

## The Development of College Students' Conceptions of Assessment in Mathematics

Marc Edly Magat  
Mary June Jordan  
Francia Inez Sumilang  
Marlon Dominique Custodio  
Maxelle Guerrero  
Ian Steve Edrosolo  
Reginald Ivan Sison  
Jean Leslie Zerrudo

*National University  
Manila, Philippines*

The purpose of the study is to develop items on a scale for higher education students' conception of assessment in Mathematics having four domains namely: (1) assessment as improvement, (2) student accountability, (3) school accountability, and (4) irrelevance. These factors are anchored in the model of Brown (2008). The scale was administered to 300 college students, public and private school in the National Capital Region in the Philippines. Coefficient alpha for the whole scale is 0.88 and 0.89, 0.89, 0.05, and 0.94 for the factors improvement, student accountability, school accountability, and irrelevance, respectively. The scale's convergent validity was analyzed and it showed that the four factors were significantly inter-correlated. Factors had gone through factorial validity and supported a four factor model with adequate fit using Confirmatory Factor Analysis (GFI = .783, AGFI = .758, BCFI = 0.851).

**Keywords:** *Mathematics, Improvement, Accountability, Irrelevance, Conception, Assessment*

**C**onception contains the views, attitudes, or understanding on a particular phenomenon which leads to begin a plan. Understanding what conception means to students is of great importance because it explains difficult categories of experience (White, 1994). People's purposes towards occurrences are expressed in their conceptions of the phenomena (Fodor, 1998). Students' conception was studied by Brown (2008) for the purpose of analyzing the conceptions of the students when it comes to assessment. In their study, conception was framed for students' conception of assessment. Students' conceptions of assessment are significant because assessment has a substantial effect on the quality of learning (Entwistle & Entwistle, 1991; Marton & Säljö, 1997; Ramsden, 1997).

There have been quite a number of studies in developing a scale to measure student conceptions. Brown (2008) conducted a study using interactive informal assessment practices among secondary students' conceptions of assessment. The opportunity of knowing and understanding the conception of the students about assessment is important. When he asked the students about how they perceive assessment, their answer has been the basis for the construction of the items. This triggered him to do a research and it was found that there are four factors that affect the conception of the students towards assessment. These four factors include: (1) assessment improves learning, (2) assessment makes students accountable, (3) assessment is negative because it is unfair or bad, and (4) assessment is liked because it is fun or beneficial. According to Brown, student's behavior, norms beliefs etc. are substantial to distinguish and understand people's conception of assessment.

In another study, Brown (2010) looked at Queensland teachers' conceptions of assessment and its impact on policy priorities on teacher attitudes has a purpose of amplifying the teacher's conceptions about their practices of assessment. In this study it is evident that teachers' conceptions of different aspects of the education process firmly influence how they teach and what students learn or accomplish (Calderhead, 1996; Clark & Peterson, 1986; Thompson, 1992). It is specified that teachers' beliefs about students, learning, teaching, and subjects influence assessment techniques and usefulness (Asch, 1976; Cizek, Fitzgerald, Shawn, & Rachor, 1995; Kahn, 2000; Tittle, 1994). This is dependable with Ajzen's (2005) model of arranged or reasoned behavior which proposes that teachers' motives, beliefs about what others think, and sense of authority to comply with their intentions determine their behavior within the school. Teacher belief systems appear to increase from their early experiences of educational processes (Pajares, 1992).

In another study, Brown (2011) looked at teachers' conceptions of assessment in Chinese contexts texting tripartite model of accountability, improvement, and irrelevance. The educational policy reform movements attempt to reduce the negative consequences of selective educational assessments by amplifying the use of assessment to inform teaching and learning progresses through an active involvement of learners in leading learning development (i.e., assessment for learning—Assessment Reform Group, 1999).

Brown (2012) further investigated teachers' beliefs about feedback within an assessment for learning environment. Brown justified that feedback is considered to be an essential element of Assessment for Learning

Review of the empirical literature on students' conceptions about the purposes of assessment has identified four major purposes, some of which can be matched to teachers' conceptions of assessment. Students are reported as conceiving of assessment as (a) improving achievement, (b) a means for making them accountable, (c) making school accountable, and (d) being irrelevant. However, the previous studies of Brown indicates that these are only applicable for primary, high school students, and teachers' conceptions of assessment. His studies are domain-general which needs specific context and subject areas. Thus, the present study made conception on assessment more specific, focusing in Mathematics and also on the need to determine conceptions of assessment for college students.

### **Assessment improves the achievement of learning**

Brown (2010), stated that assessment is a means of improving the quality of students' learning and teachers' instruction. A variety of assessment techniques are used to progress the student's learning, and to deliver excellent instructions of the teachers. The goal is for the teachers to answer the two key questions: "Who among the students have learned?" and "Who among them needs to be taught again?". According to Peterson and Irving (2008), students conceived assessment as a method of receiving feedback to help expand their knowledge. It needs to be focused, straightforward, based on formal assessment, and provided by educators.

### **Assessment makes students accountable**

Zeidner (1987) reported that the students perceived assessment as increasing their achievements. McMillan (2001) stated that students are accountable for their learning because it is expected to them by the society; it is through their performance on examinations or tests. This requires grading, scoring, or evaluating their performance against standards, objectives, targets, or expectations. Students experience positive or negative consequences (e.g., placement into classes or groups, selection for special programs, or awarding of certificates) depending on their performances.

### **Assessment makes school accountable**

The outcomes on assessment is being practiced by the conception of school accountability to publicly demonstrate that educators or schools are executing an excellent work (Butterfield, Williams, & Marr, 1999; Mehrens & Lehmann, 1984; Smith, Heinecke, & Noble, 1999). However, according to Firestone et al., (1998) and Guthrie, (2002), there are certain consequences whether the schools or instructors have attained or not the required standards.

### **Assessment is irrelevant**

As stated by Moni, Van Kraayenoord and Baker, (2002), proofs showed that learners differ on how negatively they perceive assessment. Maximum number of students becomes gradually negative towards assessment throughout the educational system because the students cannot relate to the specific topic. Assessment may be considered irrelevant to students if it is thought of as being bad or unfair. Students are most likely to consider irrelevant assessment practices such as group projects that do not give credit to individual effort (Hoffman & Rogelberg 2001), those that lack explicit criteria for assessment (MacLellan 2001), and end-of-year examinations (Kniveton 1996). Students have treated assessment as irrelevant tasks because they perceive as making high marks more difficult to obtain.

## Mathematics Context in the Philippines

There are some bases on making conception of assessment domain specific for mathematics. Brown and Hirschfeld (2007) explained that the strength of students' agreement with different purposes of assessment links their conceptions of assessment to achievement outcomes on standardized national assessments of Mathematics for high school students.

With regards to the status of Mathematics in the Philippines, Crisostomo (2000) stated that Filipino students are still weak in Mathematics and Science. This is supported by the results of the Third International Mathematics and Science Study-Repeat (TIMSS-R). Similarly, a report of the Department of Science and Technology (DOST) said that 6,601 Filipino students who took the TIMSS last 1999 ranked 36th in both Science and Mathematics tests out of a field of 38 countries. The Philippines also placed 41st in Science and 42nd in Mathematics among 45 countries in the 2003 Trends in International Mathematics and Science Study. The Philippines ranked low on the said standardized tests; this is the reason why it is needed to have a quadrennial international assessment of Mathematics and Science skills among primary and secondary schools (Torregoza, 2014). The present study focused on Mathematics to determine if conceptions of assessment have an impact on Mathematics performance of college students.

The goal of the present study is to determine the conception of assessment whether college students improves and makes the students and school accountable in learning or if it is irrelevant. Given the studies by Brown and Hirschfeld (2009), the validity of conceptions of assessment in Mathematics is further supported if it can predict students' performance in Mathematics. The present study aims to determine if the same perception of assessment in Mathematics is applicable to college students and if it predicts students' grade in mathematics. Also, one of the purposes of the study is to determine if conceptions of assessment can be tried to other field of discipline aside from Mathematics.

In the present study, a new set of items on students' conceptions in Mathematics were anchored in the research model of Brown (2008). This provides further understanding on college students' views on assessment in Mathematics. The present study also confirms if the factors matched the data and contextualize students' conception of assessment in Mathematics in the Philippines.

## Method

### Participants

The participants in the study are 300 college students in universities of National Capital Region, Philippines. The participants are from ages 16-25 years old, 103 are males and 197 are females. Most of the respondents answered the scale inside their schools and some are while they are at home.

## **Instrument**

The developed scale consists of 40 items conferred in a response-type format with five options. The 40 items were generated based on the four domains of conceptions of assessment, namely: (1) Improvement, (2) Student Accountability, (3) School Accountability and (4) Irrelevance. These four domains were introduced in Brown's study. There are 10 items created from each domain based on the definition provided by Brown (2010).

The development of the scale was initiated by determining the concept, framework or theory. The items had gone through content validity which is conducted through consultation with experts (Magno, 2012). The expert checked whether the items were enough to cover the content of the instruction provided and if items were appropriate for the definitions intended. The process also involved correcting the items if they were appropriately phrased and whether the items were relevant to the subject area tested. The respondents rated the items by selecting the options if they (5) strongly agree, (4) agree, (3) neutral, (2) disagree and (1) strongly disagree on each item presented on the scale.

## **Procedure**

The researchers randomly selected students in Manila to answer the scale. Initially, respondents were asked if they are willing to answer the scale. Secondly, they were informed that they will be rating items on a scale about their conception of assessment in Mathematics. Then, questionnaires were distributed to each of the respondents. The questionnaire also determines the respondents' information such as name, course, gender, school, year level, and GPA (last semester). The instruction in answering was clearly indicated on the scale and it was also explained to them verbally before answering. They were also advised that if they have any clarification, they can raise it before answering. All the inquiries was acknowledged and answered by the researchers. The respondents rated the items by putting a check mark to the appropriate box depending on how they conclude assessment based on their experiences as students taking or have taken Mathematics subjects. All of the items were answered appropriately. The respondents answered the scale for no more than 30 minutes.

## **Data Analysis**

The responses of the respondents were input into a spreadsheet. The data gathered were used to determine the validity and reliability of the scale or new set of items of conceptions of assessment in Mathematics. For the descriptive statistics, means, standard deviation, kurtosis, skewness, and confidence interval were obtained.

To test the reliability of the scale, the researchers computed the Cronbach's alpha.

To test the validity of the scale, the items had gone through content validity where it was examined and revised with the help of an expert. Convergent validity in which the four sub-constructs of the scale were correlated to each other and the researchers did factorial validity using Confirmatory Factor Analysis to confirm whether the factors are valid when contextualized in the Philippines. The researchers were able to determine if the items were in good fit using the fit indices Joreskog GFI, Joreskog AGFI, Akaike Information Criterion, Schwarz Bayesian Criterion, Browne-Cudeck Cross Validation Index, Independence Model Chi-Square, Independence Model df, Bentler- Bonett Normed Fit Index, Bentler - Bonett Non - Normed Fit Index, Bentler Comparative Fit Index, James Mulaik-Brett Parsimonious Fit Index, Bollen Rho, and Bollen Delta.

## Results

This part presents the results obtained from pre-testing the scale of conceptions of assessment to 300 college students in National Capital Region (NCR). The reliability of the scale was assessed using the Cronbach's alpha. The coefficient alpha determined the internal consistency of the 40 items as a whole and for each factor (10 items each). Presented also in this part is the descriptive statistics that consists of means, standard deviation, and confidence intervals. The construct validity of the scale was first assessed using convergent validity by inter-correlating the four domains. The factor structure of the scale was also tested by comparing the four-factor model with a one-factor model.

Table 1

*Descriptive Statistics: Means, Standard Deviation, Cronbach's Alpha, CI+, and CI-*

Domains	M	SD	Kurtosis	Skewness	Cronbach's Alpha	CI+	CI-
Whole Scale	3.64	0.56	0.6	-0.19	0.88	3.83	3.45
Improvement	3.88	0.64	1.85	-0.85	0.89	4.09	3.67
School Accountability	3.81	0.64	1.62	0.75	0.5	4.26	3.36
Student Accountability	3.86	0.65	1.13	-0.72	0.89	4.08	3.64
Irrelevance	3.00	1.00	-0.87	-0.02	0.94	3.24	2.76

The means of the whole scale 3.64 which indicate that the average score of the respondents is near 3 to 4. Standard Deviation or SD was used to determine the heterogeneity or unevenness of the scores. In this table, SD is 0.56 which means that

some scores are one unit away from the average score of 3.64. In Kurtosis, we noticed on table 1 that the value of the whole test is 0.6 which means that the distribution is more likely platykurtic since the kurtosis exceeds the normal distribution, that is approximately 0.263. Similarly, in Improvement, School Accountability, Student Accountability, Irrelevance, it is platykurtic because the value obtain is 1.85, 1.62, 1.13, and 0.87 which absolute values are greater than 0.263. The whole scale and all of the factors are skewed to the right because the value is positive.

For the whole scale, the Cronbach's alpha of the 40 items is 0.88. This indicates high internal consistency since the value is greater than 0.4.

For each subtest, we can see on the table that the Cronbach's alpha of each factor attained internal consistency except for school accountability. The Cronbach's alpha of the four sub-factors are 0.89, 0.05, 0.89 and 0.94 which indicates adequate internal consistency because the obtained values were greater than 0.4.

*Table 2*  
*Convergent Validity*

Domains	Improvement	Student Accountability	School Accountability	Irrelevance
Improvement	-----			
Student Accountability	.76 *	-----		
School Accountability	.80 *	.84 *	-----	
Irrelevance	.08	.27*	.23*	-----

\* $p < 0.05$

When the subscales of the conceptions of assessment were inter-correlated, significant correlation coefficients were obtained ( $p < 0.05$ ). The four factors are significantly correlated at confidence level of 95 % ( $\alpha = 0.05$ ). However, Improvement and Irrelevance are not significant at ( $\alpha = 0.05$ ).

The factor structure of the whole scale was tested using the Confirmatory Factor Analysis. A four-factor structure was compared to a one-factor structure measurement model was done to determine which factor structure best fits the data. A four-factor structure is composed of four domains name improvement, school accountability, school accountability, and irrelevance. The items were used as indicators for the domains. There were 10 items on each latent factor. In the one-factor model, all indicators were combined in one latent variable. To determine which solution explains best the factors of the scale, the goodness of fit indices were compared. The four-factor structure of the scale turned out to have the best fit as opposed to a one-factor structure. The fit indices of the four-factor structure where the subscales are treated in separate latent variables had the best fit across several

indices. The proposed model had the highest GFI, AGFI, NFI, NNFI, CFI, PFI, Bollen's rho, and Bollen's delta (see Table 3).

**Table 3**  
*Fit Indices of the Different Measure Model*

Fit Indices	Values (Four Factor Model)	Values (One Factor Model)
Joreskog GFI	0.783	0.468
Joreskog AGFI	0.758	0.410
Akaike Information Criterion	63.40	13.468
Schwarz Bayesian Criterion	7.405	14.459
Browne-Cudeck Cross Validation Index	6.431	13.553
Independence Model Chi- Square	7858.094	7858.094
Independence Model df	780.000	780.000
Bentler- Bonett Normed Fit Index	0.781	0.508
Bentler - Bonett Non - Normed Fit Index	0.851	0.534
Bentler Comparative Fit Index	0.860	0.558
James Mulaik-Brett Parsimonious Fit Index	0.735	0.482
Bollen's Rho	0.767	0.481
Bollen's Delta	0.861	0.561

*Note.* Four-Factor Model = Improvement, Student Accountability, School Accountability, and Irrelevance  
One-Factor Model = Whole Scale



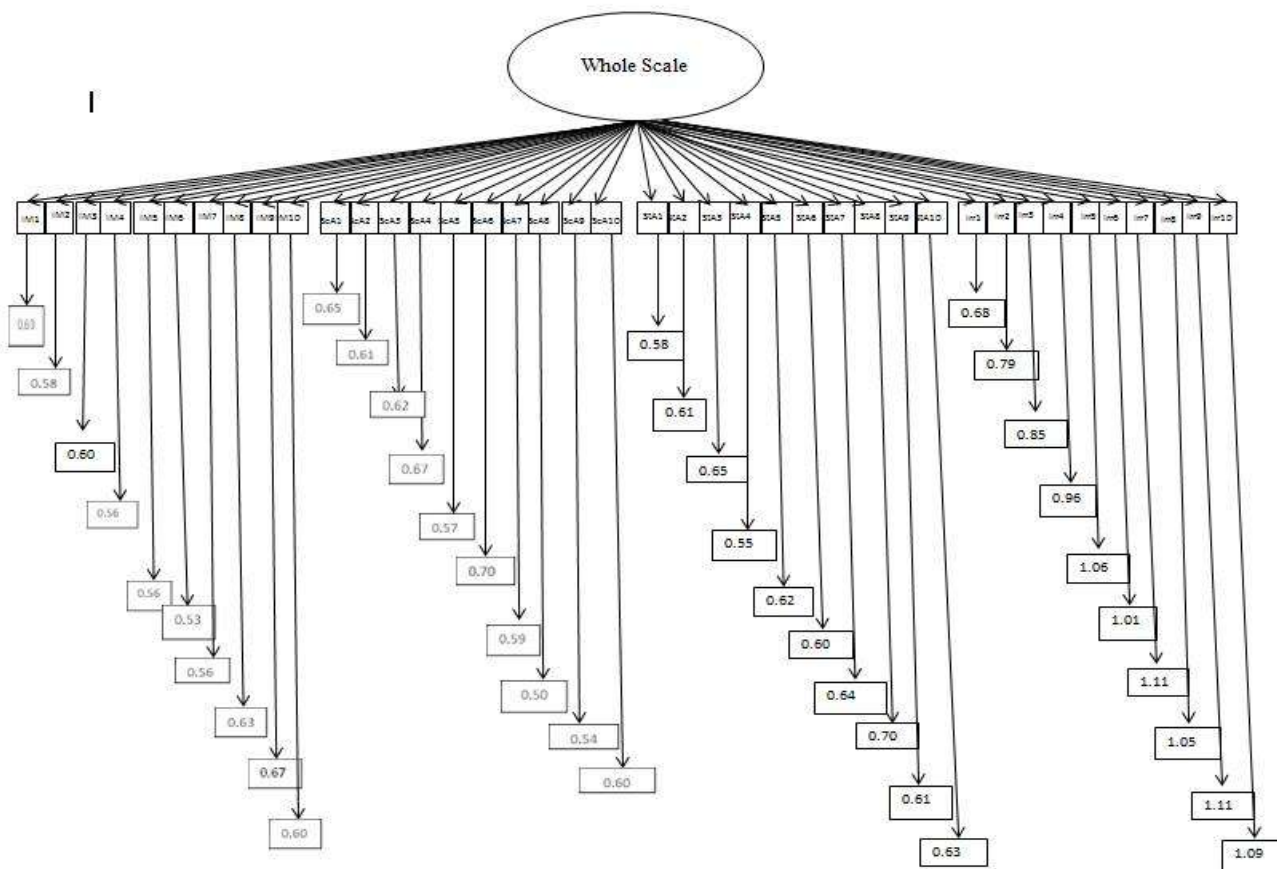


Fig.1. One-factor Model of a Scale

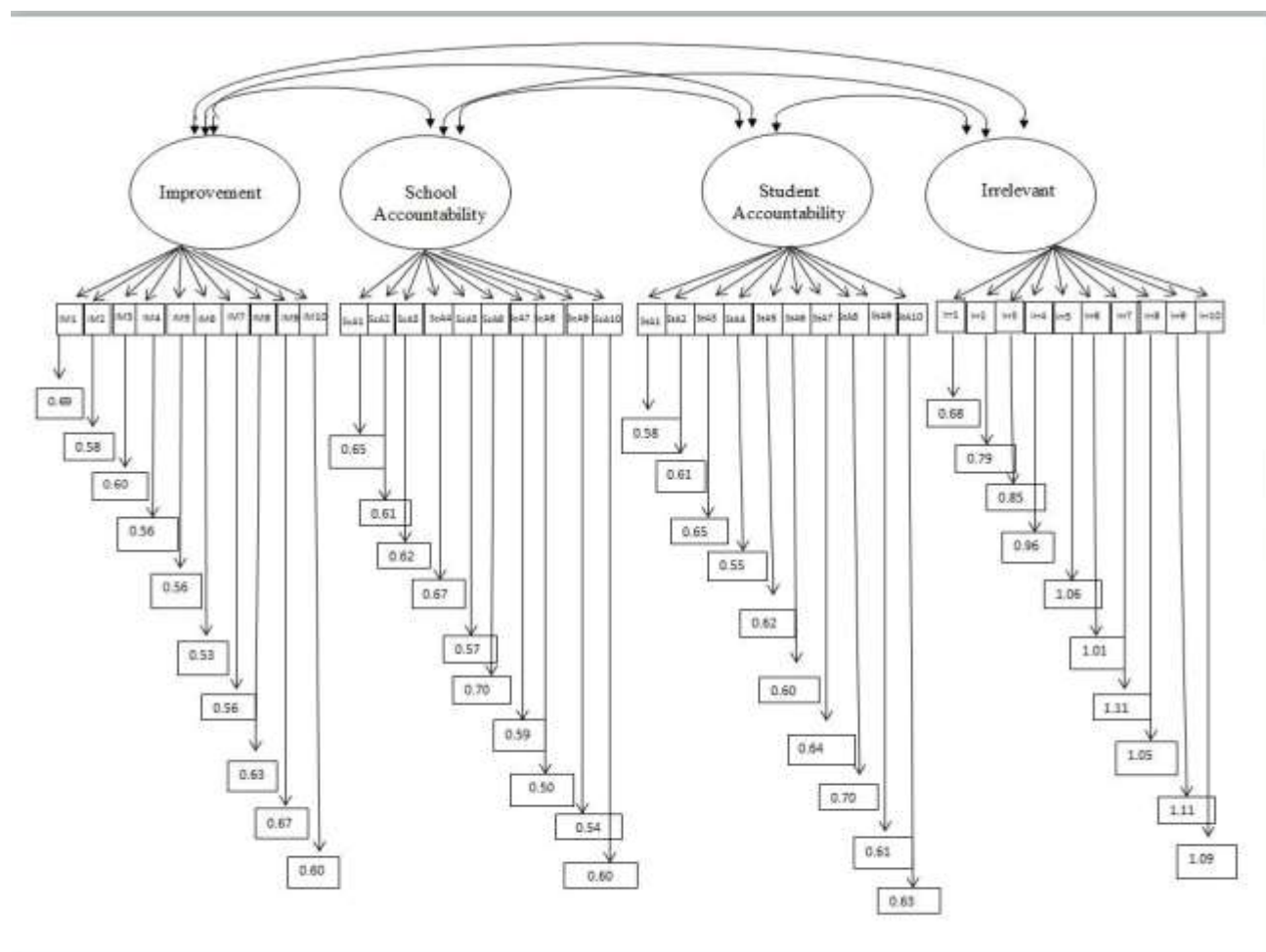


Fig.2. Four-factor Model of Whole Scale

Table 3  
Predictive Validity

Domain	Beta	SE	B	SE	t(295)	p
Intercept			1.26	0.29	4.39	0.00
Improvement	0.02	0.10	0.03	0.12	0.23	0.81
School Accountability	-0.03	0.11	-0.03	0.13	-0.25	0.80
Student Accountability	0.20	0.12	0.23	0.14	1.70	0.09
Irrelevance	0.18	0.06	0.12	0.04	3.16	0.00*

\* $p < 0.05$

Multiple regressions were conducted to establish the predictive validity of the factors of conception of assessment on students' grades in mathematics. For the four factors, Irrelevance is the only factor that is predicted by the grades which is significant at  $\alpha = 0.05$ . The rest of the factors namely improvement, school accountability, and student accountability are not significantly predicted by the grades.

### Discussion

The main purpose of the study is to develop a new set of items on conceptions of assessment in Mathematics applicable for college students. Four factors were proven where items are created to determine and analyze the conceptions of college students when it comes to assessment in Mathematics. The instrument constructed turned out to have adequate evidence of reliability and validity. The instrument is reliable where the items are internally consistent based on high values of the Cronbach's alpha. The scale was also proved to be valid based on the convergence of the four factors. Factorial validity was established where a four factor-structure explained the most adequate solution to fit the data supporting the characteristics identified by Brown (2008). Fit between the domains of the conceptions of assessment and the items are very much evident in the four factors. Theoretical implications and new findings of the study discusses in this section.

The researchers used Cronbach's alpha to determine if the items were internally consistent. The values obtained using Cronbach's alpha of the whole scale and the four domains namely: (1) improvement, (2) student accountability, (3) school accountability, and (4) irrelevance indicate high internal consistency of items. Having achieved internal consistency among the items indicates that there is similarity how the respondents answer each items within a subscale. The items are the same in reference to the construct being measured. The participants while answering the scale have in mind that their conceptions of assessment as taking any Mathematics subjects are being measured. This evidence of internal consistency implies that there is a marked coherence among the items for a given factor. This coherence served as basis for considering the reliability of the items. The internal consistencies of the factors from the study of Brown (2008) has values of 0.71, 0.63, 0.88, and 0.77 using Cronbach's alpha for the factors improvement, school accountability, student accountability and irrelevance. The present study shows higher internal consistency using Cronbach's alpha obtaining values of 0.89, 0.5, 0.89 and 0.94 for the factors improvement, school accountability, student accountability and irrelevance respectively. This means that in the present study, the new develop set of items are more internally consistent as compared to the set of items on the previous study made by Brown (2008). This is because the previous study has items that are domain-general as compared to the items of the present study that specifies in Mathematics when it comes to conceptions of assessment.

The significant inter-correlation among the factors showed that as one subscale increases, the others also increase. The four domains converge well. This evidence of convergent validity showed that the four factors have a commonality and refer to a

general conception of assessment of the respondents. The convergence of the factors is consistent the CFA as well. This shows that the four domains have potential to be developed even if one is weak in one area. In the study of Brown (2008), there was a tendency for students to respond to each conception in the same way. While improvement and accountability may be seen as antagonistic purposes, certainly for the students the correlations suggest that these conceptions have between half and one-third of the variance between them in common (0.59, 0.41, and 0.36, respectively). In contrast, the present study indicates that improvement, school accountability, and student accountability are conceived by the respondents as relevant and protagonistic purposes that are significantly correlated which values are 0.76, 0.80, 84. There is a same result in the previous study which indicates that the factor irrelevance of assessment is most associated with it not providing improvement -- oriented information to either the student or the teacher. There is a high significant inter-correlation of factors school accountability and student accountability. This means that the students' responses on items which belong to these two factors are relatively the same. The two factors are both accountability. The non-significant correlation of factors assessment as improvement and assessment as irrelevant means that the students' responses on the items which belong to this factors are relatively different. This is because the items in the factor improvement elicit positive conception on assessment while the factor irrelevant has items which are negative direction. Negatively-keyed items elicit negative conception on assessment in Mathematics. The results therefore show that the respondents have positive conceptions of assessment in Mathematics since in the items for the factor irrelevant is negative and the responses on the two factors are relatively different which means to say that there is a positive conception of assessment. For instance, the student responded strongly disagree in one of the items in the factor irrelevant, it means to say that he conceived that particular item as relevant because of the negative factor responded negatively.

Factorial validity was established where a four factor-structure explained the most adequate solution to fit the data supporting the characteristics identified by Brown (2008). The four factors that include improvement, school accountability, student accountability, and irrelevance were confirmed using a sample of 300 college students. Confirming a four-factor model as compared to fewer factors means that there are several characteristics considered in determining the sub-factors of the construct. This provides a multidimensional perspective focus on the conception of assessment in multiple factors. The four domains were proven where items created accurately depict what the learners conceive about assessment. The items of the four factors also showed fit using several fit indices. This means that the items constructed in the scale are appropriate given the range of characteristics and attitudes of the respondents. For improvement and irrelevance, only one item is not significant. This indicates that all items are appropriate measures given the college students' conception of assessment. In the previous study, it appeared that secondary students conceived assessment as first being about improvement, second about accountability while not being irrelevant, and only a little being a positive social or personal experience. In other words, students appeared to be thinking: "Assessment helps

improve my learning and it holds me accountable but it doesn't feel good. However, it is a bad thing when it doesn't help me or my teacher to improve my learning''. This interpretation has a similarity in the present study in a way that it makes the student more focusing on being accountable than improvement in learning. The interpretation of conception of assessment indicates that the respondents conceived assessment in Mathematics as irrelevant if it does not help them improve learning most especially if they cannot relate the topic to real life situations.

There are five major theoretical implications of the present study. (1) The findings validate the studies of Brown (2008) that theorized; (a) assessment as improvement, (b) school accountability, (c) student reliability, and (d) irrelevance as valid factors of conceptions on assessment, (2) The translation of conceptions of assessment of college students in Mathematics is that the model combines a holistic perspective in assessing the conception of the college students such that represents different broad domains such as the cognitive, social, and affective. How the students conceive assessment based on their prior knowledge is the cognitive aspect. For social aspect, their conception is based on their experiences in life (Quizzes). For affective, the respondents conceived assessment based on their eagerness to learn where they visualize assessment as helpful or not in their learning. These facets are considered important in the improvement of learning that helps the students attained the learning. The four factors confirmed shows that learners who conceived assessment is helpful in learning are important as in the case of thinking positively for the factors improvement, school accountability, and student accountability will result to positive output. Similarly, college students who conceived assessment as harmful or irrelevant will result more likely to a low academic performance and the possibility of non-participation. This supports the explanation of Brown and Hirschfeld (2007) that higher overall mathematics achievement was found among students who conceived that assessment makes students themselves accountable for learning and who conceived of assessment as good for them. In contrast, the more students agreed that assessment interferes with learning or assessment is ignored, the lower their mathematics achievement is. The results calls for higher education institutions to strengthen the conceptions of assessment to students as the teachers need to instill in the minds of the students that assessment is beneficial and relate it in classrooms especially in the field of mathematics in the Philippines in real-life applications. (3) The study in conceptions of assessment can be tried to other field of discipline aside from Mathematics such as Science, Mathematics, and English and so on since the "domain-general" conceptions of assessment and its factors were already validated as factors specified in Mathematics. (4) The study indicates that conceptions of assessment and its factors can be contextualized into other countries aside from the Philippines. (5) In the study of Brown (2008), the original conception of assessment as irrelevant indicates that students are quite sensitive to assessment which they perceive to be unfair, bad, or irrelevant to them. This depicts negative conception of assessment. However, based on the results in the present study, the factor irrelevance appears to be in need of further classification of items loaded to it. The mean of the responses of the respondents in the factor irrelevance is exactly 3.00 which indicate neutrality of the responses. The responses depict that the respondents

do not view assessment as harmful but rather no effects in learning. This means that several sub-factors could be produced out of the factor irrelevance. This can be done by future researchers by utilizing Principal Components and Classification Analysis. The future researchers should look into whether it could either be harmful in the learning of the students or there are no effects in learning at all.

### References

- Brown, G. T., Hui, S. K., Yu, F. W., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307-320.
- Brown, G. T., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27(1), 210-220.
- Brown, G. T., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28(7), 968-978.
- Brown, G. T., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19(2), 97-111.
- Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19(2), 97-111.
- Brown, G. T., & Hirschfeld, G. H. (2007). Students' Conceptions of Assessment and Mathematics: Self- Regulation Raises Achievement. *Australian Journal of Educational & Developmental Psychology*, 7, 63-74.
- Crisostomo, S. (2000, December). Filipino students still rate low in math, science. Retrieved from <http://www.philstar.com/news-feature/91528/filipino-students-still-rate-low-math-science>
- Hui, S. K., & Li, W. S. (2008). Conceptions of assessment of mainland China college lecturers: A technical paper analyzing the Chinese version of COA-III. *The Asia-Pacific Education Researcher*, 16(2), 185-198.
- Mao, J. (2008). The effects of assessment strategies and self- regulated learning (SRL) skills on college students' skill-based and cognitive learning outcomes and perceptions of assessment for learning. (Unpublished doctoral dissertation). The Pennsylvania State University, USA.