading item signifies a low correlation or association with the latent variable [26]. This means that the item's variance is not well explained by the factor it is intended to measure. Overall, the instrument shows capable reliability for assessing both LIS and SEL among BTLED students.

### Validity of the Instrument

The HTMT analysis provides insight into the discriminant validity of the instrument. It demonstrates that the constructs it measures are sufficiently distinct from each other [18]. By confirming discriminant validity, the analysis supports that the instrument effectively distinguishes between these constructs, enhancing the credibility and accuracy of the measurements obtained [2]. Moreover, the indication that the instrument accurately measures the intended constructs suggests that it is a reliable tool for assessing respondents' LIS and SEL. The effectiveness of the instrument in delivering accurate insights into respondents' skills and competencies enhances its practicality and value for

evaluating and meeting the educational and developmental requirements of BTLED students.

### Ethical Considerations

In the conduct of this study, the researcher ensures ethical standards are met. The researcher obtained voluntary consent from the respondents before their involvement in the study, and the goal and purpose of the research were explained to them clearly [27]. It also seeks permission from the immediate head of the institution's department where the study was conducted. Furthermore, it also ensures to maintain the confidentiality of the respondent's information by securing and anonymizing it. All the study materials, such as survey questionnaires and consent letters, were accurately constructed and reviewed to avoid any offensive or discriminatory language.

## CONCLUSIONS

In conclusion, the analysis of respondent characteristics discloses notable trends, including a predominant presence of young adult females specializing in Home Economics, predominantly in the third year of study. These demographics may influence the generalizability of study findings due to gender imbalance and academic focus biases. Regarding the reliability of the instrument, all constructs exhibit acceptable internal consistency the instrument, demonstrating reliable capabilities in assessing LIS and SEL among BTLED students. Additionally, the instrument's confirmed discriminant validity enhances its credibility which confirms its effectiveness in accurately evaluating BTLED students' skills and competencies for educational and developmental purposes within this specific academic cohort.

## RECOMMENDATIONS

Based on the result and discussion, the following recommendations may help enhance the instrument reliability and validity used in the study. First, it is recommended to address the significant gender inequality and academic specialization observed in the sample, it is advisable to diversify the participant pool by recruiting more male participants and individuals from diverse academic specializations. This approach can improve the generalizability of study findings. Second, conducting additional validation studies across different cohorts or academic disciplines can help establish the reliability and validity of the instrument beyond the current study context, confirming its effectiveness in assessing LIS and SEL among BTLED students.

## LIMITATIONS OF THE STUDY

The potential lack of diversity within the respondent's sample

limits the study. The majority of respondents being female and specializing in Home Economics may limit the generalizability of the findings to other demographic groups or academic disciplines. This narrow focus could impact the broader applicability of the study's conclusions and recommendations, particularly in contexts with more diverse student populations or different academic backgrounds.

## REFERENCES

- [1] E. Almasreh, R. Moles And T. F. Chen, "Evaluation Of Methods Used For Estimating Content Validity," Vol. 15, No. 2, Pp. 214-221, 2019.
- [2] S. El-Den, C. R. Schneider, A. Mirzaei And S. R. Carter, "How To Measure A Latent Construct: Psychometric Principles For The Development And Validation Of Measurement Instruments," International Journal Of Pharmacy Practice, Pp. 326-336, 2020.
- [3] L. B. Mokkink, C. B. Terwee, D. L. Knol, P. W. Stratford, J. Alonso, D. L. Patrick, L. M. Bouter And H. C. De Vet , "The COSMIN Checklist For Evaluating The Methodological Quality Of Studies On Measurement Properties: A Clarification Of Its Content," BMC Medical Research Methodology, Vol. 10, No. 22, 2010.
- [4] S. M. Y. Arafat, H. Z. R. Chowdhury, M. M. A. S. Qusar And M. A. Hafez, "Cross-Cultural Adaptation And Psychometric Validation Of Research Instruments: A Methodological Review," Journal Of Behavioral Health, Vol. 5, No. 3, Pp. 129-136, 2016.
- [5] W. Beyers And L. Goossens, "Concurrent And Predictive Validity Of The Student Adaptation To College Questionnaire In A Sample Of European Freshman Students," Educational And Psychological Measurement, Vol. 62, No. 3, Pp. 527-538, 2002.
- [6] L. Quintiliani, A. Sisto, F. Vicinanza, G. Curcio And V. Tambone, "Resilience And Psychological Impact On Italian University Students During COVID-19 Pandemic. Distance Learning And Health," Psychology, Health & Medicine, Vol. 27, No. 1, Pp. 69-80, 2021.
- [7] Q. Zeng, Z. Liang, M. Zhang, Y. Xia, J. Li, D. Kang, D. Yi And J. Wang, "Impact Of Academic Support On Anxiety And Depression Of Chinese Graduate Students During The COVID-19 Pandemic: Mediating Role Of Academic Performance," Psychology Research And Behavior Management, Vol. 14, Pp. 2209-2219, 2021.
- [8] H. Song, Y.-F. Mu, C. Wang, J. Cai, Z. Deng, A.-P. Deng, X. Huang, X. Meng, L. Zhang, Y. Huang, W. Zhang, W. Shen, J. Chen, B. Liu, R. Gao, J. Zhao And M. Ran, "Academic Performance And Mental Health Among Chinese Middle And High School Students After The Lifting Of COVID-19 Restrictions," Frontiers In Psychiatry, 2023.
- [9] M. Abduh, E. Purwanta And Hermanto, "In What Ways Students' Socio-Economic Status Affecting Academic Performance?," International Journal Of Evaluation And Research In Education (IJERE), Vol. 12, No. 1, Pp. 34-43, March 2023.
- [10] M. Kaya, "The Relationship Among Socioeconomic Status, Attitude Towards Science, School Climate And Students' Science Achievement: A Cross-Country Comparison Of TIMSS: A Cross-Country Comparison Of TIMSS 2019," Science Education International, Vol. 33, No. 4, Pp. 366-375, 2022.
- [11] U. R. Sudaryono , Q. Aini, Y. I. Graha And L. Ninda , "Validity Of Test Instruments," In 2018 1st Workshop On Engineering, Education, Applied Sciences, And Technology, 2019.
- [12] K. F. Cook, M. A. Kallen And D. Amtmann, "Having A Fit: Impact Of Number Of Items And Distribution Of Data On Traditional Criteria For Assessing IRT's Unidimensionality Assumption," Quality Of Life Research, Pp. 447-460, 2009.
- [13] A. C. Pickett , D. Valdez And A. E. Barry, "Does It Work For E Ork For Everyone? The Influence Of Demogr One? The Influence Of Demographic Variables On Statistical Reliability," Health Behaviour Research, Vol. 3, No. 2, 2020.
- [14] R. K. Hambleton And A. L. Zenisky, , "Translating And Validating Test Scores For Linguistic And Cultural Equivalence," In Handbook Of Test Development, Pp. 277-306, 2011.
- [15] B. Muthén And . T. Asparouhov, "Latent Variable Analysis With Categorical Outcomes: Multiple-Group And Growth Modeling In Mplus," Mplus Web Notes, 2002.
- [16] S. G. Shealy And W. Stout, "). A Model-Based Standardization Approach That Separates True Bias/DIF From Group Ability Differences And Detects Test Bias/DTF As Well As Item Bias/DIF," Psychometrika, Vol. 58, No. 2, Pp. 159-194, 1993.
- [17] F. J. Van De Vijver And K. Leung, "Methods And Data Analysis For Cross-Cultural Research," 1997.
- [18] T. R. Kelley, J. G. Knowles, J. Han And E. Sung, "Creating A 21st Century Skills Survey Instrument For High School Students," American Journal Of Educational Research, Vol. 7, No. 8, Pp. 583-590, 22 August 2019.
- [19] M. Zhou And J. Ee, "Development And Validation Of Social Emotional Competency Questionnaire," International Journal Of Emotional Education, Vol. 4, No. 2, 23 November 2012.
- [20] E. Roemer, F. Schubert And J. Henseler, "HTMT2—An Improved Criterion For Assessing Discriminant Validity In Structural Equation Modeling," Industrial Management & Data Systems, Vol. 121, No. 12, Pp. 2637-2650, 10 November 2021.
- [21] "Schooling And Education," 2024. [Online]. Available: <https://www.angloinfo.com/how-to/philippines/family/schooling-education>.
- [22] S. J. Kamper, "Sampling: Linking Evidence To Practice," Journal Of Orthopaedic & Sports Physical Therapy, Vol. 50, No. 12, Pp. 725-726, December 2020.
- [23] O. Robertson And S. Evans, "Just How Reliable Is Your Internal Reliability? An Overview Of Cronbach's Alpha (A)," Psychology Postgraduate Affairs Group Quarterly, No. 115, Pp. 23-27, June 2020.