

Construct-Related Validity Vis-A-Vis Internal Structure of the Test

Rufina C. Rosaroso (PhD)¹, Enriqueta D. Reston (PhD)², Nelson A. Rosaroso (PhD)³

¹Cebu Normal University, ^{2,3}University of San Carlos, Cebu City, Philippines
*raffycoronel@yahoo.com*¹, *edreston@yahoo.com*², *nelson_rosaroso@yahoo.com*³

Date Received: August 6, 2015; Date Revised: September 16, 2015

Abstract –*This study aimed to establish the content, construct and predictive validity evidence of a locally-made college admission and placement test for a private sectarian university. It employed empirical methods for test validation and reliability analysis using a sample of 262 college freshman students. Content-related validity evidence was initially established through the development of Tables of Specifications for the various subtests and experts' review on the alignment of test item content and objectives. Construct-related evidence of validity was established by exploring the internal structure of the test via factor analysis. The procedure yielded two factors and examination of the items comprising the subtests that loaded heavily on each factor suggested that the test measured two distinct but interrelated skills; namely: (1) Language and Literacy Skills and (2) Analytical Reasoning Skills, across three content areas in English, Mathematics and Science. Results from correlating test scores and first year cumulative Grade Point Average (GPA) of the students yielded a substantial overall predictive validity coefficient of 0.643. Reliability analysis using internal consistency measures yielded a high reliability coefficient Cronbach Alpha value of 0.968. These findings indicated that the proposed college admission test was a good predictor of student's academic performance in the first year of college work, and thus, valid for its intended purpose.*

Keywords – *content-related validity, construct-related validity, factor analysis, predictive validity, college admission test*

INTRODUCTION

College admission tests have served various purposes in higher education. Performance in these tests is one major criterion for consideration of an applicant for admission to college academic programs or grant to competitive scholarships in higher education institutions (HEIs). College admission test scores are used to appraise and make inferences on an applicant's potential for success in college work. They

also serve as basis for counseling and placing students in “appropriate” college programs. In the Philippines, most colleges and universities require applicants to take a college admission test as part of the admission requirement process. They have been used by colleges for many years to predict the success of students in college. However, locally developed admission tests need to undergo test validation process to justify the inferences and decisions made based on test scores.

The quality of a test is dependent on its validity. According to Oosterhof [1] validity is the most important criterion in scale development and interpretation, and the property of the test to which it is capable of achieving certain aims [2]. Validity is significant in test evaluation and is concerned whether the test measures what it is supposed to measure.

The nature of validity may be based from both classical and modern unified views as presented by various authors. The classical view of validity considers validity as an aspect or property of a test or any measuring instrument. Nunnally [3] described classical view of validity in terms of its types as content, predictive and construct. However, Cronbach and Meehl [4] added concurrent validity as another type of validity evidence. In contrast with the classical view of validity, the modern view of validity is described as “the accuracy of students' scores” [5] and “what the test measures and what can be inferred from the test scores” [6]. Thus, in accordance with the modern unified view of validity, the different types of validity in the classical approach may be viewed as various sources of validity evidence. These validity evidences are classified as content-related; criterion-related which may be predictive and concurrent; and construct-related evidences.

Content validity is the adequacy and presentation of a specific topic in relation to test items. Cronbach and Meehl [4] contended that content-related evidence of validity is established by identifying the test coverage and the appropriate competencies measured in the test. To ensure maximum content validity of the test, a Table of Specifications is required to identify the covered content areas and the extent to which the

test items represent the content coverage and the test objectives which specify intended competencies to be tested.

Criterion-related validity evidence, on the other hand, is classified into two types: concurrent and predictive. Kerlinger and Lee [7] contended that concurrent validity is used to validate a new test. It is computed by correlating the two sets of scores, one is the new test scores and the other is the existing test scores. Crocker and Algina [8] defined predictive validity as the degree to which test scores predict an observable criterion in the future or the extent on how a person's current performance is measured with a certain criterion at a later time [9]. Moreover, Crocker and Algina [8] claimed that predictive validity is used as basis for justifying students' future performance in their chosen degree programs. Anastasi [10] emphasized that concurrent validity is used in diagnosing students' present academic performance. The diagnosis indicates whether the student is a fast learner and/or advised for remediation.

Another type of validity evidence which is essential in aptitude, achievement and interest tests is construct-related evidence. According to Kerlinger and Lee [7], a *construct* is a characteristic presumed to exist but cannot be directly measured. Thus, construct validation involves test interpretation where factor analysis is used [4].

Kerlinger and Lee [7] distinguished the relationships among the different types of validity. There are instances when instruments intend to measure constructs used as specific predictors, as in intelligence tests. Construct validity is performed to predict success in college and in various occupations. Content validity is supportive of construct validity, in a way that the procedures required to ensure content validity are related in defining the domains identified and observed in construct validity. Messick [11] further argued that construct validity is the overarching category which encompasses all other types of validity evidence.

At the University of San Carlos (USC), a standardized intelligence test (commonly referred to as IQ Test) and a locally-made English Language Proficiency Test (EPT) have been used over the years as bases for admission of applicants to college academic programs. While different colleges have adopted different cut-off scores for IQ and EPT, performance on these tests served as a major basis for consideration of an applicant for admission in a degree program. While standardized IQ tests are

imported and there are rarely local studies that investigate their freedom from cultural bias when used within the local context for college admissions, the locally-made test, EPT, has not received any systematic investigation in terms of its validity and relevance over the past several years that it has been used by the university. It is in this context that the researchers undertook a review of the existing admission tests as basis for establishing the need for developing and validating a multidimensional research-based college admission and placement test. Within the framework of Classical Test Theory as basis for test item development and validation, the project developed and validated a college admission and placement test for USC. Along this line, this paper described how content-related evidence, construct validity evidence based on internal structure, and test criterion relationships based on predictive validity evidence support the appropriateness of decisions made on students using the proposed college admission and placement test.

OBJECTIVES OF THE STUDY

This study aimed to gather and establish various forms of validity evidence on a localized admission and placement test that may serve as basis for admission and/or placement of applicants to college academic programs in the University of San Carlos. In particular, this paper sought to determine the extent of the admission test scores reliable as bases for making decisions, admission and placement in college; to determine the extent of the following sources of validity evidence which support the validity of the college admission and placement test for assessing students' potential to do college work in an academic degree program: content-related evidence, construct-related evidence based on internal structure of the test and criterion-related evidence based on external criterion of future performance (predictive validity).

MATERIALS AND METHODS

The project employed test development and validation methods in order to develop and validate an admission and placement test for college applicants to USC. The procedures were undertaken in four phases; namely: (1) Evaluation of the current admission tests, test planning and item development; (2) Administration of a first pilot test to a sample of 4th year high school students and item analysis of results; (3) Administration of a second pilot tests to a sample of college freshmen at USC for test validation and

reliability estimation; (4) Development of the final form of the test and procedures for administration and scoring. This paper focuses on the results of Phase 3 which deals with the procedures for establishing test validity and reliability.

The new college admission and placement test was pilot tested to 262 college freshmen representing 7 colleges of the University of San Carlos. Cluster random sampling by class sections was used in the selection of college freshmen who participated in the pilot test. One section of a college freshman class was randomly selected to represent each College for the pilot test administration. The class size varied from 35-45 students and the number of students who participated in the pilot test for the 7 selected sections summed up to 262.

The new college admission test served as instrument of the study. Reliability analysis was performed using Cronbach's alpha coefficient as a measure of internal consistency of the items in the test and its subtests. Content-related validity evidence was established through alignment of test item content and objectives based on tables of specifications and experts' review. Construct-related validity evidence was established using the methods of factor analysis to identify the number and describe the nature of the underlying factors which the items in the subtests measure. The Kaiser-Guttman rule on retaining factors with eigenvalue greater than 1 was used as the criterion for dimensionality or determination of the number of factors [12]. Further, the degree of relationship of the students' subtest scores and their Grade Point Average was computed using bivariate Pearson Correlation Coefficient to measure the predictive validity of the test and its subtests. All statistical analysis procedures were performed using the Statistical Package for the Social Sciences (SPSS) Program.

RESULTS AND DISCUSSION

This section presents the results in the test validation process of the university college admission

and placement test. Since the reliability of a test is a pre-requisite to validity, the results of reliability analysis of the test and its subtests are presented first, followed by the various forms of validity with focus on the construct-related validity as the overarching category of the various types of validity evidence.

Reliability of the Revised USCCAPT

The reliability analysis of the revised admission and placement test was done using internal consistency measures to ascertain whether the items that made up the subtests were internally consistent. The procedure for reliability analysis generated a reliability coefficient *alpha*, also called the *Cronbach's Alpha*, as the measure of homogeneity or internal consistency of items. Values of the reliability coefficient range from 0 to 1, with values closer to 1 indicating high reliability. Acceptable values for reliability range between 0.80 and above [13]. Students' responses to the four subtests were subjected to reliability analysis via internal consistency measure and the result yielded reliability coefficients for the four subtests as shown in Table 1.

From Table 1, it can be seen that there is a high degree of internal consistency among the items in three subtests; namely: the *English Proficiency Test* which has the highest reliability coefficient, followed by *Mathematics and Science Proficiency Tests*, respectively.

These three subtests met the acceptable standards for reliability coefficient values. On the other hand, the reliability *Reasoning Ability Test* falls below the acceptable standard of at least 0.80. According to Ary, Jacobs and Razavieh [14], one factor that affects the value of the reliability coefficient is the length of the test; that is, the longer the test, the greater the reliability.

Since *the Reasoning Ability Test* has only 12 items, this could be one factor that explains the low reliability of the test. Further, based on their inter-item correlation values, there are items subject for revision in all the subtests.

Table 1. Reliability Coefficients of the Subtests of the Revised USCCAPT

| Subtests/Areas of USCCAPT | No. of Items | Reliability Coefficient (Cronbach's Alpha) | Interpretation | No. of Items to be revised |
|---------------------------|--------------|-----------------------------------------------|----------------|----------------------------|
| English | 114 | 0.912 | high | 17 |
| Mathematics | 37 | 0.902 | High | 6 |
| Science | 41 | 0.872 | high | 12 |
| Reasoning | 12 | 0.666 | moderate | 3 |
| Overall Test | 204 | 0.968 | high | 38 |

Among the four subtests, *English Proficiency Test* has the most number of items to be revised, followed by *Science*. *Mathematics and Reasoning Tests* have the least number of items for revision.

Content-related Validity Evidence

In this study, content-related validity evidence of USCCAPT was initially established while the test framework was being planned through the development of Tables of Specifications to determine the relationship of the test items with the identified competencies. Further, test content was reviewed and analyzed by the entire project committee where experts' judgment/s on accuracy and expected competencies were considered. Further, items presentation to college deans and department chairs was done to elicit further comments and suggestions.

Based on the table of specifications, the entire test covered four subtests: namely: (1) *English Proficiency Test* (EPT) developed by selected faculty member/s of the Department of Languages and Literature (DOLL) which comprised five sub-areas designed to assess the applicants' ability to use the English Language Proficiency; (2) *Mathematics Proficiency Test* (MPT), developed by selected faculty members of the Department of Mathematics, designed to assess the student applicant's ability to use mathematical facts, principles and procedures in number operations and problem solving applications, as well as in drawing inferences or reasoning about relationships between variables based on algebraic expressions, trigonometric and geometric relations, or statistical data; (3) *Science Proficiency Test*, developed by selected science faculty members representing the Departments of Biology, Chemistry and Physics, which assessed student applicants' knowledge and understanding of basic concepts and processes of Science as well as the ability to use and reason with scientific information, principles and procedures to solve problems that relate science to real life; and (4) *Reasoning Test*, which includes logical reasoning and visual-spatial reasoning to measure the student applicants' ability to use words, symbols and patterns to reason out, infer or draw logical conclusions based on given information. To ensure maximum content-related validity evidence, the test table of specifications was made in such a way that the test items represented across the different content areas and the range of cognitive processes involved based on the Bloom's Taxonomy, namely; Knowledge,

Comprehension, Application, Analysis, Synthesis and Evaluation.

Construct-Related Validity and Evidence Based on Internal Structure of the Test

Construct-related validity evidence was established by analyzing the internal structure of the test using factor analysis to determine whether the items measure one, several or many constructs. In this study, construct-related evidence of validity was established based on the scores of 262 college freshmen in the 4 areas of the USCCAPT. Factor Analysis was performed to identify the number and describe the nature of the underlying variables, called factors, among its sub-tests which the items measure. Cronbach [15] defines factor analysis as a systematic method for examining the meaning of a test by studying its correlations with other variables. Further Ary, Jacobs and Razavieh [14] contend that this procedure analyzes the intercorrelations among a large number of measures to identify a smaller number of common factors. Factors are hypothetical constructs assumed to underlie different measures or tests and factor analysis is used to identify these underlying categories.

Table 2. Factor Loadings of the USCCAPT Areas and Its Sub-areas

| <i>Areas and Sub-areas</i> | <i>Factor 1</i> | <i>Factor 2</i> |
|------------------------------|-----------------|-----------------|
| English Proficiency Test | | |
| Spelling | 0.454 | 0.116 |
| Finding Error | 0.661 | 0.182 |
| Vocabulary | 0.557 | 0.195 |
| Grammar | 0.810 | 0.290 |
| Reading Comprehension | 0.577 | 0.319 |
| Reasoning Ability Test | | |
| Verbal Reasoning | 0.507 | 0.397 |
| Visual Spatial | 0.346 | 0.385 |
| Science Proficiency Test | | |
| Biology | 0.542 | 0.411 |
| Chemistry | 0.432 | 0.574 |
| Physics | 0.460 | 0.497 |
| Mathematics Proficiency Test | | |
| Arithmetic and Measurement | 0.463 | 0.593 |
| Algebra and Trigonometry | 0.240 | 0.841 |
| Geometry | 0.067 | 0.451 |

*Based on the 0.300 the minimum acceptable value for significant factor loadings

Table 2 shows the factor loadings obtained from the rotated component matrix obtained as output of Factor Analysis. The results revealed that two factors

were extracted using Principal Axis Factoring and the solution was rotated using Varimax Method for interpretability. The rotated factor matrix, the main output of factor analysis, yielded factor loadings or structure coefficients which measured the degree of relationship between the item response score (observed variables) and the derived factors (the latent variables or underlying dimensions) in the sub-tests.

The results in Table 1 provide rich information on the internal structure of the test. It shows that two factors were extracted and the values of the factor loadings for each of the different areas and their subtests. These factor loadings, with values from -1.00 through 0 to 1.00, indicate the degree of correlation between the sub-tests and the factor. As such, they are interpreted like correlation coefficients that measure the degree of correlation between the knowledge/skills specified in the sub-tests and the underlying dimensions or factors that these sub-tests measure. Kerlinger and Lee [7] contend that if two or more sub-tests are substantially correlated, they bear a common variance, and thus, measuring something in common. The minimum acceptable value for factor loadings is set at 0.30 in the literature, although some authors also suggest that they should be 0.40 or higher [12].

The results of factor analysis identified two factors measuring common dimension or trait that items in the four areas of USCCAPT and some of its sub-tests have in common. For the *English Proficiency Test*, all its 5 sub-tests loaded heavily on Factor 1 indicating that they share common factor variance that is, measuring one common dimension or trait. Sub-tests such as Spelling, Finding Error, Vocabulary, Grammar and Reading Comprehension measured one common trait that was identified based from factor analysis results. Moreover, Reading Comprehension has also significant factor loadings in Factor 2 which indicates that this sub-test measures to some extent, some behaviors or traits under Factor 2.

For the *Reasoning Ability* sub-tests, both sub-tests Verbal Reasoning and Visual Spatial loaded significantly in both Factors 1 and 2, that is; these sub-tests measure both factors to some extent. For the *Science Proficiency Test*, the sub-tests in Biology, Chemistry and Physics, loaded heavily in both Factors, indicating that the items measure the two dimensions or categories of behavior summed up in Factors 1 and 2. However, for the *Mathematics Proficiency Test*, all three sub-tests loaded heavily on Factor 2; moreover, Arithmetic and Measurement subtest loads significantly in Factor 1 as well.

A closer examination of the items comprising these sub-tests that loaded heavily on Factor 1 suggests that this construct may comprise *Language and Literacy Skills* which encompass those items that relate to knowledge or recall of facts, concepts, principles and procedures as well as understanding of the context and language upon which the question is asked. On the other hand, the items comprising the sub-tests that loaded heavily on Factor 2 is indicative of higher order thinking processes involving analysis, reasoning, problem-solving and other applications. The items that load heavily on this Factor require the students' ability to understand, analyze and reason with words (as in Grammar, Reading Comprehension and Verbal Reasoning), scientific or mathematical facts, principles or procedures (as in the Science and Mathematics sub-tests) or with numbers and figures (as in Arithmetic and Measurement, Visual/Spatial Reasoning, Trigonometry and Geometry). All of the items in Reasoning Ability, Science and Mathematics Proficiency Sub-tests loaded heavily in Factor 2. Further, the same results for items in Reading Comprehension sub-test for English Proficiency Test were obtained. For these reasons, Factor 2 is labelled as *Analytical Reasoning Skills*.

Criterion-Related Validity Evidence Predictive Validity

Criterion-related evidence refers to how well performance on a test correlates with performance on relevant criterion measures external to the test [1] or the effectiveness of a test in relation to performance in some criterion, some independent measure of whatever the test assesses [16]. For admission and placement tests, the external criterion measure is the test's ability to predict the applicant's future college performance in the degree program he/she is applying for. Thus, the type of criterion-related validity evidence that is relevant here is predictive validity. De la Rama [17] pointed out that predictive validity measures the extent to which the test scores can predict some subsequent criterion variable of interest and tells how closely test scores correspond with some future criterion performance.

Predictive validity was assessed using the 262 college freshmen's USCCAPT scores and their first year college cumulative Grade Point Average (GPA) which was the future criterion performance. Using the bivariate Pearson Correlation Coefficient, the degree of relationship of the sub-test scores and GPA was explored and used as a measure of predictive validity.

Table 3 shows the predictive validity coefficients of the USCCAPT and its sub-tests with respect to the cumulative freshmen GPA.

Table 3. Predictive Validity of USCCAPT and its Subtests

| <i>Test</i> | <i>Validity Coefficient</i> r^* | <i>Coefficient of Determination</i> r^{2*} | <i>Interpretation</i> |
|-------------|--------------------------------------|-------------------------------------------------|-----------------------|
| USCCAPT | 0.639 | 0.409 | substantial |
| Subtests | | | |
| English | 0.600 | 0.360 | substantial |
| Reasoning | 0.442 | 0.196 | low |
| Science | 0.548 | 0.300 | substantial |
| Mathematics | 0.497 | 0.247 | substantial |

* Sig. $p < 0.001$

As shown in Table 3, the predictive validity coefficients of USCCAPT and its subtests show a substantial degree of correlation of students' test scores with their cumulative freshmen GPA. The over-all predictive validity coefficient of 0.639 is substantial which indicates that the USCCAPT, as an admission test, is a good predictor of academic performance of students in first year college work. Validity coefficients for the subtests range from 0.442 to 0.600 which are all above the minimum acceptable value of 0.40 as indicated in the literature [1]. Moreover, the predictive validity of the *Reasoning* subtest is relatively low compared to the other subtests.

Table 3 shows the values of the coefficient of determination r^2 for the overall test and its sub-test to establish the explanatory power of the test in predicting students' achievement in their first year of college work. The coefficient of determination r^2 is defined as "the percent of the variation in one variable that is accounted for (predicted) by the other variable" [18]. The r^2 values in Table 3 when expressed in percent, shows the percentage in the variation of students' cumulative GPA that can be explained or accounted for by their test performance. For the overall test, around 40.9% of the variation in the students' GPA is accounted for by the college applicants' performance in USCCAPT. These coefficient values of r and r^2 values are significant at 0.001 level, which indicated possible generalization of these findings to other groups of college applicants to USC, and thus supports the predictive validity of the test.

These results further supported the decision to delete the entire *Reasoning Ability Test*. Besides that it yielded low reliability below the acceptable standards

and relatively low predictive validity, the results of factor analysis showed that the items in these subtests are redundant as they both share common factor variance with both Factors 1 and 2. Further, reasoning was identified as the underlying trait in the items that loaded highly in Factor 2. With these findings, the structure of USCCAPT basically included the improved items for the English, Science and Mathematics Proficiency subtests.

CONCLUSION AND RECOMMENDATION

The validation of an admission and placement test for applicants to the college degree programs of USC supports the need for more relevant and substantial basis for admission and placement decisions of college students. Findings on the validation and reliability estimation procedures within the framework and assumptions of Classical Test Theory yielded various sources of validity evidence for USCCAPT. Evidence based on content, internal structure and criterion-related evidence (predictive) were established. As an admission test, USCCAPT is found to be a good predictor of a student's academic performance in the first year of college work.

Further, construct-related validity evidence was established by analyzing the internal structure of the test using factor analysis to determine whether the subtests measure one, several or many constructs or dimensions. Construct-related evidence supported that the entire admission and placement test measured primarily two underlying traits expected of college freshmen across three content domains; namely: English, Science and Mathematics. These traits or dimensions of the test were identified based on the nature of the items as *Language and Literacy Skills* and *Analytical Reasoning*.

In conclusion, the USCCAPT passed the standards of test validation as the results of pilot test administrations yielded high reliability coefficient and substantial content-related, construct-related and predictive validity evidences. With minor revisions for improvement of some items, it is recommended that USCCAPT may be used as an admission and placement test for USC. Moreover, any test or assessment tool has its own implicit assumptions, limits of applicability and potential hazards of misinterpretation. The use of Classical Test Theory as the framework for the validation process has its inherent limitations being sample-dependent, that is, results of empirical validation and reliability analysis were highly dependent on the data obtained from the

sample of students who participated in the pilot tests. Hence, the following recommendations were put forth for future courses of action: First, a continuous review, evaluation and revision/updating of the items has to be taken a multi-disciplinary admission test committee of the university. Second, a manual of test administration and score interpretation be drafted based on the University Testing Center's decision to use either norm-referenced measures like percentile ranking or criterion-referenced interpretation based on a cut-off score. Finally, further studies may be made based on data derived from USCCAPT and other relevant educational outcomes.

REFERENCES

- [1] Oosterhof, A. (2001). *Classroom applications of educational measurement*. 3rded. Columbus, Ohio: Merrill Prentice Hall, p. 45, p.69, p. 347.
- [2] Calmorin, L. P. (1994). *Educational research, measurement and evaluation*. 2nd ed. Mandaluyong City: National Book Store, pp. 49-51.
- [3] Nunnally, J. C. (1988). *Psychometric theory*. 2nd ed. New York: McGraw Hill Book Co., p.86, p.90-91, p.98.
- [4] Cronbach, L. J. & Meehl, P. E. (1988). "Construct Validity in Psychological Tests." *Journal of Measurement*, 20 (5), 55-65.
- [5] Popham, J. W. (2000). *Modern educational measurement: Practical guidelines for educational leaders*. Boston: Allyn and Bacon, p.23.
- [6] Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. United Kingdom: Cambridge University Press, Page 102, p. 143.
- [7] Kerlinger, F. N. & Lee, H. (2000). *Foundations of behavioral science*. 4th ed. Forth Worth: Harcourt College Publishers, pp. 668-669.
- [8] Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, p. 217, pp. 224-228.
- [9] Salvia, J. & Ysseldyke, J. E. (1998). *Assessment*. Boston: Houghton Mcfflin Co., p.139, p.173.
- [10] Anastasi, A. (1988). *Psychological testing*, 6th edition, New York: Macmillan Publishing Company, ISBN No. 0-02-303020-8, p. 109
- [11] Messick, S. (1996a). Standards-based score interpretation: Establishing valid grounds for valid inferences. Proceedings of the joint conference on standard setting for large scale assessments, Sponsored by National Assessment Governing Board and The National Center for Education Statistics. Washington, DC: Government Printing Office, p. 77.
- [12] Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. (2005). *Multivariate data analysis*. 6thed. New Jersey: Prentice Hall, Inc., p.129.
- [13] Bryman, A. & Cramer, D. 1997. *Quantitative Data Analysis with SPSS for Windows: A Guide for Social Scientists*. London: Routledge., p. 63.
- [14] Ary, D., Jacobs, L.C. & Razavieh, A. (2002). *Introduction to research in education*. USA: Wadsworth/Thomson Learning, p. 260
- [15] Cronbach, L. J. (1988). "Construct Validation after Thirty Years." In R. L. Linn (Ed.), *Intelligence Measurement Theory and Public Policy*, Urbana: University of Illinois Press, pp. 362-363
- [16] Munarriz, N. J. (1994). Validity. *UPROBE*, Part 1. Manila: ALJUN Printing Press, p.54
- [17] Dela Rama, P. A. (2006). The Validity of the Admission Test of Silliman University: Proposed Qualifying Framework. Unpublished Dissertation, University of San Carlos, Cebu City, p.78
- [18] Walpole, R. E. (2000). *Introduction to statistics*. 3rd ed. Singapore: Prentice-Hall International, p. 377.