

ISSN 20194-5876

Educational Measurement and Evaluation Review

VOLUME 7 JULY 2016

THE PHILIPPINE EDUCATIONAL
MEASUREMENT AND EVALUATION
ASSOCIATION (PEMEA)



The Educational Measurement and Evaluation Review (EMEReview) is the official publication of the Philippine Educational Measurement and Evaluation Association (PEMEA). It is international, refereed, and abstracted/indexed. The EMEReview publishes scholarly reports about contemporary theories and practices in the field of education and social science that highlights measurement, assessment, and evaluation. It welcomes articles that are about test and scale development, quantitative models of a construct, evaluation studies, best practices in evaluation, issues and policies on assessment, contemporary approaches in educational and psychological measurement, and other studies with direct implication to assessment in education, social science, and related fields. EMEReview is indexed/abstracted in the Open J-Gate, JournalTOCs, Google Scholar, InfoBase Index, Social Science Research Network, Open Academic Journals Index, and ejournals.ph

Copyright © 2015 by the Philippine Educational Measurement and Evaluation Association.

This journal is open-access and users may read, download, copy, distribute, print, search, or link to the full texts, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

The articles in the EMEReview are open access at
<http://www.pemea.org/emereview>



Publication Division of PEMEA
Philippine Educational Measurement and Evaluation Association

◆ Editorial Board

Editor:

Dr. Adonis P. David, Philippine Normal University
editorpemeajournals@gmail.com

Managing Editor:

Dr. Carlo Magno, National University
crlmgn@yahoo.com

Associate Editors:

Dr. Marcos Lopez, Centro Escolar University-Malolos, Philippines
Dr. Richard Gonzales, Development Strategists International Consulting
Dr. Marilyn Balagtas, Philippine Normal University, Philippines
Dr. Teresita T. Rungduin, Philippine Normal University, Philippines
Mr. Jesus Alfonso Datu, The University of Hong Kong, Hong Kong
Ms. Belen Chu, Philippine Academy of Sakya Philippines. Arellano University
Ms. Marife Mamauag, HELP University, Malaysia

Editorial Advisory Board

Dr. John Hattie, University of Melbourne, Australia
Dr. Jack Holbrook, University of Tartu, Estonia
Dr. Anders Jonsson, Malmo University, Sweden
Dr. Timothy Teo, University of Macau, China
Dr. Tom Oakland, University of Florida, USA
Dr. Jimmy dela Torre, Rutgers University, USA
Dr. Jose Pedrajita, University of the Philippines-Diliman, Philippines
Dr. Shu-ren Chang, Department of Testing Services, American Dental Association, USA
Dr. Karma El Hassan, Office of Institutional Research and Testing, Americal University of Beirut, Lebanon
Dr. Alexa Abrenica, Professional Regulation Commission
Dr. Marie Ann Vargas, University of Sto. Tomas, Philippines

Reviewers for this issue:

Dr. Adonis David, Philippine Normal University
Dr. Carlo Magno, National University
Mr. Ryan Cayubit, University of Sto. Tomas
Dr. Jonathan Macayan, Mapua Institute of Technology
Dr. Marilyn Balagtas, Philippine Normal University
Dr. Violeta Valladolid, De La Salle University

Editorial Note

The Editors 1

The Role of Perceived Autonomy, Supportive Teachers' Behaviors, Situational Interest, and Self-Regulated Learning Strategies on Mathematics Academic Achievement

Carmen Riva 3

Exploring the Predictors and Outcomes of Academic Resilience among College Students

Analyn O. Leysa and Filmor A. Malnegro..... 24

Understanding the Structure of Autonomy among Filipino Adolescents

John Jamir Benzon R. Aruta..... 37

Multigroup Invariance of Course Experience Questionnaire Across Gender Using Malaysian Sample

Lei Mee Thien.....
50

Assessment Schemes in the Senior High School in the Philippine Basic Education

Carlo Magno and Tristan Piosang 66

Assessment Systems in Southeast Asia: Models, Successes and Challenges

Educational Research Unit, *SEAMEO INNOTECH* 88



Editorial Note: Open Access Advocacy

The volume 7 of the Educational Measurement and Evaluation Review (EMER) strengthens the theory in educational measurement that is available for various readers. The Journal through its editorial board takes part in the promotion of open access publications. The association advocates in making knowledge available to different scholars all over the world especially in the areas of educational measurement and assessment. According to Super (2002), “open access to scientific journal articles means online access without charge to readers or libraries. Committing to open access means dispensing with the financial, technical, and legal barriers that are designed to limit access to scientific research articles to paying customers. It means that, for the sake of accelerating research and sharing knowledge, publishers will recoup their costs from other sources.” Knowledge generated and scientific information needs to be shared to the public for replicability, faster growth in theoretical development, expand discussions, and wider collaboration of scholars around the world. The Philippine Educational Measurement and Evaluation Association take part with the initiative of the Global Young Academy on three advocacies on open access publications: (1) Making knowledge available, (2) avoiding inequality in publishing, and (3) enjoining scholars to participate in the advocacy.

The articles in this volume touches on testing and advancing theoretical model and assessment practice.

The article by Riva (2016) expanded the construct situational interest and its impact on students’ mathematics achievement in the grade school. Through careful assessment and adaptation of the scales for grade school students, it is noteworthy to conclude that supportive environments facilitate situational interest and self-regulation.

The article by Leysa and Malnegro (2016) advances the construct on academic resiliency among college students by providing new insights on the model by Andrew Martin. They found that anxiety plays an important part in developing academic resilience among college students. Consistent with literature, academic resilience predicts

the desirable educational outcomes of enjoyment in school, class participation, and general self-esteem.

The article by Aruta (2016) provided new insights on the structure of autonomy using the self-determination theory by the Deci and Ryan. Using qualitative data analysis and exploratory factor analysis, two dimensions of autonomy emerged—inward autonomy and outward autonomy.

The article by Thien (2016) analyzed the measurement invariance of the College Experience Questionnaire with gender as a source of variation. Findings showed that configural invariance was fully supported whereas metric invariance was partially supported across female and male groups.

The article by Magno and Piosang (2016) provided many possible opportunities to assess senior high school students with the senior high school in its initial implementation in the Philippines. School administrators, assessment specialists, psychometricians, and teachers are guided with the assessment programs that can be implemented at this stage. The levels of assessment with varied purposes provided support on evaluating how well the new curriculum is implemented in a school as well as in the national level.

Finally the article of the South East Asian Ministers of Education explains how the assessment system in Asia translates into student learning. They defined the concepts of assessment “for” learning and assessment “as” learning as new reforms on assessment in schools. The article emphasize on the utility of assessment in helping students further learn. Recommendations are provided.

References

- Super, P. (2002). Open access to the scientific journal literature. *Journal of Biology*, 1, 3.
- Global Young Academy. (2016). Open access pledge. Retrieved from <https://globalyoungacademy.net/news-release-gya-members-pledge-for-open-access/>

The Editors



The Role of Perceived Autonomy, Supportive Teachers' Behaviors, Situational Interest, and Self-Regulated Learning Strategies on Mathematics Academic Achievement

Carmen J. Riva

Romblon State University

Abstract

The present study tested a model highlighting that effects of specified teachers' autonomy supportive behaviors on mathematics are explained by the mediating roles of situational interest and self-regulated learning strategies. The participants were 417 fourth year students in the secondary level who beforehand were evaluated to be low in perceived teachers' autonomy supportive behaviors. These participants were exposed into series of mini-lessons utilizing 4A's strategy instruction. The Learning Climate Questionnaire (LCQ), Situational Interest Survey (SIS), Academic Self-Regulation Scale (A-SRL-S), and mathematics test were administered after the students have engaged in the mini-lessons. Using Structural Equations Modeling, results confirmed that situational interest and self-regulated learning strategies explain the effect of autonomy supportive behaviors on mathematics academic achievement. The tested model attained an adequate fit ($\chi^2 = 376.795$, $df = 147$ or $CMIN/df = 2.535$, $RMSEA = 0.061$, $TLI = 0.925$, $CFI = 0.936$, and $IFI = .937$).

Keywords: Autonomy supportive behaviors; situational interest; self-regulated learning strategies; mathematics achievement

Introduction

Autonomy is a universal psychological need that, when satisfied, provides an individual a feeling of independence and identity, but when stifled will result in maladaptive practices. Learners

have high need for autonomy, which is critical for their learning. In this light, teachers' autonomy supportive behaviors should be examined to find whether they provide the necessary autonomy support during mathematics instruction. Autonomy influences situational interest (Tsai et al., 2008), a construct which learners perceive is absent during math instruction. In fact, some authors presumed that students achieved poorly in mathematics because they do not perceive the subject as attractive, interesting, and engaging (De Corte, Verschaffel, & Depaepe, 2008). Autonomy supportive behaviors not only influence students' interest during instruction, it promotes a chain effect on the students' behavior prompting them to use self-regulated learning strategies (Dohn, Madsen, & Malte, 2009). When students perceive the classroom environment as interesting, they engage in activities and match their skills against the challenges set by teachers (Turner et al., 1998). It is this experience of interest which drives learners to adopt and enhance their use of self-regulated learning strategies (Sierens, Vansteenkiste et al., 2009).

Majority of learners have poor study habits resulting in low mathematics achievement, indicating that they need to be trained to use self regulated learning strategies (Camahalan, 2006). The lack of well-defined self regulated learning strategies may also prove lack of interest on their part as learners, and which can be traced from lack of support to their need for autonomy. This scenario provides the reason why the necessity to investigate more teachers' autonomy supportive behaviors. A number of studies have been conducted showing the direct impacts of autonomy supportive behaviors, situational interest and self regulated learning strategies on math achievement (Tsai, Kunter, Ludke, Trautwen, & Ryan, 2008; Hidi & Harakiewicz, 2000), however, no indirect effects of autonomy to math achievement via situational interest and self regulated learning strategies were specifically known.

This study aims to explain how perceived teachers' autonomy supportive behaviors affect math achievement. It provides explanations on the direct and indirect paths between autonomy supportive behaviors, situational interest and SRL which occur in an autonomously supportive learning context.

Teachers Autonomy Supportive Behavior

Self determination theory (SDT) accounts that autonomy pertains to actions that are self-endorsed and are based on self-

integrated values or interests. It is a basic and universal human need and therefore opportunities to experience it is critical to well-being (Ryan & Deci, 2000). SDT argues that people are naturally prone to self-organized actions and that the sense of choice, congruence, and initiative that characterize autonomy are necessary aspects of healthy functioning. SDT strongly points out that if autonomy is not supported, it will have deleterious impact on well-being (Deci, Koestner, & Eyan, 1999). In classrooms, a teacher who is autonomy supportive ascertains learners to be free of internal hindrances and external manipulations, coercion, and gives adequate options to choose from (Adams, 2007). Autonomy is not limited and bounded by culture, whatever the kind of society is, support for autonomy is a fundamental psychological need among adolescents (Chirkov & Ryan, 2001). It is manifested when teachers develop classroom opportunities for students to align their inner resources with their classroom activities (Assor, Kaplan, & Roth, 2002).

Studies documented positive effects of autonomy on academic achievement. Autonomy enhancing instructional behaviors directly affect students' cognitive and behavioral engagement with academic tasks (Assor, Kaplan, & Roth, 2002). During mathematics instruction, students who perceive teachers to be autonomy supportive of their learning obtained higher academic performance (Tsai et al., 2008; Wilsons & Simons, 2002). In a study of 135 sixth-grade and 91 ninth-grade regular education students in Southern California, stepwise regression analyses indicated that teachers' autonomy support among others, predict mathematics performance (Wong, Weist & Cusick, 2002).

Evidences proved that autonomy supportive teaching behaviors trigger development of situational interest (Stefanou, Perencevich, DiCintio, & Turner, 2004; Assor, Kaplan, & Roth, 2002; Reeve & Jang, 2006; Grolnick, Ryan, & Deci, 1991). Studies conducted on 261, 7th grade students in Germany found that autonomy supportive behaviors predicted students' situational interest in math (Tsai, et al, 2008). Positive feelings, valuing tasks, and engagement characterize high situational interest. Similarly, a study conducted among 95 sophomore students at Dutch University found that variance in students' situational interest was generated due to their satisfaction in terms of psychological need in which one is autonomy (Boekaerts & Minnaert, 2006).

Teachers' autonomy supportive behaviors can stimulate students to adopt self-regulatory learning strategies. In a study on regulatory fit, students displayed equally autonomously regulated learning strategies when they perceived the environment to be autonomously supportive, but, demonstrated maladaptive strategies when the learning climate was perceived to be controlling (Pierro, et al., 2009). In a particular study of 526 Belgian students, teachers' autonomy support was found to be positively correlated with students' self-regulated learning. These findings were in line with Self-Regulated Learning Theory stating that these processes are assumed to be influenced by environmental stimuli (Zimmerman, 1989). A study further suggests that if teachers want their students to evaluate themselves, to plan their study activities, and to think about themselves as learners, they are expected to provide instruction and expectations in an autonomy-supportive way (Pierro, et. al, 2009).

Situational Interest

Interest as a psychological construct is of two types: the individual or personal type and situational interest. Situational interest is characterized by positive emotions and heightened concentration. It is spontaneous, environmentally activated, assumed to be transitory, context specific, and is provoked by external and situational stimuli rather than by individual variables (Hidi & Renninger, 2006). Once elicited, situational interest facilitates cognitive functioning and learning (Hidi, 1990). It can be initiated by something in the environment, and is more amenable to change (Schraw & Lehman, 2001).

Development of interest usually goes into phases, such as triggered-situational interest, maintained-situational interest feelings and maintained situational interest values. Triggered-situational interest (Trig-SI) usually comes out at the onset of individual's association with the environmental stimuli, and is similar to "catch". It refers specifically to *initiating* interest (Hidi & Harackiewicz, 2000; Hidi & Renninger, 2006). Maintained-situational interest feeling (MSI-F) develops due to the extent wherein the material itself was enjoyable and engaging. Lastly, maintained-situational interest-values (MSI-V) develops based on whether the material was viewed as important and valuable (Linnenbrink-Garcia, Durik, Conley, Barron, Tauer, Karabenick, & Harackiewicz, 2010).

Situational interest is relevant to learning. It may best contribute to the beginning stages of knowledge acquisition (Hidi, 1990). The triggered situational interest serves as a “spark” which would ignite student’s curiosity to learn more and engage in mathematical concepts, facts, ideas, processes and principles being presented. More engagement in challenging tasks, experiences of positive affect, enjoyment in dealing with difficult tasks and seeing the tasks as significant are evidences of interested students that lead more to higher achievement in math (Harakiewicz, Barron, Durik, Tauer, and Linnenbrink-Garcia, 2008).

Situational interest affects students’ use of learning strategies (Schiefele, 1991). It energizes learners to adopt self-regulated learning strategies (Sorić & Palekčić, 2009). Results of a study revealed that the level of interest or enjoyment in math is a significant predictor of students’ reports of strategy use which accounted for 14% of the variance (Cleary & Chen, 2009). Another findings proved that when situational interest is enhanced, it promoted the use of self-regulated learning strategies specifically when the instructional setting is less structured (Dohn, Madsen, and Malte, 2009).

Self-Regulated Learning Strategies

Self-regulated learning strategies are activities or actions which students use to perform well and obtain successful academic outcomes (Zimmerman, 1989). They can be understood as an organized, conscious, and intentional whole of what the student does to efficiently accomplish a learning objective in a given social context (Gargallo, Suárez-Rodríguez, & Pérez-Perez, 2009). These strategies are used by independent learners to control their learning, focus in their studies, plan and study (Magno, 2009; 2011a). Generally, it is through self-regulated learning in which students activate and sustain their cognitions and behaviors and systematically orient themselves toward the attainment of their learning goals (Zimmerman, 1989).

Academic Self Regulated Learning Scale (A-SRL-S) identified self regulated learning strategies which include memory, goal setting, self-evaluation, seeking assistance, environmental structuring, learning responsibility, and organizing.

Self-regulatory learning strategies predict academic success. They are proven as the best predictor of standardized achievement test scores (Zimmerman & Martinez-Pons, 1986). Higher level of use of

self-regulated learning strategies results to better academic achievement (Soric' & Palekic', 2009). The same results were shown in the study that all seven self-regulation strategies significantly predicted college students grade at the end of the term or semester (Magno, 2011a). Specifically, another study found that self regulated learning activities such as self-evaluation, monitoring, planning, goal setting positively affected mathematics achievement (Arzal, 2011).

Conceptual Framework

A conceptual framework was formulated which shows the direct and indirect effects of teachers' autonomy supportive behaviors, situational interest and self-regulated learning strategies on math achievement. This model shows the probability of the existence of indirect effects of teachers' autonomy supportive behaviors to impact math achievement by affecting first the students' situational interest, or their self regulated learning strategies. These indirect effects visualize how the influence is carried over to math achievement. The study may expose the probability of the mediating roles of situational interest and self regulated strategies in an autonomously supportive learning context.

The first indirect path presumed to exist in this study as the result of logically linking the constructs during model formulation is that students' perception of teachers' autonomy supportive behaviors could trigger development of situational interest (Stefanou, Perencevich, DiCintio, & Turner, 2004; Reeve & Jang, 2006; Grolnick, Ryan & Deci, 1991) and as a consequence, math performance improved (Hidi & Harakeiwicz, 2000; Harakeiwicz, Barron, Durik, Tauer, & Linnenbrink-Garcia, 2008). Situational interest serves as the construct through which effects of autonomy supportive behaviors would be carried over to math achievement. This indirect effect is expected to assume a certain magnitude which could explain how changes in autonomy affect situational interest, which in turn may affect math achievement.

When students perceive teachers to be autonomy supportive, they are stimulated to self-regulate (Grinsven & Tillema, 2006). More autonomy supports will be encoded as favorable experiences and thus prompt them to match their skills, to effectively engage in self-regulated learning strategies, and improve academic performance (Sierens et al., 2009). In addition, students who perceive teachers in

mathematics to be generally good at scaffolding their learning also engage more in self-regulated learning strategies. Teachers who explain the purposes and relevance of the tasks as well as link present concepts and prior knowledge facilitate students to put more meaning in and value their learning, resulting in more use of self-regulated learning strategies. Hence, increasing autonomy supports increases the use of a variety of learning strategies (Grinsven & Tillema, 2006), and in effect, the frequency of use of SRL consequently improves students' achievement (Zimmerman & Martinez-Pons, 1986; Soric' & Palekcic', 2009; Camahalan, 2006; Arzal, 2011).

Autonomy supportive behaviors influence students' use of self regulated learning strategies through situational interest. The model explains that students can be prompted to adopt self-regulated learning strategies when at that particular context, they experience situational interest. The model describes that autonomy supportive behaviors can at the same time induce learners to adopt self-regulated learning strategies by ensuring the presence of situational interest. Situational interest in this path serves as the construct through which autonomy supportive behaviors carry over its effect on self regulated learning strategies.

The formulated conceptual model identified the path indicating that students situational interest can influence math achievement through the presence of self-regulated learning strategies. This relationship shows that when students use more self-regulated learning strategies, situational interest impacts best mathematics achievement. It proposed a strategic way on how situational interest effects to math achievement could be maximized.

Teachers' autonomy supportive behaviors can trigger students to generate situational interest and make them adopt SRL strategies which would improve math achievement. During mathematics instruction, although students may perceive the lessons as difficult, instructions would appear as interesting and challenging through teachers supporting their autonomy. When teachers understand their feelings and show confidence on them, students continue to persevere in difficult tasks. Experiencing situational interest in classroom settings, positive influence on students' self-regulated learning strategies results, and mathematics achievement are also attained (Schiefele, 1991; Soric' & Palekcic', 2009).

Teachers' autonomy supportive behaviors stimulating students' inner resources to develop situational interest, and utilize SRL

strategies are embodied in Self Determination Theory. This organismic-dialectical theory views human being as proactive organisms but their natural or intrinsic functioning can be either facilitated or impeded by the social context (Deci & Ryan, 1985; 1991). These tendencies do not operate automatically, instead require ongoing social nutriments and supports (Ryan & Deci, 2000). This theory explains that individuals have high need for autonomy aside from competence and relatedness. Once all the necessary nutriments are provided continuously such as teachers' autonomy support, the most volitional and high quality forms of motivation and engagement result. The output would be persistence and creativity leading to higher performance. Thus, SDT serves as the bases for explaining relationships and connections as well as outcomes between perceived teachers' autonomy supportive behaviors, situational interest, and self regulated learning strategies.

Statement of the Problem

The main purpose of this study was to examine and explore how autonomy supportive behaviors affect mathematics achievement. It tried to examine how this construct behaves in the presence of situational interest and self-regulated learning strategies, concepts which have great roles in students' attainment of knowledge especially inside the classrooms.

Specifically, this study sought answers to the following research questions:

Direct Effects:

1. Do teachers' autonomy supportive behaviors directly increase students' achievement in mathematics?
2. Do teachers' autonomy supportive behaviors increase students' situational interest?
3. Do teachers' autonomy supportive behaviors stimulate students' adoption of self-regulated learning strategies?
4. Does situational interest directly increase students' achievement in mathematics?
5. Does situational interest directly improve student's use of SRL strategies?
6. Do self-regulation strategies directly increase students' achievement in Mathematics?

Indirect Effects:

7. When situational interest is optimal, how do autonomy supportive behaviors influence students' academic performance in mathematics?
8. When students' level of use of self-regulated learning strategies is high, how do autonomy supportive behaviors encourage learners to attain satisfactory academic performance in mathematics?
9. When situational interest is optimal, how do autonomy supportive behaviors advance learners to use self-regulated learning strategies?
10. When students' level of use of self-regulated learning strategies is high, how do situational interest affect academic performance in mathematics?
11. In general, how do autonomy supportive behaviors, situational interest and self-regulated learning strategies impact mathematics achievement?

Hypothesis

The following hypotheses were tested in this study;

Direct Effects

Hypothesis 1: Teachers' autonomy supportive behaviors directly improve students' academic achievement in mathematics.

Hypothesis 2: Teachers' autonomy supportive behaviors during mathematics instruction increase students' development of situational interest.

Hypothesis 3: Autonomy supportive behaviors instigate students to adopt self-regulated learning strategies in mathematics.

Hypotheses 4: When students' situational interest during mathematics lessons is high, learners adopt more and deeper level of self-regulated learning strategies.

Hypothesis 5: When students' situational interest is high, academic achievement in mathematics is expected to improve.

Hypothesis 6: Students adoption of self-regulated learning strategies will result to an improved achievement in mathematics.

Indirect effects:

Hypothesis 7: When situational interest is high, autonomy supportive teachers' behaviors promote students to attain better academic performance in mathematics.

Hypothesis 8: When students adopt more self-regulated learning strategies, autonomy supportive teachers' behaviors increase their academic achievement in math.

Hypothesis 9: When students' situational interest is high, autonomy supportive behaviors would encourage learners to adopt self-regulated learning strategies.

Hypothesis 10: When students' level of use of self-regulated learning strategies in mathematics is high, increased levels of situational interest sustain them to achieve better.

Hypothesis 11: In general, when autonomy supportive behaviors, situational interest and use of SRL strategies are experienced at high levels, students' mathematics achievement is improved.

Method

Research Design

A learning context was structured and maintained where teachers' autonomy supportive behaviors were ensured to exist and were held constant so as to affect student's development of situational interest, use of self-regulated learning strategies, and improve mathematics academic achievement. The explanatory, cross-sectional design was utilized and analysis of data were made across variables of autonomy supportive behaviors, situational interest, self-regulated learning strategies, and math achievement. Finally, Structural Equation Modeling (SEM) tested fitness of the proposed model to the data.

Participants and Procedure

A permission to conduct a research with the fourth year students as the participants was secured from authorities of Manggahan High School, Division of Pasig City, Philippines and proposed plans for gathering data were discussed.

The Learning Climate Questionnaire (LCQ) was administered and those students whose means were below 3.5 in the LCQ were

selected as participants. This is to ensure a context low in autonomy supportive behaviors. Only 417 out of the original number of 423, (58.04% males, mean age=15.38 years old) were able to complete the scales. Aside from having low levels of perceived autonomy supportive teachers' behaviors ($M=3.23$, disagree to strongly disagree to items of LCQ), the participants have varied socioeconomic background, with more than half (56.17%) of them having family monthly income lower than Php10,000.

Instruments

Learning Climate Questionnaire (LCQ). This is the short version scale that measures the degree to which the students perceive their teachers as supporting their autonomy. Items were adapted from the Learning Climate Questionnaire developed by Tsai et al., (2008), (e.g., “*I felt that my teacher provided me choice and options*”). The scale has good internal consistency, Cronbach's $\alpha=0.983$. Responses were given on a 7-point Likert type scale. Goodness of fit indices supported unidimensionality (CMIN/df=4.7; NFI=.96; IFI=.97 and TLI=.95).

Situational Interest Survey (SIS). This 12- item Likert type scale was developed by Lenninbrink-Garcia, Durik, Conley, Barron, Tauer, Karabenick, and Harackiewicz (2010) and is made up of three factors (triggered situational interest, maintained situational interest-feelings, and maintained situational interest-values). All indices of reliability of item scores were good. The subscales measured situational interest in Math domain; Triggered-SI consisted of 4 items; $\alpha=.813$ (e.g. “*My math teacher is exciting*”), maintained-SI-feelings have 4 items; $\alpha=.850$ (e.g. “*I am excited about what we are learning in math*”) and maintained-SI-values consists of 4 items; $\alpha=.811$ (e. g. *The things we are studying in math are important to me*). Responses were given on a 6-point Likert scale that ranged from 1 (*disagree strongly*) to 6 (*agree strongly*).

Academic-SRL-Scale (A-SRL-S). This is a self-report 55-item scale consisting of 7 subscales that measures self-regulated learning strategies in academic settings (Magno, 2010) and is made up of seven factors. The sub-scales have high internal consistency: Memory Strategies ($\alpha=.881$); Goal setting strategies ($\alpha=.677$); Self Evaluation ($\alpha=.848$); Seeking assistance strategies ($\alpha=.704$) Environmental structuring ($\alpha=.742$), Learning responsibility strategies

($\alpha = .714$); and Organization strategies ($\alpha = .725$). Responses are given on a 4-point Likert scale and ranged from 1 (never) to 4 (always).

Math Test. This is a 35-item test used to measure students' competencies in Mathematics 4 on the topic, "Graphing Quadratic Function". The test items were aligned to curriculum standards, checked by experts, and TOS has been prepared, thus attained content validity. The discrimination and difficulty indices were determined using the Classical Test Theory (CCT).. The reliability was established by determining its Cronbach's alpha (Table 1) ($\alpha = .712$ $p < .05$).

Data Collection

Collection of data went through 3 phases; Phase 1 is orientation, training, and seminar-workshop for teachers. The existence of an autonomy supportive learning context, situational interest and utilization of self regulated learning strategies was ensured in this phase. It included development of lesson plan, preparation of the modules and training design for teachers and conduct of the orientation, training and seminar workshop for math teachers. Phase 2 involved conducting and monitoring the 5-day series of mini lessons which exposed the participants under highly autonomous supportive teachers while phase 3 was the administration of the instruments. The scales were distributed at the end of each period while mathematics test was administered on the following week after the mini lessons.

Data Analyses

Two tests were done initially, to satisfy the assumptions of multivariate and SEM analyses; the directionality and data related assumptions. Temporal precedence and unidirectionality were ascertained. Tests for normality and homoscedasticity of the data were carried out. Normality was induced and homoscedasticity was diagnosed. Descriptive statistics and reliability estimates of each scale were calculated. Bivariate correlations were done and SEM with Maximum Likelihood Technique was utilized to estimate the coefficients between parameters to answer the research questions and hypotheses posed in the beginning of this study. Parameter estimates between latent variables indicating direct and indirect effects and coefficients of their corresponding indicators were analyzed with the

technique. This was followed by assessment of model fit (Joreskog & Sorbom, 1989). Interpretations and recommendations by Schermelleh-Engel, Moosbrugger, & Müller (2003) were applied.

Results

The model tested the directionality related assumptions specified in the hypothesis. Autonomy supportive behaviors, an exogenous variable, has arrows pointing to situational interest, self regulated learning strategies and math achievement which are the endogenous variables or the downstream variables. The latent constructs have arrows pointing to their respective manifest variables characterizing a reflective model. Disturbance and error terms have no covariance symbols connecting them. The model is recursive; theories and previous findings served as the bases for the specification.

Levene's test was similarly done to find for the homogeneity of variances across teachers' autonomy supportive behaviors, results are shown in Table 1.

Table 1

Levene's Test for Homogeneity of Variance Across Autonomy Supportive Learning Behaviors

Variables	Levene's Statistic	df 1	df 2	p
(1) Triggered-Situational Interest	1.909	27	385	.005
(2) Maintained Situational Interest - Feelings	1.645	27	385	.024
(3) Maintained Situational Interest- Values	1.421	27	385	.082
(4) Memory	1.504	27	385	.053
(5) Goal Setting	1.477	27	385	.061
(6) Self evaluation	2.050	27	385	.002
(7) Seeking Assistance	.988	27	385	.484
(8) Environmental Structuring	1.034	27	385	.420
(9) Learning Responsibility	.975	27	385	.502
(10) Organizing	1.156	27	385	.272
(11) Competence 1	2.233	27	385	.001
(12) Competence 2	.961	27	385	.524
(13) Competence 3	2.293	27	385	.000

The obtained Levene's statistics reflected that majority of the variables are non significant ($p > .05$) indicating that the variances of

these variables across autonomy supportive behaviors are homogenous, thus, decision to proceed to the major analysis was warranted.

Table 2

Means, Standard Deviations, Confidence Intervals, Cronbach's alpha of Autonomy Supportive Teaching Behaviors, Situational Interest, Self-regulated Learning Strategies and Mathematics Academic Achievement (N=417)

Variables	M	SD	Minimum	Maximum	Confidence Interval	Cronbach's Alpha
Autonomy Supportive Behaviors	5.39	1.02	2.02	7.0	.013	.983
Triggered-Situational Interest	4.65	.92	1.67	6.0	.002	.813
Maintained Situational Interest- Feelings	4.74	.86	1.50	6.0	.002	.850
Maintained Situational Interest- Values	4.93	.85	1.00	6.0	.002	.811
Memory	2.71	.61	1.40	4.0	.002	.831
Goal Setting	2.90	.61	1.00	4.0	.002	.677
Self Evaluation	2.96	.50	1.67	4.0	.002	.848
Seeking Assistance	2.85	.48	1.50	4.0	.001	.704
Environmental Structuring	3.14	.57	2.00	4.0	.002	.740
Learning Responsibility	3.04	.54	1.40	4.0	.002	.714
Organizing	3.06	.51	1.50	4.0	.002	.725
Competence 1	8.06	2.82	0	10	.008	.813
Competence 2	11.43	4.64	0	20	.014	.852
Competence 3	1.49	1.71	0	5	.005	.840

Table 2 shows that the participants perceived their teachers to be autonomy supportive of their learning. Similarly, they have developed and possessed different forms of situational interest and have utilized to certain levels self-regulated learning strategies. It is also shown that they acquired certain competencies on the three dimensions of math achievement. The standard deviations reflect that there was no abnormal spread of the data.

Reliability estimates for the scales ranged from $\alpha=0.68-0.98$ indicating that items have acceptable to excellent internal consistency.

Truncated estimates for the path coefficients during SEM analysis are avoided (Kline, 2011, p. 123).

Confirmatory factor analysis validated the factors for each latent variables while bivariate correlations between dimensions are found to be linear. The correlations among and between the dimensions or factors of situational interest, self-regulated learning strategies, and mathematics achievement were tested at $p = .01$ and $p = .05$ (2-tailed level of significance). Table 3 showed that the bivariate correlations are positive and are higher between indicators of the same construct supporting convergent validity. The positive correlations between dimensions support linearity and are significant indicating the certainty that regression weights of considerable estimates can be calculated, a realization of another assumption for SEM.

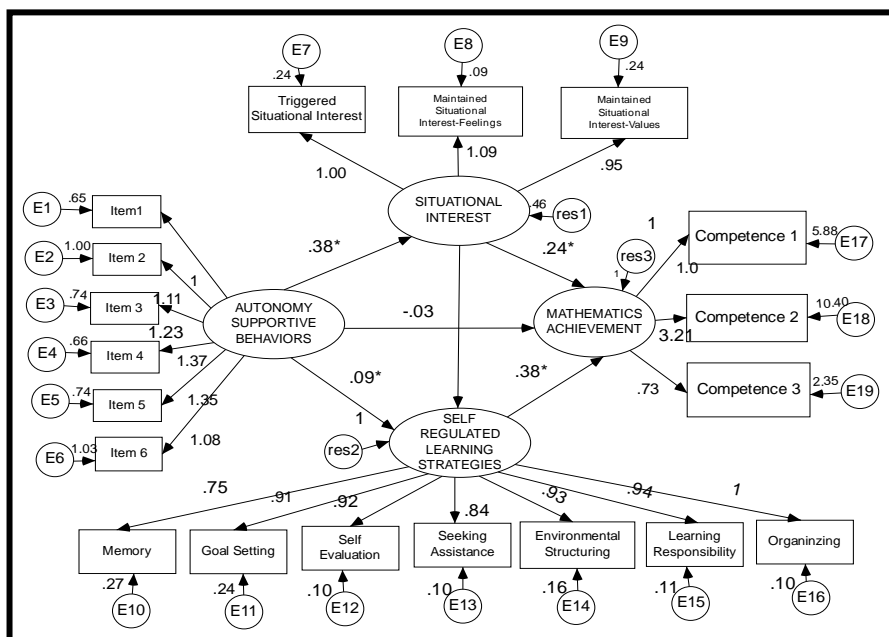


Figure 1. Effects of Autonomy Supportive Behaviors on Mathematics Achievement via Situational Interest and Self Regulated Learning Strategies

In the figure, teachers' autonomy supportive behaviors' direct effect on mathematics achievement (parameter estimate = -.03) is not significant, hence there is no supporting evidence as whether to confirm or not Hypothesis 1. However, the direct effects of the rest of the variables on each other and on math achievement are all positive, significant, and high, indicating that Hypothesis 2 to Hypothesis 6 are supported. Based on these direct effects, results implied that for every 1 unit increase in perceived teachers' autonomy support, there is a corresponding 38% increase in students' situational interest ($p < .001$) and 9% increase in self regulated learning strategies ($p < .05$), respectively. On the other hand, situational interest significantly and directly affected students' SRL strategies (parameter estimate = 0.15, $p < .001$) and Math achievement (parameter estimate = 0.24, $p < .05$), thus supporting Hypothesis 4 and Hypothesis 5, respectively. It implied that for every 1 unit increase in situational interest, a corresponding 15% and 24% increase resulted in self-regulated learning strategies and math achievement, respectively. Lastly, a direct effect of self-regulated learning strategies on math achievement is likewise significant, (estimate = .38, $p < .001$), supporting Hypothesis 6 and indicating that for every 1 unit increase in self regulated learning strategies, a corresponding 38% increase is carried over to math achievement.

The hypothesis on the existence of indirect effects of autonomy on math achievement via situational interest and SRL, as well as the indirect effect of autonomy on SRL via situational interest, and situational interest on math achievement via SRL are all supported by the results. Evidences for indirect causal relationships can be interpreted through the products of indirect effects (Kenny, 2008). Thus, following this statement, as shown in Figure 1, it is implied that during increased levels of situational interest, autonomy supportive behaviors impact positively math achievement (indirect effect, $.38 \times .24 = .091$), supporting Hypothesis 7 and implying further that for every unit increase in autonomy, 9.1% is carried over to math achievement, through situational interest. This indicates that situational interest served as the presumed mediator, which explains the effect of autonomy to math achievement. This also implies that autonomy supportive behaviors impact math achievement through increased situational interest.

On the other hand, the indirect effect of autonomy to math achievement via SRL strategies (indirect effect, $.09 \times .38 = .0342$)

supported Hypothesis 8. The results implied that when students are prompted to use available SRL strategies due to autonomy supportive behaviors, an effect of 3.42% is carried over to math achievement. In addition, as reflected in the model, at high levels of situational interest, autonomy supportive behaviors correspondingly would affect SRL (indirect effect=.057), supporting Hypothesis 9. This indicates that a 5.7% effect is carried over to SRL via situational interest for every unit increase in autonomy support. The findings suggest that the use of SRL strategies can also be significantly improved by autonomy supportive behaviors by increasing the levels of situational interest. In addition, as manifested in Figure 1, situational interest impacts math achievement via SRL (indirect effect= .057), supporting hypothesis 10. This connotes that through increased self-regulated learning strategies, the 5.7% effect of situational interest is carried over to math achievement.

The calculated total effects=.74, presaging that a 74% increase is attained in math achievement by the influence of autonomy supportive behaviors through situational interest and self-regulated learning strategies supporting hypothesis 11. The resulting goodness of fit indexes evaluated that the model is supported by the data. Since absolute fitness of test, $\chi^2 = 376.795, df=147$, or $CMIN/df=2.535$, $p<.000$ is significant, other indices were examined; Root Mean Square Error of Approximation (RMSEA) = 0.061; Tucker Lewis Index (TLI) = 0.925; Comparative Fit Index (CFI) = 0.936 and Incremental Fit Index (IFI)= .937. Majority of these indices are within the acceptable range which indicates that the model fits the sample well. The data also met the required assumptions for the SEM and multivariate analysis, hence a strong argument is made that the model cannot be rejected and is as good as true.

Discussion

The model, which describes how perceived teachers' autonomy supportive behaviors improve mathematics achievement is the present study's major contribution to empirical research. The developed model hypothesized that similar with previous findings, mathematics achievement is predicted by the direct effects of autonomy supportive behaviors, situational interest and self-regulated learning strategies, and emphasized, which is unique in this study, the indirect effects or

associations between these variables can provide more meaningful pathways to explain mathematics achievement.

The tested model explains that the effects of teachers' autonomy supportive behaviors to math achievement was plausibly mediated by situational interest and self regulated learning strategies as shown by the significant magnitudes of these two constructs to math achievement. With their inclusion in the causal process, an indirect effect of autonomy supportive behaviors is carried over to math achievement. This confirmed the seemingly facilitating roles of situational interest and self-regulated learning strategies in transmitting the positive influences of teachers' autonomy supportive behaviors to mathematics achievement whenever they are present. It suggests that when teachers are more autonomy supportive, students' interests are triggered which serves as the "spark" that ignites the learner's curiosity to know more thus engage in all activities (Hidi & Harakiewicz, 2000). When students engage more they experience positive affect, enjoy even more in difficult tasks and see the tasks as meaningful and valuable. The more that they interact with other learners and with the materials during activities, there is hands on, hearts on, and minds on, characterizing a learner centered curriculum (Ocampo, 2008). In this context, to sustain their satisfaction, learners are led to self regulate their actions. With the continuous and increased support from the outside source, which in this situation, the teachers' autonomy supportive behaviors, students maintain the feelings of interestingness in the 3 dimensions, resulting in more use of self regulated learning strategies. The learners acquired more tools in the form of SRL strategies which aid them in learning; consequently, they are able to achieve the competencies required of them. This context finally led to higher achievement in math (Harakiewicz, Barron, Durik, Tauer, & Linnenbrink-Garcia, 2008). Although the direct effect of teachers' autonomy supportive behaviors which leads to better mathematics achievement (Tsai et al, 2006; Assor, Kaplan, & Roth, 2002; and Wilsons & Simons, 2002) was not provided with evidence for confirmation in this study, however, how teachers' autonomy affects math achievement is clearly manifested. These outcomes confirmed the similar influences to math achievement, obtained in the study conducted by Hidi & Harakiewicz (2000), Schraw, Flowerday, and Lehman (2001) and Dohn, Madsen, and Malte (2009). On the other hand, when students utilize SRL often as the result of interestingness

of the learning context, they are aided in facing and overcoming challenges in math instruction, enabling them to analyze, process, comprehend, do mathematical computations with accuracy thereby attaining the competencies expected of them, and finally improve mathematics achievement. These results were also found in the studies conducted by Soric' and Palekcic' (2009), Camahalan (2006), and Arzal (2011). However, it should be reiterated that positive effects of situational interest and self regulated learning strategies on math achievement become significant only when the learning context is autonomy supportive. These triadic relationships which are not yet explicated well in the research arena affords valuable findings that support self-determination theory (Deci & Ryan, 2000), stating that positive outcomes are possible when necessary social nutriment from the outside are provided. Teachers' autonomy supportive behaviors supply the necessary "nutriments" needed for the development and improvement of situational interest making learners to self-regulate their learning strategies, become self determined and consequently attain higher mathematics achievement. These behaviors provide the necessary encouragements for learners to overcome their fears, be one in the group, and/or recognize their potentials which lead to more development of situational interest and adoption of more self-regulated learning strategies. This explains why the urgent need to support autonomy of our learners.

Teachers' autonomy supportive behaviors effect on math achievement is explained best by the development of situational interest and SRL strategies among learners. Findings of this study fit well the Self Determination Theory (Deci & Ryan, 1985; 1991; 2010) and has wider range of application in the field of education.

References

- Adams, H. (2004). *Justice for children: The development of autonomy*. (Unpublished doctoral dissertation). Rice University, USA.
- Arzal, S. (2011). The impact of self-regulation instruction on mathematics achievements and attitudes of elementary school students. Retrieved from egitimvebilim.ted.org.tr/index.php/EB/article/download/4/2
- Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing

- teacher behaviors predicting students' engagement in schoolwork. *British Journal of Educational Psychology*, 72, 261–278.
- Boekaerts, M., & Minnaert, A. (2006). Affective and Motivational Outcomes of Working in Collaborative Groups. *Educational Psychology*, 26(2), 187–2008.
- Camahalan, F. (2006). Effects of self-regulated learning on mathematics achievement of selected Southeast Asian children. *Journal of Instructional Psychology*, 33(3), 194–205.
- Cleary, T., & Chen, P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, 47(5), 291–314.
- Deci, E. L., & Ryan, R. M. (2000). The “what ” and “why of goal pursuits: Human need and the self-determination of behavior. *Psychological Inquiry*, 11, 227–268.
- De Corte, E., Verschaffel, L., & Depaepe, F., (2008). Unraveling the relationship between students' mathematics-related beliefs and the classroom culture. *European Psychologist*, 13(1), 24–36. DOI 10.1027/1016-9040.13.1.24
- DepEd Memo # 49, s. 2002. Operationalization of the DepEd and JICA Cooperation on the Strengthening of Continuing School-Based INSET for Science and Mathematics Teachers.
- Dohn, N., Madsen, P., & Malte, H. (2009). The situational interest of undergraduate students in zoophysiology. *Advances in Physiology Education*, 3, 196–201.
- Grinsven, L., & Tillema, H. (2006). Learning opportunities to support student self-regulation: comparing different instructional formats. *Educational Research*, 48, 77–91
- Grolnick, W. S., Ryan, R. M., & Deci, E. L. (1991). The inner resources for school achievement: Motivational mediators of children's perceptions of their parents. *Journal of Educational Psychology*, 83, 508–517.
- Harakiewicz, J., Barron, K., Durik, A., Tauer, J., Linnenbrink-Garcia, L. (2008). The role of achievement goals in the development of interest: reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology*, 100(1), 105–122.
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60, 549–571.

- Hidi, S., & Harackiewicz, J. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, 70, 151–179.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41, 111–127.
- Linnenbrink-Garcia, L., Durik, A., Conley, A., Barron, K., Yauer, J., Karabenick, S., & Harackiewicz, J. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*, 20, 1-25. DOI: 0.1177/0013164409355699
- Magno, C. (2009). Developing and assessing self-regulated learners. *The Assessment Handbook*, 1, 26-45.
- Magno, C. (2010a). Assessing academic self-regulated learning among Filipino college students: The factor structure and item fit. *International Journal of Educational and Psychological Assessment*, 5, 61-76.
- Magno, C. (2011). The predictive validity of the academic self-regulated learning scale. *The International Journal of Educational and Psychological Assessment*, 9(1), 48-56.
- Magno, C. (2011b). Validating the academic self-regulated learning scale with the Motivated Strategies for Learning Questionnaire (MSLQ) and Learning and Study Strategies Inventory (LASSI). *The International Journal of Educational and Psychological Assessment*, 7(2), 56-77.
- Magno, C., & Ouano, J. (2010). *Designing written assessment of student learning*. QC, Manila: Phoenix.
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85, 424–436.
- National Education Testing and Research Center. (2012). NAT Overview and 2012 Test Results. Retrieved from <http://philbasiceducation.blogspot.com/2013/07/the-national-achievement-test-in.html#ixzz3GmUewBhp>
- Ocampo, C. S. (2008). Country papers: School Based Training Program (SBTP) for Science and Mathematics Teachers: The Philippines Experience. Retrieved from <http://hrd.apec.org/images/d/d0/45.19.pdf>
- Pierro, A., Presaghi, F., Higgins, T. E., & Kruglanski, A.W. (2009). Regulatory mode preferences for autonomy supporting versus

- controlling instructional styles. *British Journal of Psychology*, 79, 599-617.
- Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98, 209-218.
- Ryan, R. M., & Grolnick, W. S. (1986). Origins and pawns in the classroom: Self-report and projective assessments of individual differences in children's perceptions. *Journal of Personality and Social Psychology*, 50, 550-558.
- Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13, 23-52.
- Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., & Dochy, F. (2009). The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology*, 79, 57-68.
- Sorić, I., & Palekčić, M. (2009). The role of students' interests in self-regulated learning: The relationship between students' interests, learning strategies and causal attributions. *European Journal of Psychology of Education*, 24, 545-565.
- Stefanou, C. R., Perencevich, K. C., DiCintio, M., & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist*, 39, 97-110.
- Tsai, Y. M., Kunter, M., Lu'dtke, O., Trautwein, U., & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology*, 100(2), 460-472.
- Turner, J. C., Meyer, D. K., Cox, K. E., Logan, C., DiCintio, M., & Thomas, C. T. (1998). Creating contexts for involvement in mathematics. *Journal of Educational Psychology*, 90, 730-745.
- Ulep, S. A. (2006). Good Mathematics Teaching Practices - In The Making: A Philippine Experience. National Institute for Science and Mathematics Education Development. University of the Philippines. Retrieved from http://www.criced.tsukuba.ac.jp/math/apec2006/progress_refort/Specialist_Session/Ulep.pdf.
- Wilson, A., & Simons, R. (2002). University Students' Perceptions of the Learning Environment and Academic Outcomes:

- Implications for theory and practice. *Studies in Higher Education*, 27(1), 27-52.
- Zimmerman, B. J. (1989). A social cognitive view of self-regulated-academic learning. *Journal of Educational Psychology*, 81, 329-339.
- Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614-628.



Exploring the Predictors and Outcomes of Academic Resilience among College Students

Analyn O. Leysa

Filmor A. Malnegro

Ateneo de Davao University

Abstract

This study examines Martin's motivation and engagement framework (2002) on academic resilience and its effects on school enjoyment, class participation, and general self-esteem. There were 487 College students enrolled in General Psychology classes who were asked to answer a 22-item questionnaire based on the specifications of Martin's model (2002). Path analysis was used to determine how well the model fits the data obtained from the sample. Results indicate that the model is not a good fit for the sample of the study. In the context of the sample at hand, Low Anxiety decreases the Academic Resilience of the learners and Control is not a significant predictor of Academic Resilience. This implies that students tend to be more resilient when they worry more about school and their fear of failure. Moreover, the extent of perceived Control they have over their learning experiences does not signify the level of their resilience when it comes to schoolwork. Consistent with literature, academic resilience predicts the desirable educational outcomes of enjoyment in school, class participation, and general self-esteem.

Keywords: academic resilience and its effects on school enjoyment, class participation, and general self-esteem

Introduction

Generally, resilience is defined as the process of, capacity for, or outcome of successful adaptation despite threatening or challenging circumstances (Howard & Johnson, 2000). In the academic context, it

is “the likelihood of success in school and other life accomplishments despite environmental adversities brought about by early traits, conditions and experiences” (Wang, Haertel, & Walberg, 1994). In the same manner, those “who sustain high levels of achievement motivation and performance despite the presence of stressful events and conditions that place them at risk of doing poorly in school” (Alva, 1991), are considered to be academically resilient. There is still much to learn from students who manage to perform in school despite the difficult circumstances that they find themselves in. To a large extent, literature on resilience has established associations between individuals who are at a disadvantage or exposed to various familial and environmental stressors like divorce or drug addiction. Thus, it is of value that an investigation on academic resilience in the school context be done because students also experience difficulties, pressure, or low performance. School is an important place where resilience in young people can be enhanced, benefitting them in the long run (as cited in Martin & Marsh, 2003).

Academic resilience research is the study of high educational achievement despite the presence of risk factors that normally indicate low academic performance. It can be understood as the process and results that are part of the life story of an individual who has been academically successful despite obstacles that prevent the majority of others with the same background from succeeding (Masten & Coatsworth, 2008). Most of studies on academic resilience established its links to different domains that range from academic and educational constructs to psychological factors, socio-demographic as well as family and peer-group characteristics (Finn & Rock, 1997). As a result of studying the construct across a variety of domains, it was difficult to give sufficient detailed attention to any of the domains. In the psychological domain, studies are typically limited to focusing on just self-esteem and locus of control. Additionally, studies on academic resilience are predominantly focused on the mental health and wellbeing of the learner and not in terms of academic development. Martin (2001, 2002, 2003a, 2003b) developed an expansive model of psychological and behavioral engagement known as the Student Motivation and Engagement Wheel which reflects the thoughts, feelings, and behaviors underpinning academic achievement in school and used confirmatory factor analysis to determine which aspects

correspond to academic resilience of the learners (Martin & Marsh, 2006).

The model at hand separates motivation into factors that reflect enhanced motivation (adaptive) and those that reflect reduced motivation (maladaptive). As discussed extensively by Martin (2002, 2002, 2003b), adaptive dimensions include self-efficacy, valuing of school, mastery orientation, persistence, planning and study management. On the other hand, the maladaptive dimensions comprise anxiety, uncertain control, failure avoidance and self-handicapping. Self-efficacy is students' beliefs and confidence in their ability to understand or to do well in their school work, to meet challenges they face, and to perform to the best of their ability. Students are certain in Control when they are sure about how to do well or how to avoid doing poorly in school tasks. Planning is how much students plan their schoolwork, assignments and study and how much they keep track of their progress as they are doing them. Low anxiety involves feeling nervous and worrying about not doing very well in their schoolwork, assignments or exams. How much students keep trying to work out an answer or to understand a problem even when that problem is difficult or is challenging describes Persistence. The Wheel is proposed as an integrative and encompassing way to understand the diversity of psychological engagement dimensions that underpin academic resilience. This model was used in this study to determine its effect on academic resilience.

A host of important educational and psychological constructs that are conceptually relevant to academic resilience should be also carefully looked into. Martin and Marsh (2006) proposed a three between-network constructs that provide a theoretically relevant basis for further examination: class participation and enjoyment of school (educational "outcome" construct) and general self-esteem (a psychological "outcome" construct). Class participation is an important between-network, school-related behavioral "outcome" measure that is found to enhance students' commitment to learning. Enjoyment of school is a school-related cognitive-affective "outcome" that shapes students' willingness to attend school as well as the goals they have in relation to their academic experience while at school and in their further education and training beyond school. Students' general feelings about themselves because of their school experience refers to the between-network "outcome" construct General Self-Esteem.

These three represent a breadth of students' experience at school and are hypothesized to follow from students' capacity to deal effectively with challenge and adversity in the school setting. This model further hypothesized that students who do not deal effectively with such challenges are less inclined to participate in class, less likely to enjoy school and more likely to experience general negative affect in relation to the self (Martin & Marsh, 2006).

Academic resilience has been linked to various factors within a range of domains from academic to family, peer, socio-demographic and psychological. In conducting research across domains, there is typically less detailed attention specifically given to the psychological domain, where the common focus is on just self-esteem and locus of control (Martin & Marsh, 2006). Previous research aimed at examining a diverse array of psychological and educational dimensions. Martin and Marsh (2006) forwarded a direction of exploring class and school level climates relevant to individual-level variation in academic resilience.

Hence, the aim of this study is to look at Martin's motivation framework, whether Self-efficacy, Control, Planning, Low Anxiety and Persistence predicts Academic Resilience. It is also hypothesized that academic resilience will predict school enjoyment, class participation and general self-esteem contextualize in a school General Psychology class level.

Method

Participants

Respondents of the study were 487 college students enrolled and have taken a General Psychology class for the current semester. Teachers administered the instrument to students during their respective classes. Students were asked to complete the instrument on their own and return the completed instrument to the teacher. A sample size of 487 with a desired large effect size also yielded a large statistical power of 1.0000 ($p = 0.05$).

Materials

A 22-item questionnaire was administered to students in one testing session in their class. The instrument is composed of randomly

arranged items that measures Self-efficacy (*If I try hard, I believe I can do my school work well*), Control (*I'm often unsure how I can avoid doing poorly at school*), Planning (*Before I start an assignment I plan out how I am going to do it*), Low Anxiety (*When exams and assignments are coming up, I worry a lot*), Persistence (*If I can't understand my schoolwork at first, I keep going over it until I understand it*), Enjoyment of School (*I enjoy being a student*), Participation in School (*I get involved in things we do in class*), General Self-esteem (*Overall, I have a lot to be proud of*) and Academic Resilience (*I think I am good at dealing with school work pressure*). These items were adopted to closely resemble the Student Motivation and Engagement Scale (SMES) and Academic Resilience Scale (ARS) utilized by Martin and Marsh (2006). Students rated themselves on a scale of 1 (*Not True of Me*) to 7 (*Extremely True of Me*) for all items.

Except for the scale for Academic Resilience, which had six items, all the other scales were composed of two items each. The scale for Self-efficacy obtained a mean inter-item correlation of .47 and an alpha coefficient of .64. For Control, the scale had an inter-item correlation of .35 and an alpha coefficient of .52. For the scales for Planning, Low Anxiety, and Persistence, the values for mean inter-item correlation and alpha coefficient are .45 and .62, .16 and .28, .46 and .63, respectively. For the outcome variables involved in the model, the scale for Enjoyment of School had an inter-item correlation of .42 and an alpha of .59. The items for Class Participation obtained an inter-item correlation of .48 and an alpha of .65. For the General Self-esteem scale, inter-item correlation is at .30 and alpha is .44. The Academic Resilience scale has acceptable fit values (CFI = 0.97; NNFI = 0.97), its total item correlation ranged from 0.59-0.78 and has an alpha coefficient of .89.

Data Analysis

Pearson r was used to establish the correlations between the factors and path analysis was used to determine if the proposed model showing the effect of Self-efficacy, Control, Planning, Low Anxiety, and Persistence on Academic Resilience and the effect of Academic Resilience on students' Enjoyment of School, Class Participation, and General Self-esteem, is a good fit for the data obtained from the sample of this study.

To determine how well the data corresponds with the model, model fitting was done using fit indices. The main absolute fit index is the chi-square (CMIN), which tests for the level of misspecification. A significant CMIN suggests that the model does not fit the data. Then again, CMIN has been found to be too sensitive to sample size that is why it cannot be used as the sole indicator of model fit (Teo, Tsai, & Yang, 2013). Other absolute fit indices commonly used include Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), and the Root Mean Square Error of Approximation (RMSEA). GFI evaluates the relative amount of observed variances and covariances that can be explained by the model. AGFI takes into account the level of model complexity by considering degrees of freedom. For both GFI and AGFI, the value should be greater than .90. RMSEA counters the tendency of CMIN to reject models with large samples and many variables by using the residuals as an indicator of the accuracy of the model. A lower RMSEA value (<.05) suggests that the model is a good fit for the sample. The Comparative Fit Index (CFI) assesses whether the hypothesized model is better than a baseline model, which assumes that all observed variables are not correlated. A value that is greater than .90 for CFI is associated with a good model. Another index that compares the proposed model to the baseline model is the Tucker-Lewis Index (TLI). The TLI is not normed. For this reason, its value can fall below zero and above one, but typically, good models have TLI values that approach 1.0.

Results

Table 1
Means and Standard Deviations of the Variables

Variables	<i>M</i>	<i>SD</i>
Self-efficacy	2.68	1.35
Control	3.84	1.12
Planning	3.50	1.26
Low anxiety	3.60	1.24
Persistence	2.89	1.26
Academic Resilience	3.54	0.93
Enjoyment of School	2.97	1.32
Class Participation	3.38	1.17
General Self-esteem	2.90	1.28

Table 1 shows the mean scores and corresponding standard deviations reported by the students for each variable involved in the proposed model. For all the variables, the mean scores fall at the lower end of the scale, indicating that the respondents reported scores that are below the moderate level. In comparison to the means obtained by Martin and Marsh (2006), the values reported by the sample of this study were generally lower.

Table 2
Intercorrelations of the Factors

Variables	1	2	3	4	5	6	7	8	9
1 Self-efficacy	---								
2 Control	.12*	---							
3 Planning	.51*	.21*	---						
4 Low anxiety	-.34*	.42*	.26*	---					
5 Persistence	.71*	.15*	.64*	.33*	---				
6 Academic Resilience	.56*	.10*	.49*	.11*	.60*	---			
7 Enjoyment of School	.61*	.17*	.48*	.31*	.67*	.57*	---		
8 Class Participation	.50*	.16*	.45*	.16*	.60*	.59*	.58*	---	
9 General Self-esteem	.66*	.22*	.46*	.32*	.62*	.52*	.69*	.56*	---

* $p < .05$, ** $p < .01$

Table 2 shows the zero-order bivariate correlations which determine the pairs of variables that have significant relationships. Although varying in strength, analysis suggested that all bivariate relationship among the variables are positive and significant. This implies that when one variable increases, the other variables also increase in different degrees. Considering the relationships of the proposed predictors of Academic Resilience, Control and Low Anxiety both have a weak correlation, although it is statistically significant at 95% confidence level. Moreover, a pattern arises in the relationships of Control to the other variables in the model. All of the bivariate relationships involving Control are marginal in strength, except for the correlation with Low Anxiety, which is moderate.

Path analysis was conducted to test a hypothesized model showing the effect of Self-efficacy, Control, Planning, Low Anxiety, and Persistence on Academic Resilience, as well as the effect of

Academic Resilience on Enjoyment of School, Class Participation, and General Self-Esteem. The model is tested for goodness of fit using the chi-square (CMIN), Goodness-of-Fit Index (GFI), Adjusted GFI, and Root Mean Square Error Approximation (RMSEA). The indices showed that the proposed model is a bad fit for the data obtained from the sample, as revealed by a high chi-square value that is significant (CMIN = 1327.44, $p < .001$). The GFI (.55) and AGFI (.27) were both low and did not satisfy the required values for these measures of good fit. The RMSEA (.31) is huge relative to the value of .05, which is required for good fit. This further emphasizes that the model is a bad fit for the sample. Other fit indices also did not support the fit of the model like the CFI (.39) and the NFI (.38).

Although the zero-order relationships of the proposed predictors with Academic Resilience were all statistically significant, one regression path involved in the model is not significant. The regression weight for Control in predicting Academic Resilience is not significantly different from zero at the .05 level. It can also be noted that there is a shift in the direction of the relationship of Low Anxiety and Academic Resilience during path analysis. Results indicate that Low Anxiety decreases Academic Resilience by 0.13. Furthermore, Academic Resilience significantly predicts all of the three outcome variables in the model.

Discussion

The present study sought to examine the effect of self-efficacy, control, planning, low anxiety and persistence on academic resilience. The data demonstrated that the identified predictors significantly predicted academic resilience. Path analysis also showed that academic resilience subsequently predicts three educational and psychological outcomes over and above the motivation and engagement factors supporting academic resilience. This result holds up to the findings of Martin and Marsh (2006) that specific interventions that aims to increase students' academic resilience should revolve around students' self-efficacy, control, persistence, planning, and anxiety. The ability to do school work, the degree of control one has to avoid doing poorly in school tasks, to keep track of one's progress and to persist in understanding a difficult problem or task influences students' ability to

maintain high levels of performance despite stressful conditions that will place them at risk of doing poorly in school.

Self-efficacy is the belief of the student in his or her own academic capacity (Bandura, 1997). Learners with a higher sense of self-efficacy are generally more resilient because they have more confidence in their ability to meet the challenges and perform to the fullest of their potential. In addition, Self-efficacy also influences how learners set the goals that they want to achieve and how they commit themselves to the tedious process of achieving these goals. Planning and Persistence are two other key dimensions of Academic Resilience. Effective planning improves resilience in the context of school because it provides learners with essential details on how they can work on the things that they aim to accomplish. Having a roadmap to serve as a guide allows for better management of time, effort, and other resources even in the midst of the many challenges in school. Persistence is the extent that learners continue to keep trying even when the problem or task is considerably difficult. Persistent students keep going in spite of minor or major setbacks until they achieve the goals that they have set, making them more resilient and generally more successful in school.

Low Anxiety is the extent of being calm and comfortable in school and with schoolwork. Theoretically, student who experience low anxiety generally report high academic resilience. Contrary to the findings of Martin and Marsh (2006), results of this study indicate that low anxiety decreases academic resilience. This suggests that in the context of this sample, learners increase their resilience when they worry more about school and schoolwork. In contrast to existing literature, the fear of failure, exams, and grades turns out to be an important reason for the respondents of this study to be work harder and become more resilient. Similarly, Church and Katigbak (1992) found that Filipino college students tend to persist more after failure than after success. That is, Filipino students perceive the thought of failure as a source of motivation rather than a barrier to academic success. Moreover, this echoes the findings of Jowkar et. al. (2013) when predicting the academic resilience of Iranian high school students from goal orientation and perceptions on failure.

Control is the extent of being sure about how to do well in school. In theory, students who feel a greater sense of control over their learning experiences are more resilient. Then again, the findings

of this study indicate that control is not a significant predictor of the academic resilience of the sample. That is, it cannot be guaranteed that the level of control that the respondents perceive influences the level of their resilience in school. As stated by Perry, Hall, & Ruthig (2005), there are existing dispositional differences in the notion of control among students. It is possible for students to view that, at times, control is unimportant. The value that they place in having control over the process of learning affects the extent that control influences various desirable outcomes including sense of worth, efficacy, competence, performance and possibly, resilience. Generally, it is assumed that students want to control their own educational experiences. Then again, it does not necessarily happen all the time. Instances in which this is not the case, as revealed in this study, might be indicative of certain concerns that will have to be addressed including factors like quality of instruction, teacher effectiveness, grading standards, course level, curriculum structure, classroom discipline, class composition, and others (Perry, 2003).

Consistent with literature, academic resilience, to a large extent, predicts students' enjoyment of school, class participation, and general self-esteem. Results exhibit that improving the academic resilience of learners can lead to significant developments involving the desirable outcomes in the context of academic environments. Resilient students tend to enjoy being in school and doing schoolwork more than those who are less resilient. They are able to appreciate the tasks involved in school and the challenges that arise in the process of learning. Fostering academic resilience in learners provides them with a mindset that allows them to be more comfortable and at-ease in the midst of the learning environment despite difficulty. Academic resilience also leads to improved class participation and engagement. Resilient learners completely understand that participating in classroom discussions and school activities is an important part of the learning process that is why they continue to engage even if various barriers and risks are involved. Also, academic resilience affects the general self-esteem of the learners. Students who are more resilient feel more confident in their capacity of accomplish the goals that they have set for themselves, which in turn helps them develop a healthy sense of pride when it comes to their learning experiences.

The results of the study at hand reveal that the model hypothesized by Martin & Marsh (2006) on academic resilience is not a

good fit for the data obtained from the sample. Analysis of regression paths show that there are points which do not corroborate with the existing literature on the proposed model. First, it would be interesting to note that in the context of the sample of this study, low anxiety decreases academic resilience, which indicates that the resilience of these learners increases when they worry about school and schoolwork more. This supports the findings of studies that suggest students can have different paradigms and corresponding reactions to the thought of academic failure. Another point worth noting is that control does not significantly contribute to the academic resilience of the respondents. This can be attributed to dispositional differences that influence how student value being in control of their educational environment and learning experiences. In general, having a good sense of control is beneficial for the learning process as well as for the students themselves. As such, other concerns experienced by the respondents will have to be addressed. Lastly, consistent with existing literature, academic resilience significantly predicts the desirable educational outcomes included in the model. This clearly shows that enhancing the academic resilience of students can lead to significant improvements that would allow them to thrive in school and achieve goals which are integral to their development as learners.

It must be recognized though that this study was only able to utilize a few indicators for the identified predictors. Likewise, it was limited to college students enrolled in a general education subject. Thus, for future research on this construct, there is a need to include more items for the indicators that will measure the factors that focus on the family, the school climate, and the immediate community to integrate into the model aspects which are beyond the individual learners, as well as involving students taking up other subjects, specifically professional/core subjects relative to their degree programs.

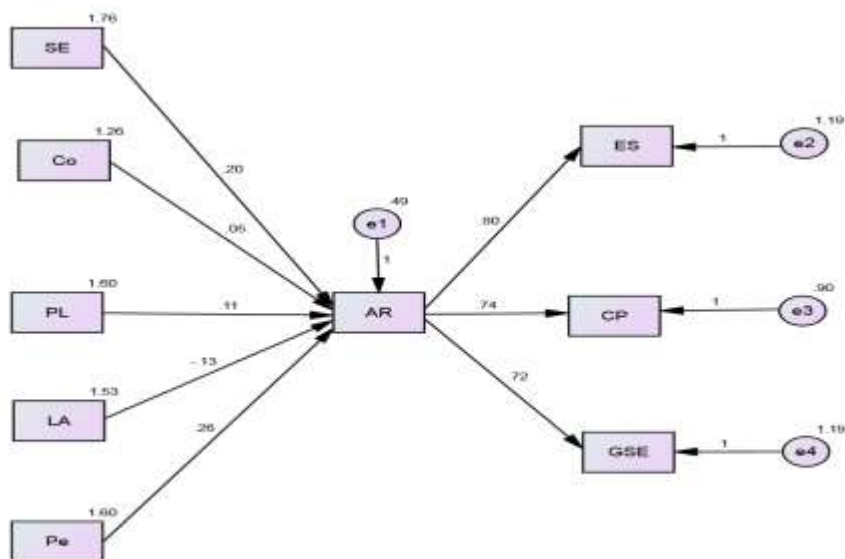


Figure 1. Path coefficients for motivation and engagement factors on academic resilience and academic resilience on educational and psychological outcomes.

References

- Alva, S.A. (1991). Academic invulnerability among Mexican-American students: The importance of protective and resources and appraisals. *Hispanic Journal of Behavioral Sciences*, 13, 18-34.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman & Co.
- Jowkar, B., Kojuri, J., Kohoulat, N., & Hayat, A. (2013). Academic Resilience in Education: The Role of Achievement Goal Orientations. *Journal of Advances in Medical Education*, 2(1), 33-38.
- Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools*, 43, 267-281.
- Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favourable and unfavourable environments:

- Lessons from research on successful children. *American Psychologist*, 53, 205-220.
- Sandoval-Hernandez, A. & Cortes, D. (2015). Factors and conditions that promote academic resilience: A cross-country perspective. *International Congress for School Effectiveness Improvement*, 25.
- Sinay, E. (2009). Academic resilience: Students beating the odds. *Organizational Development/Research and Information Services*, 5(1), 416-394.
- Solberg, V. S., Davis, A., & McLemore, C. (2010). *Resiliency as an indicator of academic success*. NJ: ScholarCentric.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1994). *Educational resilience: An emergent construct*. Hillsdale, NJ: Erlbaum.



Understanding the Structure of Autonomy among Filipino Adolescents

John Jamir Benzon R. Aruta

De La Salle University, Manila

Abstract

This study examined the structure of autonomy among Filipino adolescents through the framework proposed by Markus and Kitayama (1991b). Screening through median split was used to identify participants with high level of interdependence. The first set of participants ($n=21$) responded in the qualitative phase while the second set of participants ($n=201$) participated in the quantitative phase. After performing exploratory factor analysis (EFA), two dimensions of autonomy emerged—inward autonomy *and outward autonomy*. The results provide evidence about the unique definition of autonomy in an interdependent culture like the Philippines. Implications to assessment and student motivation are discussed.

Keywords: Self-Determination Theory, Interdependence, Autonomy, Filipino Adolescents

Introduction

Self Determination Theory or SDT (Deci & Ryan, 1985; Ryan & Deci, 2000) described autonomy (deCharms, 1968; Deci, 1975) as a motivational state that produces behaviors that are self-endorsed and willingly enacted. People are described to be most autonomous when an action is based on their own decisions, genuine interests, values, and desires. Furthermore, it had been studied in many perspectives and had been linked primarily with well-being (e. g., Chirkov, Ryan, Kim, & Kaplan, 2003; Reis, Sheldon, Gable, Roscoe, & Ryan, 2000). In addition, Ryan and Deci (2000) proposed autonomy as one of the

three basic psychological needs that, when satisfied, is assumed to contribute to overall well-being.

Autonomy has been challenged by many theories. For example, Wegner (2002) views that behavior can be displayed non-consciously, and therefore emphasized that human will is an illusion. Other theorists forwarded another perspective when they criticized the idea that autonomy is gender and culture bound and insisted that autonomy is a cultural element rather than a process of behavioral regulation (e.g., Iyengar & DeVoe, 2003; Jordan, 1991). Despite these divergences in theoretical perspectives, autonomy remained to be a fundamental human need and was claimed to be universal and a cross-developmental need (Ryan & Deci, 2006). In SDT perspective, autonomy is characterized as a state that can vary in degree. Ryan and Deci (2000) emphasized that it is a regulatory style continuum that ranges from controlled regulation (heteronomy) to true self-regulation (autonomy). SDT classifies people's motivation into different forms such as being externally regulated, introjected, identified, and integrated. Individuals with externally regulated form of motivation are considered the most heteronomous. People with introjected kind of motivation, on the other hand, displays partial assimilation of external influences. Moreover, individuals who internalize a personal valuing of actions are described as having an identified form of motivation. Those with well-synthesized values and beliefs are characterized to have an integrated form of motivation. (Ryan & Deci, 2000). The last form of regulation, intrinsic motivation, is considered highly autonomous, as it is independent and is fueled by interest in the behavior (Ryan & Deci, 2006). Although autonomy has been formulated as a universal need, one cannot assume that it operates the same way when observed among people who live in an interdependent social system. Interdependent cultures like the Philippines value interpersonal relationships as a primary cultural goal. Therefore, the motivation and decision to display any behavior while navigating in this kind of culture would give elevated consideration on interpersonal perception and consequences. Enriquez (1977) explained that Filipinos have higher sensitivity to other people and makes decision in consideration of others.

Markus and Kitayama (1991a, 1991b) argued that people with independent (individualistic) and interdependent self-construal (collectivistic) perceive the self in different ways. People with

independent self-construal view the self as unique individual with unique set of characteristics and was described as having lower attention to context. On the other hand, people with interdependent self-construal view the self as interconnected with others. They are described to be more sensitive with the context and tend to give more value to harmony. For this reason, cultural psychologists suggested a review of psychological processes and constructs that had been studied in the past because most of these constructs and processes are believed to behave differently when observed in a cultural perspective (e.g., Na & Kitayama, 2011; Kitayama, Duffy, Kawamura, & Larsen, 2003). Although autonomy had been studied already in a cultural perspective (e.g., Chirkov, et al., 2003), many of these studies investigated it using western formulations and instruments. Thus, studying the structure of autonomy in an interdependent culture such as the Philippines is an important research direction. The present study assumes that autonomy is shaped differently in an interdependent culture. That is, people do not only endorse behaviors that are self-oriented but also behaviors that are other-oriented or those that concern the benefit of other people.

Method

Participants

Twenty one participants (Mean Age = 17.86 years; Male = 11, Female = 10) in the qualitative phase were chosen after performing median split (Median = 5.68). The 21 participants were deduced from 45 individuals. They were chosen on the basis of their scores that are higher than the median, suggesting high level of interdependence. Moreover, 201 participants (Mean Age = 17.74; Male = 52, Female = 149) in the quantitative phase were selected as participants in the study. They were chosen from the original 405 individuals after performing median split (Median = 5.36). The participants were tertiary students from two private and one public university in Manila, Philippines. They came from different year levels and were selected through convenience sampling.

Instruments

The 11-item Relational-Interdependence Self-Constraint (RISC) scale by Cross, Bacon and Morris (2000) was used to screen the participants with high level of interdependence. The researcher evaluated the 7-point scale RISC in order to ascertain that the Filipino sample can understand the items (e.g., *My close relationships are an important reflection of who I am.*). The participants answered the instrument by rating each item from 1 = strongly disagree to 7 = strongly agree. The obtained Cronbach's alpha coefficient of 0.83 suggests that the RISC items have relatively high internal consistency.

Situational sampling was used to gather qualitative data that would describe autonomy among the participants. It is a method typically used in cultural studies. This instrument was used in the present study by instructing the participants to read the definition of autonomy provided in the questionnaire then they were asked to list down at least five situations where autonomy is being displayed. After the coding analysis of the qualitative data, two major categories of behaviors were generated: *inward and outward autonomy*.

In the quantitative phase, the researcher developed a 19-item, 5-point scale instrument to measure autonomy of the participants in the quantitative phase. The 19 items were based on the two major categories that were generated from the qualitative responses in the situational sampling. 10 items were developed under ego-oriented autonomy (e.g., *I make all my decisions for my own welfare.*) and 9 items were developed under other-oriented autonomy (e.g., *I do things willingly for other people.*). The participants answered each item by rating from 1 = not true to me to 5 = very true to me.

Procedure

The study involved three phases—qualitative, item development, and the quantitative phase, all involving a participant screening process. In the qualitative phase, 45 participants were asked to answer the RISC and the situational sampling questionnaire. The RISC was used to determine the participants with higher level of interdependence through performing median split. 21 out of the 45 participants in this phase were selected for the analysis of situational sampling (qualitative) data. The qualitative method was used to gather

data about autonomy. During the administration of the two instruments, the participants were reminded that there are no right or wrong answers and that it is essential to respond with utmost honesty.

After analyzing the qualitative data, two major categories such as *inward* and *outward autonomy* were identified. Next, the researcher developed items based on the two major categories derived from the qualitative phase. The researcher checked the items to determine if they truly represent the two identified dimensions of autonomy. A new set of participants was selected to participate in the quantitative phase. Pretesting of online and face-to-face administration was done and the participants responded that both methods were almost equally effective in getting genuine responses. Two to five minutes was spent by the participants in answering the RISC and the autonomy questionnaire combined. Data on autonomy were gathered both through online and face to face administration. The same with the first face, RISC was also administered first to screen the individuals with high level of interdependence by performing median split. During the administration, an informed consent was provided to the participants. They were oriented about their voluntary participation and the nature of the study. Participants' data were analyzed using exploratory factor analysis through SPSS Version 20.

Results

Qualitative Results

The results show the experience of autonomy of the participants based on their listings of situations and behaviors where they fully endorse a behavior. Results from 21 participants indicate that the experience of autonomy or fully endorsing one's behavior revolves around two distinct dimensions: *inward* and *outward autonomy*. Table 1 shows the two dimensions represented by the categories that emerged after the analysis. It also shows definitions and sample behaviors per dimension. The first dimension is termed as *inward autonomy*. It is characterized by decisively endorsing behaviors with a motivation to gain personal benefit. It was found out that participants tend to display behaviors that are driven to have a sense of personal fulfillment and expression. However, the participants also reported fully endorsed behaviors that are interpersonal in nature or *outward autonomy*. It is

defined as a form of autonomy that prioritizes others or gives higher consideration to the welfare of other people. It must be noted that while the participants experience self-oriented autonomy, most of the listings of behaviors that they provided were other-oriented in nature. This suggests that since the participants came from an interdependent culture, which prioritizes interpersonal relationships as a cultural norm, it should not be surprising to observe behaviors that gear towards the benefit of other and yet be fully endorsed.

Table 1
Generated Categories, Definitions, and Sample Behaviors

<i>Generated Categories</i>	<i>Category Definition</i>	<i>Sample Behavior Listings</i>
Self-Oriented Autonomy	<i>Self-oriented autonomy</i> is characterized by decisively endorsing behaviors with the motivation to gain personal benefit.	When I joined the glee club in high school to have a sense of personal achievement and expression I want to do good so that I can live up to my ideals I learned more about football to be knowledgeable and to feel fulfilled about my lack of ability I study harder to be a great game developer so I can prove to my self that I can do it I independently choose the degree program I am taking now I enrolled myself in a violin class because I want to learn
Other-Oriented Autonomy	Other-oriented autonomy is described as endorsing behaviors with the motivation and high consideration of other people.	I took computer science for my family I spent my vacation at home to take care of my mom I intentionally graduated late to spend more time

with someone I loved
I volunteered for competitions
to see my
grandmother smile
I talk to and include the shy kid
in class
I spoke in a presentation for my
scared groupmates
despite being scared too
I spend extra effort to animate
a project for my group
and my own grade
Working over the summer to
fulfill my working dreams
and to help out family
Choosing priorities over games
to help out groupmates
and support them in their study
I choose my course because I
wanted to give my
time to people with special
needs

Preliminary analysis

Descriptive statistics was used to test the normality of the data. All the indicators of normality appeared to be good except for the distribution of scores on the other-oriented dimension tends to be slightly leptokurtic. This would mean that the scores in the said dimension reached a higher peak compared to a normally distributed set of data.

Table 2
Descriptive Statistics of Autonomy Dimensions

Variable	M	SD	Skewness	Kurtosis	SE
Self-oriented	3.97	0.433	-0.459	0.174	0.341
Other-oriented	3.8	0.436	-0.567	1.387	0.341

Prior to the main analysis, the reliability of the self-oriented and other-oriented dimensions was tested and the obtained Cronbach’s

alpha coefficients were 0.70 and 0.66, respectively. The computed Bartlett's Test of Sphericity value of 777.692 ($p < .001$) and the Kaiser-Meyer-Oklín (KMO) value of 0.644 suggest that the data are acceptable to undergo factor analysis.

Main Analysis

A two-stage exploratory factor analysis (EFA) was performed. The first stage aimed to extract the factors and identify the items that loaded in each factor, while the second stage involved analyzing the retained items after ruling out the those that did not load in any of the factor and those that cross-loaded. EFA through Principal Axis Factoring (PAF) was performed with Oblimin rotation. The researcher decided to use PAF method because the multivariate normality of data distribution in the study cannot be assumed. In addition, Oblimin Rotation method was used because there is a reason to believe that the dimensions in the present study are correlated. Using the minimum Eigenvalue of 1 and screen plot as criteria, two factors which accounted for 58.36% of the variance were extracted.

The obtained pattern matrix in the first EFA obtained shows factor loadings ranging from 0.10 to 0.63 in Factor 1 and factor loadings ranging from 0.12 to 0.72. Four items from self-oriented dimension (e.g., *"I make all my decisions for my own welfare."*; *"I do not allow others to influence my decisions in life."*; *"I enroll in college to work on the skills and abilities that I lack."*; and *"Only a few of my decisions in life are self-motivated."*) and four items in other-oriented dimension (e.g., *"I rarely consider other people in making personal decisions."*; *"I do not do an action when I know that other people would be compromised."*; *"I do things willingly for other people."*; and *"I rarely let others affect my decisions."*) with factor loadings that are $< .40$ were omitted and were excluded in the second EFA. No items were excluded due to cross-loading.

Excluding the items with loadings $< .40$ in the initial EFA, the obtained pattern matrix in Table 2 shows the second EFA where six items loaded in Factor 1 with factor loadings ranging from .46 to .61, while five items loaded in Factor 2 with factor loadings ranging from .41 to .71. The second EFA was performed to check if similar results would appear after excluding the items with low factor loading values. The two factors accounted for 60.48% of the variance and the same items that loaded in both factors in the initial EFA also loaded in the

second one. Factor 1 loadings (Eigenvalue = 4.85; % of variance = 40.46) showed six items that came from one of the hypothesized dimensions which describe behaviors that are self-oriented in nature. A sample is “*I do things to prove to myself that I can do it.*” The six items in this factor represent the inward autonomy dimension. The cronbach alpha of .70 for the said dimension remained even after the second EFA, suggesting that the first autonomy dimension has acceptable level of internal consistency.

Table 3
Factor Loadings for Exploratory Factor Analysis with Rotation of Autonomy Dimensions

Items	Factor 1 Inward	Factor 2 Outward
I am pursuing my studies to have a sense of personal achievement.	0.47	0.24
If I am to join an organization, I will do it for my own improvement.	0.58	0.035
Whenever I make big decisions, I make sure that it is good for my own future.	0.46	0.09
I want to succeed in life to fulfill my own dreams.	0.61	0.018
I do things to prove to myself that I can do it.	0.51	0.13
I accomplish things to have a sense of self-worth.	0.56	0.21
*Only a few of my decisions in life are self-motivated.	0.04	0.54
I consider other people whenever I do important decisions.	0.12	0.55
I will not leave my group because I know that they need me.	0.04	0.41
I work harder to see my family happy.	0.2	0.71
It is important for me to not cause problem to others whenever I make actions.	0.19	0.5

Note. Factor loadings >.40 are in boldface. Items with asterisk (*) are reversed coded.

Moreover, five items that load in Factor 2 (Eigenvalue = 1.16; % of variance = 20.02) came from the hypothesized dimension that characterized behaviors that are concerned with the welfare of other people. Hence, this factor is labeled as *outward autonomy*. A sample item is “*It is important for me to not cause problem to others whenever I make actions.*” This dimension refers to self-endorsed behaviors that are guided by an individual’s concern to others’ welfare or situation. The cronbach alpha coefficient of 0.66 for this dimension suggests a moderate level of

internal consistency during the first EFA and increased to .69 in the second EFA. The overall reliability of the two dimensions combined is 0.69 in the first EFA and .72 in the second, which also suggests an acceptable level of internal consistency.

Discussion

The present study aimed to define autonomy by testing its structure using data from individuals from an interdependent culture. This paper argued that the definition of autonomy should be explored in a nation living in an interdependent culture like the Philippines (Markus & Kitayama, 1991b) since many of the motivations behind the behaviors displayed in such cultural orientation were theorized to have much more concern about the context and welfare of the people around. By conducting an initial qualitative study, item development derived from the qualitative data, and performing EFA twice with the data from the quantitative phase, two major dimensions of autonomy was found to operate among the participants—*inward* (ego-oriented) and *outward* (other-oriented) *autonomy*. Thus, autonomy is defined in this study as a motivational state that promotes behaviors that are fully endorsed based one's own interest and/or the welfare of other people, and situational context—with inward and outward autonomy as its dimensions. Inward autonomy is a motivational state that produces self-endorsed behaviors that concern one's own interests, values, and desires, whereas outward autonomy is another motivational state that fully promotes based on context and other people's welfare and situation.

The findings suggest that Filipinos' behaviors may not only be endorsed for one's interest but it may also be influenced by how other people would be affected by our decisions and actions. The findings reflect the Filipino notion of "*kapwa*" (Enriquez, 1977) which explains Filipino's sensitivity to other people. The findings were consistent with the theoretical formulation of Markus and Kitayama (1991a, 1991b) that people from interdependent culture have high context. This means that individuals with interdependent self-construal tend to be highly sensitive to the overall situation before endorsing an action or behavior. They are more likely to screen the behaviors that they would display in order to maintain harmony and good social relations. These findings pose a challenge to the conceptualization of autonomy in the

context of SDT which is formulated using mostly Western samples. Since most studies on autonomy using SDT framework are conducted using Western samples, it may be possible that its formulation is biased towards autonomy as a self-oriented construct. The findings of the present study suggest a more culturally sensitive framework in understanding autonomy.

Understanding the interdependent nature of autonomy among Filipino adolescents can help teachers, educators, and allied professionals in dealing with the lack motivation of students. Lack of motivation causes serious long-term problems (e.g., high drop out rate, diversion to drugs and violence, poverty, etc.) among students in any level and poses a challenge to parents, educators, and the society. Student motivation can be increased not only by promoting internal sources of motivation but also by promoting external sources such as the will to help one's family and others. External sources of motivation can also be provided by setting up an environment that teaches students to be more sensitive with the situation of other people and with the context. Being sensitive with the context can motivate students to independently endorse behaviors for the benefit of others and inhibit those that do not promote harmony. This study does not suggest to set aside the internal causes of motivation because it is equally important with its external counterpart.

The present study also has implications to testing. Researchers, educators, and psychometricians should be careful in using standardized measures. Fully relying on standardized tests without further validation using local norms may lead to misleading information. Most standardized tests made used Western samples with an independent view of the self. There is a need to use more culturally sensitive frameworks that give attention to the unique understanding of the self from an interdependent cultural perspective. Since people from interdependent countries like the Philippines view the self not as a unique entity but as a self that is defined by interpersonal relationships (Markus & Kitayama, 1991), using tests that were developed using Western samples may bring errors in our understanding of different constructs and phenomena as they operate in this culture. The present study suggests developing locally-made assessment tools that can best capture the indigenous characteristics of Filipino samples. The researcher also suggests that future research should verify the results of this study by revisiting the items used, and

selecting other types of populations since this study only focused on adolescents. Furthermore, confirmatory factor analysis using a new set of sample should be done to verify the results. While the results of this study are preliminary and have limitations, we now have a new measure of autonomy that is more culturally relevant to the experiences of Filipinos. It is the hope of the researcher that this new measure becomes a useful tool in the assessment of autonomy among Filipinos.

References

- Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology*, 78(4), 791-808
- deCharms, R. (1968). *Personal causation*. New York: Academic Press.
- Deci, E.L. (1975). *Intrinsic motivation*. New York: Plenum.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination theory in human behavior*. New York: Plenum.
- Enriquez, V. (1977). Filipino Psychology in the third world. *Philippine Journal of Psychology*, 10(1), 3-18.
- Iyengar, S. S., & DeVoe, S. E. (2003). Rethinking the value of choice: Considering cultural mediators of intrinsic motivation. In V. Murphy-Berman & J. J. Berman (Eds.), *Nebraska symposium on motivation: Cross-cultural differences in perspectives on self* (Vol. 49, pp. 129–174). Lincoln: University of Nebraska Press.
- Jordan, J. V. (1991). The relational self: A new perspective for understanding women's development. In J. Strauss & G. Goethals (Eds.), *The self: Inter-disciplinary approaches* (pp. 136–149). Cambridge, MA: Harvard University Press.
- Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A Cultural look at new look. *Psychological Science*, 14(3), 201-206
- Markus, H. R., & Kitayama, S. (1991a). Cultural variation in the self-concept. In J. Strauss & G. R. Goethals (Eds.), *The self: Interdisciplinary approaches* (pp. 18–24). New York: Springer-Verlag.
- Markus, H. R., & Kitayama, S. (1991b). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253
- Na, J., & Kitayama, S. (2011). Spontaneous trait inference is culture-

- specific: Behavioral and neural evidence. *Psychological Science*, 22(8), 1025-1032. DOI: 10.1177/0956797611414727
- Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, 26, 419–435.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist*, 55, 68–78.
- Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does Psychology need choice, self-determination, and will? *Journal of Personality*, 74, 6. DOI: 10.1111/j.1467-6494.2006.00420.x
- Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA. MIT Press



Multigroup Invariance of Course Experience Questionnaire Across Gender Using Malaysian Sample

Lei Mee Thien

SEAMEO RECSAM

Abstract

The purpose of this study is to examine the measurement invariance of CEQ across gender of Malaysian undergraduate students in two public universities. A survey method was employed for data collection. A total of 315 undergraduate students in two Malaysian public universities responded to 23 items measuring five scales of Good Teaching, Generic Skills, Clear Goals and Standards, Appropriate Workload, and Appropriate Assessment. Covariance Based Structural Equation Modelling was used as the analysis method. A multigroup analysis of invariance was performed to examine measurement invariance across male and female students using AMOS 20.0 computer software. Findings showed that configural invariance was fully supported whereas metric invariance was partially supported across female and male groups. This study has shown the satisfactory psychometric properties of CEQ 23 with only one item that was found not convey the same meaning across gender. Findings could be used as the empirical evidences to compare with the future empirical-based CEQ studies.

Keywords: multigroup invariance, Course Experience Questionnaire, teaching quality, configural invariance, metric invariance

Introduction

Surveys of student perceptions of their program or course experience in higher education are found widely used over the world such as Australia and United Kingdom (Yorke, 2009). The results from

the survey are particular importance as it serves as the key performance indicators of the higher education institution (Yorke, 2009). In relation to this, Course Experience Questionnaire (CEQ) was developed and adopted as a domain-neutral indicator of university course quality. CEQ is an Australian-based instrument that has been used to probe university students' perceptions of their programme of study experience in higher education. CEQ has been widely used as a performance indicator and benchmarking of students' learning experience in Asia Pacific and Western countries particularly Australia, UK, and Canada universities context (Ginns, Prosser, & Barrie, 2007). Deriving from the theory of learning and teaching, CEQ was originated from the item pool of Course Perception Questionnaire (Ramsden & Entwistle, 1981), School Experience Questionnaire (Ramsden, Martin, & Bowden, 1989), as well as Experiences Studying and Higher Education Questionnaire (Entwistle & Tait, 1990). CEQ 30 which was advanced by Trigwell and Prosser (1991) and Ramsden (1991) is the first version that evolved from the item pools mentioned. CEQ 30 comprised five scales, such as: Good Teaching (8 items), Clear Goals and Standards (5 items), Appropriate Workload (5 items), Appropriate Assessment (6 items) and Emphasis on Independence (6 items). The psychometric properties of CEQ 30 was warranted and has undergone series of improvement from the statistical perspective over decades (e. g., Ainley & Long, 1994; Richardson, 1994; Wilson, Lizzio, & Ramsden, 1997).

The improved version refers to CEQ 23. CEQ 23 was developed in consultation with the Department of Employment, Education, and Training (DEET) and has been used in national survey of Australian graduates since 1993. Differs from former version, CEQ 23 consisted of the scales of Good Teaching (6 items), Generic Skill Scale (6 items), Clear Goals and Standards (4 items), Appropriate Workload (4 items), and Appropriate Assessment (3 items) with the exclusion of scales of Emphasis on Independence (6 items). Interestingly, CEQ 23 has been extensively piloted in different contexts using a variety of data analysis. However, the main concern of the existing studies is limited to the levels of internal consistency and satisfactory of its factor structure (Wilson et al., 1997). For instances, in British university context, Broomfield and Bligh (1998) confirmed the five factor structures of CEQ 23 and further split Good Teaching scale into two subscales: teacher interaction and presentation; and the quality

of feedback given to students. Distinctly, Eley (2001) examined three parallel versions of CEQ 23 to examine whether altering question format and phrasing can improve the effectiveness of CEQ. The first version refers to Graduate Careers Council of Australia (GCCA) with 5-point Likert scale from agree (1) to disagree (6). The second and third versions modified the regular CEQ 23 to devise Behavioural Observation Scale (BOS) and Dimensional Rating Scale (DRS), respectively. The internal consistency appeared to be satisfactory for the three versions of regular CEQ, BOS, and DRS version. However, the five factor solutions for each version warrant comments with the presence of several items which loaded on unintended scales. Similarly, Byrne and Flood (2003) confirmed the satisfactory levels of reliability and construct validity for use in the accounting discipline in an Irish university. Meanwhile, in Hong Kong universities context, Ho (1998) as well as Law and Meyer (2011) have provided the satisfactory evidence of reliability of the CEQ 23. On the other hand, Thien and Ong (2016) found that only two dimensions of CEQ, namely: Good Teaching and Generic Skills were applicable in Malaysian higher education context.

Notably, most of previous studies have typically assumed CEQ 23 was operating in exactly the same way across the groups of interest. Such assumption sounds less convincing as extant literature informed gender differences in many areas of higher education research (Grebennikov & Skaines, 2009). For instance, female students now outnumber male students (Bradley, 2000). Female students were also found over-represented in most of the Malaysian public universities (Ismail, 2015). In relation to this, multigroup invariance which refers to the extent to which the content of each item is being perceived and interpreted in the same way across groups should be initially emphasised (Byrne & Watkins, 2003). The ignorance of multigroup analysis of invariance across groups often lead to contradictory findings that subsequently misleading the direction of future studies (Byrne, 2008, 2010). As such, the multigroup analysis of invariance need to be conducted using a rigorous statistical techniques suggested by Byrne (2004). The significance of this empirical study hinges upon the fact that it contributes methodological knowledge in higher education literature using a Malaysian undergraduate student sample.

Conceptualization and Operationalization

Course Experience Questionnaire (CEQ) advanced by Ramsden (1991) has five underlying scales: (1) Good Teaching, (2) Generic Skills, (3) Clear Goals and Standards, (4) Appropriate Workload, and (5) Appropriate Assessment.

According to the report of Graduate Course Experience Sydney 2009, Good Teaching is conceptualized as the nature of teaching experienced during a course. Good Teaching is operationalized as the degree to which the graduates feel that the teaching staff or lectures of their course provided a high level of teaching quality. Meanwhile, Generic Skills is conceptualized as the enhancement of selected generic skills with its operationalization as the extent to which the course adds to the generic skills that graduates might be expected to possess.

The third scale, namely, Clear Goals and Standards is conceptualized as the clarity and meaningfulness of course structure. Clear Goals and Standards is thereby operationalized as the extent to which graduates feel that they were provided with enough information regarding the learning objectives of their course and the standards of work expected from them.

On the other hand, Appropriate Workload is conceptualized as the level of workload that hindered deeper forms of learnings and it is operationalized as the degree to which graduates felt the workload involved in their course were excessive. The fifth scale, namely, Appropriate Assessment is conceptualized as the level that assessment that promoted deeper forms of learning. It is operationalized as the extent to which courses depend on the recollection of factual knowledge for assessment purpose. Table 1 shows the number of items for each scale and its description. Each scale was capitalized throughout the paper to avoid interpretation confusion.

Table 1
Item Descriptions

Scale	Item	Description
Good Teaching (GTS)	GTS1	The lecturer put a lot of time into commenting on my work.
	GTS2	The lecturer of this course motivated me to do my best work.
	GTS3	The lecturer made a real effort to understand difficulties I might be having with my homework.

Scale	Item	Description
	GTS4	The lecturer normally gave me helpful feedback on how I was doing.
	GTS5	The lecturers were extremely good at explaining things.
	GTS6	The lecturers work hard to make their subject interesting.
Generic Skills (GSS)	GSS1	The program developed my problem-solving skills.
	GSS2	The program sharpened my analytic skills.
	GSS3	The program helped me develop my ability to work as a team member.
	GSS4	As a result of my program, I felt confident about tackling unfamiliar problems.
	GSS5	The program improved my skills in written communication.
	GSS6	The program helped me to develop the ability to plan my own work.
Clear Goals and Standards (CGS)	CGS1	It was easy to know the standard work expected.
	CGS2	I usually had a clear idea of where I was going and what was expected of me in this program.
	CGS3*	It was often hard to discover what was expected of me in this program.
	CGS4	The lecturers made it clear right from the start what they expected from students.
Appropriate Workload (AWS)	AWS1*	The workload was too heavy.
	AWS2	I was generally given enough time to understand the things I had to learn.
	AWS3*	There was a lot of pressure on me as a student in this program.
	AWS4*	The sheer volume of work to be done though in this program it means it couldn't all be thoroughly comprehended.
Appropriate Assessment (AAS)	AAS1*	To do well in this program all you really needed was a good memory
	AAS2*	The lecturers seemed more interested in testing what I had memorized than I had understood.
	AAS3*	Too many lecturers asked me questions about facts.

Note: * represents recoded items.

The Present Study

The purpose of the present study is to examine the item measures for each dimension of CEQ, namely (1) Good Teaching, (2) Generic Skills, (3) Clear Goals and Standards, (4) Appropriate Workload, and (5) Appropriate Assessment across gender of Malaysian undergraduate students in two public universities.

Method

Participants

The CEQ 23 was administered on 350 students in second, third, and fourth year undergraduate programs in academic year 2013 in two Malaysian public universities. The cohort of second, third, and fourth year students were selected because they have the experience with university learning culture. Therefore, their evaluation on teaching quality can be considered reliable. The participants were selected with convenience basis due to the time and cost constraints. The survey was administered by the current researchers with the assistance from the researcher officers who worked in the university. Questionnaires were distributed to the randomly selected social science and science-based faculties, including the faculties of education, social science, humanities, engineering, biological science, and computer sciences. A total of 315 of the completed questionnaires were returned with the response rate of 90%. According to Rasoolimanesh, Jaafar, Kock, and Ramayah (2015), only 166 sample for model testing is needed to generate a power value of 0.95. Therefore, in this study, the power value of sample size of 315 was exceeded 0.95 and thus considered sufficient.

The student sample consisted of 109 male and 206 female students. The dominant ethnic group was Malay (224) followed by Chinese (66), Indian (9) and the other minority ethnic groups (16). The composition of academic year of study for the participant was second year (222), third year (91) and fourth year (2) undergraduate students with the average age of 22 years old.

Measures and Data Collection Procedures

Table 1 shows the defining items of the CEQ scales according to the report of Graduate Course Experience Sydney 2009. The data collected were entered into SPSS data file. Items that had an opposite meaning to that of the relevant scale were then recoded for the ease of finding interpretation. The recoded items were referred as AAS1, AAS2, AAS3, AWS3, AWS4 and CGS3. Each item was scored on a 5-point Likert scale from strongly disagree (1) to strongly agree (5). For convenience, the terms of 'staff' in original version of CEQ 23 has been modified as 'lecturer' in this study due to the contextual nature in local public universities. Similarly, the terms of 'course' has been modified to 'program' as the evaluation is based on the entire program instead of a specific course offered in each faculty.

The Hypothesized Model

Figure 1 shows the hypothesized model schematically. There are five scales of CEQ: Generic Skills scale (6 items), General Teaching Scale (6 items), Clear Goals and Standards scale (4 items), Appropriate Workload scale (4 items), and Appropriate Assessment scale (3 items). Figure 1 shows the five scales or factors are intercorrelated as indicated by the double-headed arrows, and the measurement error terms associated with the observed variables (e 's) are uncorrelated.

Data Analysis Procedures

This study employed Byrne's (2010) steps for testing measurement invariance using AMOS 20.0 computer software. In Step 1, the overall model, male- and female model are initially tested to ensure the model fit to provide an overview of how consistent the models results are. If consistency of factor structure is found, then the analysis proceeds to Step 2. Otherwise, the models need to re-specify rigorously for consistency. Step 2 involves the establishment of configural and metric invariance when conducting two or more different groups of sample (Hair, Black, Babin, & Anderson, 2010).

Configural invariance is satisfied when the basic model structure, indicating the pattern of fix and non-fixed parameters is invariant across group. The initial baseline model has no between-

group invariance constraints on estimated parameters. The configural invariance is therefore critically important because it provides the basis for comparison with all subsequent models. On the other hand, conducting metric invariance is to assure whether the items underlying each scale are used similarly across groups (Vandenberg, 2002; Hair et al., 2010). The test of metric invariance is conducted by constraining the factor loadings to be equal across groups because the pattern coefficients carry the information about the relationship between latent scores and observed scores. When metric invariance is established, the different scores on the item can be meaningfully compared across groups. Vandenberg and Lance (2000) proposed that configural and metric invariance should be established before comparisons across groups can be meaningful.

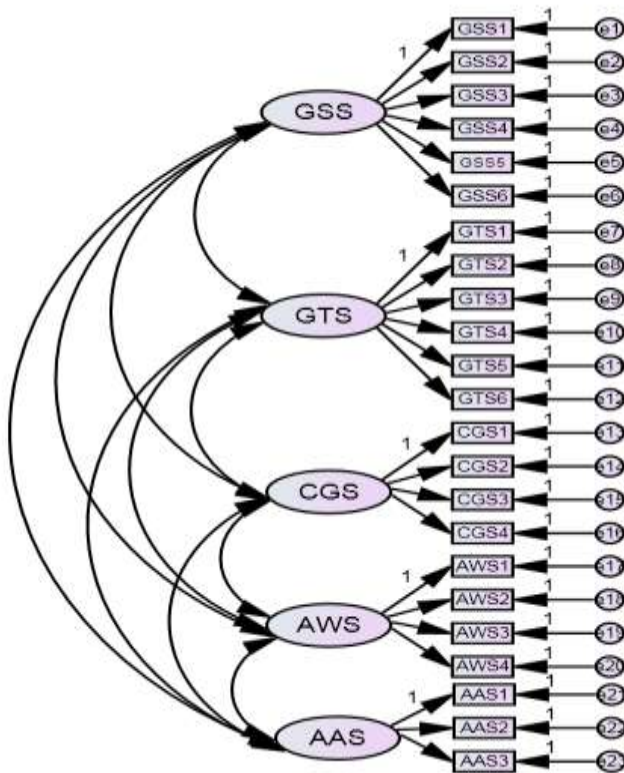


Figure 1 Hypothesized Model

The analysis procedures of Step 2 begins with a baseline model with chi-square value which derived by computing model fit for the pooled sample of all groups. Evidence of non-invariance is claimed if the chi-square difference value is statistically significant. The analysis was then followed by a hierarchical series of additional test that aimed at targeting which parameters are accounting for the non-invariant findings.

Results

Prior to multigroup analysis, a preliminarily Confirmatory Factor Analysis (CFA) for the overall data, female, and male groups were conducted. Table 2 shows the hypothesized model for overall data, female and male groups yield only a marginally good fit to the data. The initial poor CFA results showed the needs for model re-specification.

The overall data model was re-specified with the exclusion of Item CGS3 and inclusion of error covariance between GTS1 and GTS3 yielded some improvement in goodness-of-fit, including Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean-Square Error of Approximation (RMSEA). This results in a modestly adequate fit for both female and male groups. Accordingly, the re-specified female and male models were deemed the most appropriate baseline models for both groups.

The first step of testing was configural invariance where no equality constraints were imposed. This configural model simply incorporates the baseline models for female and male groups and allows for their simultaneous analyses. Table 3 shows the configural model which also termed as unconstrained model fits reasonably well with $\chi^2(396, 206) = 1002.286$, TLI = 0.81, CFI = 0.84 and RMSEA = 0.07 and 95% C.I. between 0.064 and 0.075. The results concluded that the number of factors and pattern of their item loadings were similar across female and male groups.

For measurement invariance, we compared the factor loadings equivalence model to the unconstrained model. This initial test was termed full metric invariance. The results indicated the factor loadings were significantly different or not invariant with the difference in χ^2 difference was 39.90 for female and male group and significant at the 0.05 level.

Table 2
Summary Goodness-of-Fit Statistics in Determination of Baseline Models

Model description	χ^2	df	CFI	TLI	RMSEA	90% CI
<i>Overall data (N=315)</i>						
Hypothesized five-factor model (Model 1)	854.73	220	0.83	0.80	0.090	0.089, 0.103
Model 1 with the exclusion of Item CGS3 (Model 1a)	735.82	199	0.84	0.82	0.089	0.087, 0.101
Model 1a with one error covariance specified (Item GTS1 and GTS3)	703.05	198	0.86	0.84	0.087	0.083, 0.097
<i>Female undergraduate students (n=206)</i>						
Hypothesized five-factor model (Model 2)	735.43	220	0.79	0.76	0.107	0.098, 0.115
Model 2 with the exclusion of Item CGS3 (Model 2a)	663.39	199	0.81	0.77	0.106	0.098, 0.115
Model 2a with one error covariance specified (Item GTS1 and GTS3)	626.67	198	0.82	0.80	0.103	0.094, 0.112
<i>Male undergraduate students (n=109)</i>						
Hypothesized five-factor model (Model 3)	442.95	220	0.84	0.81	0.097	0.084, 0.110
Model 3 with the exclusion of Item CGS3 (Model 3a)	393.22	199	0.85	0.83	0.095	0.081, 0.109
Model 3a with one error covariance specified (Item GTS1 and GTS3)	375.62	198	0.87	0.84	0.091	0.077, 0.105

Note. *df* represents degree of freedom, CFI represents Comparative Fit Index, TLI represents Tucker-Lewis Index, RMSEA = Root Mean-Square Error of Approximation, and CI represents confidence interval. Threshold of CFI, TLI, and RMSEA is 0.90 (Hu & Bentler, 1999). 0.90 (Hu & Bentler, 1999), and 0.03 (Browne & Cudeck, 1993), respectively.

Table 3

Invariance Tests for Loadings across Female and Male Groups

Model description	Comparative model	χ^2	df	$\Delta\chi^2$	Δdf	p
Unconstrained model (model1)		1002.286	396	n.a.	n.a.	n.a.
Metric invariance (model 1A)	1A Vs 1	1042.104	413	39.818	17	$p < 0.001$
Constrained all loadings of GSS to be equal (model 1B)	1B Vs 1	1007.756	401	5.469	5	n.s.
Constrained all loadings of GSS and GST to be equal (model 1C)	1C Vs 1	1015.175	406	12.889	10	n.s.
Constrained all loadings of GSS, GST and AWS to be equal (model 1D)	1D Vs 1	1037.063	409	34.777	13	$p < 0.001$
Constrained all loadings of GSS, GST & AWS2 to be equal (model 1E)	1E Vs 1	1036.110	407	33.824	11	$p < 0.001$
Constrained all loadings of GSS, GST & AWS3 to be equal (model 1F)	1F Vs 1	1015.274	407	12.988	11	n.s.
Constrained all loadings of GSS, GST & AWS3 & AWS4 & CGS to be equal (model 1G)	1G Vs 1	1019.192	410	16.906	14	n.s.
Constrained all loadings of GSS, GST & AWS3 & AWS4, CGS & AAS to be equal (model 1H)	1H Vs 1	1022.976	412	20.69	16	n.s.

Note. n.s. represent nonsignificant at .05 level and n.a. represents not applicable.

Since we do not have full metric invariance, the analysis was further proceeded to partial metric invariance. Partial metric invariance is used to identify at least two equal factor loadings between all constructs (Vandenberg & Lance, 2000). Thus, given a finding that the test rejected the null of equality, we further investigated the scales following the multisteps recommended by Byrne (2010). We investigated scale with problematic loading(s) to detect problem scale in advance in female and male groups by imposing constraint on all loadings within each scale in order. A chi-square different test fails to reject the null of equality of all loadings in a certain scale, then all the items in the scale were invariant between female and male group. If any chi-square different test rejected the null of equality of all loadings in a scale between female and male group, we then used a series of analyses by placing constraints on individual loadings in sequence in the scale. Finally, we compare chi-square values of constrained models with the baseline model (Byrne, 2010). Through the multistep process, all tests for invariance of the 22 loadings across female and male were completed. Table 3 shows only one item, namely, Item AWS2 is found non-equivalent between female and male groups.

Discussion

In reality, it is possible that female and male respondents could perceive the contents of the items in CEQ 23 differently. In line with this point of view, Bentler (2004) reiterates that there is no guarantee that the instrument particularly in the form of questionnaire operates equivalently across different groups such as gender and ethnic groups. Considering this, there is a need to establish the consistency with respect to the relationship between a latent variable and its corresponding subscales or items across different groups. In other words, the items should be equally valid across different groups before conducting further analysis. Therefore, it is worthy for this present study to examine the mutligroup analysis of the measures of CEQ 23 using the measurement invariance across gender.

This study has shown one item, namely, AWS2 did not operate the same way across gender. AWS 2 which stated as "*I was generally given enough time to understand the things I had to learn*" deserves further investigation. The female and male undergraduate students may have misinterpreted the phrases. This phrase could be broadly

conceptualized and not well defined which carries different interpretation from female and male students based on their preferences of program studied. In this regard, item developers could revise and specify AWS2 based on a preference of program studied among the undergraduate students.

However, this study was limited to the invariance of factor loadings, representing a minimum condition for multigroup analysis. Therefore, it would be insightful to extend the present study that involved the testing of scalar invariance, factor covariances, and error covariance to achieve a full multigroup analysis of CEQ 23 in future studies. This study is also limited with the small sample size of the number of male students which only half of the female students and only involved two public universities. Future studies can be extended by including a larger sample size across gender. In addition to gender, measurement invariance across ethnics can be another avenue that deserves to be explored in future studies. Viewing from methodology perspective, the testing of measurement invariance is based on the chi-square statistic which are known to be sensitive to sample size. For this reason, using alternative fit indices in measurement invariance investigations is recommended for future research endeavor (Cheung & Rensvold, 2002).

For methodology implication, this study has shown the satisfactory psychometric properties of CEQ 23 with only one item that was found not convey the same meaning across gender. The findings of this present study would serve as the empirical evidence that can be used to compare and contrast with the findings of future empirical studies related to CEQ. More importantly, the method and procedures used of multigroup analysis as shown in this study could serve as an exemplar and replicable in future studies in addition to contribute methodological knowledge in higher education literature.

References

- Ainley, J., & Long, M. (1994). *The course experience survey 1992 graduates*. Canberra: Australian Government Publishing Service.
- Bradley, K. (2000). The incorporation of women into higher education: Paradoxical outcomes? *Sociology of Education*, 73, 1–18.
- Bentler, P. M. (2004). *EQS: Structural equations programs manual*. Encino, CA: Multivariate software.

- Broomfield, D., & Bligh, J. (1998). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey-Bass.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of accessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equations models* (pp. 136-162). Newbury Park, CA: Sage.
- Byrne, B. M. (2004). Testing for multigroup invariance using AMOS graphics: A road less traveled. *Structural Equation Modeling*, 11(2), 272-300.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicotherma*, 20(4), 872-882.
- Byrne, B. M. (2010). *Structural Equation Modeling with AMOS: Basic concepts, applications and programming* (2nd ed.). NY: Routledge.
- Byrne, B. M., & Flood, B. (2003). Assessing the teaching quality of accounting programmes: An evaluation of the course experience questionnaire. *Assessment and Evaluation in Higher Education*, 28, 135-145.
- Byrne, B. M., & Watkins, D. (2003). The issue of measurement equivalence revisited. *Journal of Cross-cultural Psychology*, 34, 155-175.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255.
- Eley, M. (2001). The course experience questionnaire: Altering question format and phrasing could improve the CEQ's effectiveness. *Higher Education Research and Development*, 20(3), 293-312.
- Entwistle, N. J., & Tait, H. (1990). Approaches to learning, evaluations of teaching and preferences for contrasting academic environments. *Higher Education*, 19, 169-194.
- Gitins, P., Prosser, M., & Barrie, S. (2007). Students' perceptions of teaching quality in higher education: The perspective of currently enrolled students. *Studies in Higher Education*, 32(5), 603-615. doi: 10.1080/03075070701573773.
- Grebennikov, L., & Skaines, I (2009). Gender and higher education experience: A case study. *Higher Education Research and Development*, 28(1), 71-84.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*. (7th ed.). New Jersey: Prentice Hall.

- Ho, A. S. P. (1998). *Changing teachers' conceptions of teaching as an approach to enhancing teaching and learning in tertiary education* (Unpublished doctoral thesis). The University of Hong Kong, Hong Kong.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Ismail, L. (2015). Gender gap in higher education: Perspective on factors influencing enrolment in Malaysian universities. *The Online Journal of Quality in Higher Education*, 2(4), 35-45.
- Law, D. C. S., & Meyer, J. H. F. (2011). Adaptation and validation of the course experience questionnaire in the context of post-secondary education in Hong Kong. *Quality assurance in Education*, 19(1), 50-66. doi: 10.1108/09684881111107753.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16, 129-150.
- Ramsden, P., & Entwistle, N. J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology*, 51, 368-383.
- Ramsden, P., Martin, E., & Bowden, J. (1989). School environment and sixth Form pupils' approaches to learning. *British Journal of Educational Psychology*, 59, 129-142.
- Rasoolomanesh, S. M., Jaafar, M., Kock, N., & Ramayah, T. (2015). A revised framework of social exchange theory to investigate the factors influencing residents' perceptions. *Tourism Management Perspective*, 16, 335-345.
- Richardson, J. T. E. (1994). A British evaluation of the course experience questionnaire. *Studies in Higher Education*, 19, 59-68.
- Thien, L. M., & Ong, M. Y. (2016). The applicability of Course Experience Questionnaire (CEQ) for a Malaysian university context. *Quality Assurance in Education*, 24(1), 41-55.
- Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: The influence of learning context and student approaches to learning on learning outcomes. *Higher Education*, 22, 251-266.
- Vandenberg, R. J. (2002). Towards a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, 5(2), 139-158.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance

- literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-69.
- Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and application of the course experience questionnaire. *Studies in Higher Education*, 22, 33-53.
- Yorke, M. (2009). 'Student experience' surveys: Some methodological considerations and an empirical investigation. *Assessment & Evaluation in Higher Education*, 34(6), 721-739.



Assessment Schemes in the Senior High School in the Philippine Basic Education

Carlo Magno
Tristan Piosang
National University

Abstract

The present report provides the different opportunities where assessment can be conducted in the Senior High School. The levels of assessment range from entry to end of cycle assessment and from classroom level to international level assessment. The levels of assessment are described in terms of best practices, purposes, how it translates into students learning, and accountability from the classroom teacher to policy makers. The levels of assessment in the Senior High School described in this report include: (1) Placement of students in the senior high school tracks, (2) Classroom-based assessment, (3) Assessment of achieved competencies, (4) Participation in international benchmarking of competencies, (4) College readiness assessment, and (5) Career assessment.

Keywords: Assessment, Senior High School, SHS tracks

Introduction

The Philippines has recently embarked on an educational reform in improving the curriculum and spreading the number of years of learning from 10 years to 13 years. This reform is nationally implemented through passage of the “Enhanced Basic Education Act of 2013” or RA 10533. The Act institutionalized the implementation of the K to 12 education program in the Philippines. The K to 12 covers kindergarten and 12 years of basic education. These 12 years is divided into three levels, which are: six years of primary education, four years of Junior High School, and two years of Senior High School (SHS).

The lengthened years is an effort to decongest and enhance the basic education curriculum in order to provide sufficient time for mastery of concepts and skills, develop lifelong learners, and prepare graduates for tertiary education and employment (Republic Act No. 10533; Magno, 2011).

The K to 12 program aims to produce graduates who are holistically developed, equipped with 21st century skills, and prepared for higher education, middle level skills development, employment, and entrepreneurship. This aim is mostly carried out by one salient feature of the program which is the Senior High School (SHS) that includes grades 11 and 12.

The SHS includes two years of specialized upper secondary education. It is one where the student has the option to choose a specialization based on their aptitude, interests, and school capacity. The students' choice of career track will define the content of the subjects they will have to take which will fall under either the core curriculum or specific tracks (SEAMEO INNOTECH, 2015).

Under the present SHS model, the student can choose among four tracks, namely: (1) Academic, (2) Technical-Vocational-Livelihood, and (3) Sports, and (4) Arts and Design. Under the Academic track are four strands, namely: General Academic Strand, Accountancy, Business, and Management (ABM), Humanities and Social Sciences (HUMSS), and the Science, Technology, Engineering, and Math (STEM).

In the SHS core curriculum, there are 15 core subjects which will have to be taken by a student irrespective of the track they are in. These core subjects are distributed to eight core learning areas. Furthermore, there are 16 track subjects, seven contextualized subjects, and nine specialization subjects which total to 31 subjects each student will have to take to earn a SHS diploma. The student will take these subjects in four semesters.

The SHS program is the realization of what is stated in Section 2, paragraph (a) of the RA 10533 that the state shall “broaden the goals of high school education for college preparation, vocational and technical career opportunities as well as creative arts, sports and entrepreneurial employment in a rapidly changing and increasingly globalized environment.” The main thrust, therefore, of the SHS is to produce “productive and responsible citizens” who are equipped with

the essential competencies, skills and values which will make them both a life-long learner and employment-ready.

Given the additional years of senior high school and new features of the basic education program, it is important to report on the opportunities on different levels of assessment that can be conducted in order to ensure quality implementation. By reporting on the levels of assessment in the SHS, policy makers, curriculum developers, teachers, and school administrators can start to develop and implement better assessment as integral part of the curricular programs in schools and within education at a national level. There is also a greater call for looking at assessment within a multilevel approach to monitoring learning outcomes at the local, national and international levels (Ho, 2012). Looking at the senior high school as additional leg to the Philippines basic education, there are assessment processes that will take on a different approach given the age of students and the system of implementing the SHS. At the same time, some assessments that were implemented in the previous years can be adapted and moved within the duration of the SHS.

The present report is conducted on the following purpose: (1) Guide policy-makers on the different opportunities where information can be taken coming from various evidences of learning in the newly implemented levels in senior high school. The data coming from assessment of learning in senior high school directs better decision on programs. (2) Provision of various levels of assessment enable schools offering senior high school to carefully evaluate the program and effectiveness of the implementation. (3) Having identified the various areas where assessment can be conducted in the senior high school, these areas serve as indicators when conducting large scale evaluation of the new program for the K to 12.

Senior High School in United States and Australia

The SHS model in the Philippines was patterned from models which are developed and being used in some first-world countries like the case of the United States of America (USA) and Australia.

In the USA, the model includes three to four years of SHS, depending on the state. The student spends grades 10 to 12 in order to earn an SHS diploma. SHS students must take core curriculum courses or subjects for a prescribed number of years (depending on the state).

These generally include English, Mathematics, General Science, Health, Physical Education, and Social Sciences. And as reported by the International Student and Scholar Services (ISSS) of the University of Minnesota, some high schools stream students for academic subjects where the brightest students are put on a ‘fast track and are given the opportunity to take enriched classes in their academic subjects. After the satisfactory completion of SHS, the student will be given a diploma which will enable him or her to take tertiary education (Corsi-Bunker, 2009). The national assessment taken in high school in the USA includes the Scholastic Assessment Test (SAT) for college admission and the American College Testing (ACT) to assess college readiness.

In Australia, the SHS model includes grades 11 and 12. There are 14 ‘senior secondary’ which fall under four core learning areas, namely: (1) English, (2) Mathematics, (3) Science, and (4) History. These 14 subjects will have to be completed (in addition to the Foundation to Year 10 of schooling in order for the student to be qualified for tertiary education. The Australian Curriculum, Assessment and Reporting Authority manages and delivers national assessments in Australia and overseas. They implement the National Assessment Program that provides tests endorsed by the Ministerial Council for Education, Early Childhood Development and Youth Affairs including the National Assessment Program Literacy and Numeracy (NAPLAN) and three-yearly sample assessments in science literacy, civics and citizenship, and ICT literacy (see acara.edu.au).

Assessment in the K to 12

There have been several reports that dealt with different kinds of classroom assessment in the K to 12. The report of Magno and Lizada (2015) described and explained the important features of formative assessment when used with instruction. The study came up with nine assessment principles that explain both theory and practice in the conduct of formative assessment inside the classroom. The study of Trey, Schmitt, and Allen (2012) explained that “those crucial elements which remain result in a description of a classroom assessment task that involves the student deeply, both in terms of cognitive complexity and intrinsic interest, and are meant to develop or evaluate skills and abilities that have value beyond the assessment itself.

It is this type of assessment experience that is, realistically speaking, authentic.” Assessment is supposed to engage the student works when the student has found it to be rewarding for its own sake. This practice of assessment is presently translated in the classroom where a large part of students grading and summative assessment is devoted on performance-based tasks in the Philippine setting (see DepEd order no. 8).

Assessment has salient and crucial contribution to the operation implementation, and success of the curriculum (Mikre, 2010). Through documenting the kinds of assessment that can be conducted in SHS, the teachers can further develop better functions and purposes of assessment. This will contribute to the successful conduct of classroom assessment which will ultimately result to success in the implementation of the SHS subjects and the K to 12 curriculum as a whole. Teachers are also better able to implement classroom assessment if they are equipped with assessment literacy which entails understanding the appropriate use of classroom assessment.

Assessment and the SHS

The Enhanced Basic Education Act of 2013 spells out features of the K to 12 program in terms of the goals, intentions, outcomes, curriculum, and learning areas. These features needs to ensure that students improve better overtime where the necessary 21st century outcomes are achieved by the students that includes college employment readiness. Assessment takes an important role to ensure that students have developed the 21st century skills. Assessment in the K to 12 becomes more functional where it is seen as an integral part of contributing to student learning. This is even indicated in the Department of Education Order No. 8 where assessment “allows the teachers to track students progress... and assessment informs the learners, their parents and their guardians of their progress.” Furthermore, assessment is defined as “a process that is used to keep track of learners’ progress in relation to learning standards and in the development of 21st century skills; to promote self-reflection and personal accountability among students about their own learning; and to provide bases for the profiling of student performance on the learning competencies and standards of the curriculum.” Because of

the many critical roles played by assessment, it is described as an essential component of curriculum practice (Akker, 2003).

The addition of the SHS levels in the basic education provides several pathways on making assessment functional. The present report specifies assessment opportunities in the SHS in order to better realize appropriate instruction for learners and better scaffolding students learning in the classroom and national level. The level of assessment in the SHS described in this report is seen as an array covering assessment in an individual level, classroom level, and a larger scale such as the national and international arena. These assessment levels range from entry level assessment, classroom level, national level assessments, and international level. The array also describes another dimension in terms of the accountability. The smaller is the scope of the assessment within the classroom, the more specific interventions can be provided for learners. The larger is the scale of the assessment covering national and international level, the more policy makers are informed to make better decision on educational development and programs across the country. The smaller is the scope of the assessment, the more the teachers and curriculum implementers are accountable on students' progress. The larger is the scale of the assessment, the more policy makers are accountable for creating educational support.

1. Placement of Students in the Senior High School Tracks

The students entering the senior high school has gone through a curriculum where the learning areas are spiralled within 11 years including the language and literacy, mathematics, science, social studies and history, technology and livelihood education, music, arts, physical education, and health. On the other hand, the SHS is structured where the learning areas (subjects) change every semester (or trimester in some schools). Aside from common subjects taken (core and applied), the students go through a track that they have selected. The tracks include Academic Track (Science, Technology, Engineering, and Mathematics [STEM], Accounting and Business Management Track [ABM], Humanities and Social Science [HUMSS], General Academic Strand [GAS], Technical-Vocational Livelihood Track (Home Economics, Industrial Arts, Agri-fishery, and Information and Communications Technology), Sports track, and Arts and Design.

Students are differentiated in these tracks depending on their aptitude and interests. The students entering the SHS need to have the necessary aptitude, interest, and cognitive capability matching the tracks that they will be entering (SEAMEO INOTECH, 2014). The cognitive capacity, aptitude, and interest are basis that can provide information in order to determine the appropriate tracks that students can enter. The combination of cognitive capacity, aptitude, and interest are appropriate indicators in the entry level assessment in the SHS. The cognitive capacities of the students in the core learning areas directly give information on the specific strengths and weaknesses at the end of grade 10. The grade 10 is the last level before entering the SHS and assessment of the recent competencies provides an accurate basis of what students has mastered overtime in terms of the common core capacities (language, mathematics, and science). The assessment of cognitive capacity can be operationally defined in the form of an achievement test in English, mathematics, and science covering the competencies in grade 10. Part of the assessment is a component of aptitude. Aptitude measures the degree of students' potential for future training (Barret, 2011). The aptitude test results give information whether learners will be successful in the tracks they have chosen. The skills involved in aptitude measures fluid intelligence where capacities are innate among the learners. These innate capacities that learners excel should match the capacities required in the track they have selected. Both achievement and aptitude are measures of mental capacities where students' strengths and weaknesses are determined. The assessment model for entry level in SHS needs to include an affective domain. The specific affective construct that can be measured is 'interest' in the areas of the SHS track. Interests are measures of preferences of different situations (Holland, Fritzsche, & Powell, 1997).

If achievement, aptitude, and interest serve as basis for the entry level assessment to determine the appropriate track of students, specific constructs in the forms of variables, skills, and competencies required in each track needs to be specified. For instance, students entering the science, technology and engineering track at least needs to have mastered competencies in science and mathematics, posses an aptitude in syllogism (inferring conclusions) and deciphering series of events, interest in life sciences, medicine, and engineering. A model is needed that specifies how skills and characteristics are required for

each track. A model is provided by the Center for Learning and Assessment Development - Asia on the specific factors of achievement, aptitude and interest required in each track (Center for Learning and Assessment Development-Asia, 2015). The model used the learning competencies of the Department of Education in the K to 12 curriculum guide for the achievement (following a Standards-Based Assessment approach). The taxonomy of aptitude items were used for the components of aptitude (see Magno, 2009). And the factors of the basic interest markers by Liao, Armstrong, and Rounds (2008), which is intended as public domain, can be used for the interest. The factors identified under each track provide a strong indicator of the skills needed for learners to be successful on each track. The factors for each track were decoded using a sample of students who participated in the early implementation of the SHS. Factor analysis was conducted among the component scores and the specific factors that highly loaded on a track ($>.40$) was the basis for the selection of the factor. The selected factors of achievement, aptitude and interest are shown below.

Table 1
Factors of Achievement, Aptitude, and Interest for each SHS Track

Tracks	Indicators		
	Achievement	Aptitude	Interest
Academic Tracks			
▪ Science, Technology, Engineering, and Mathematics (STEM)	Students entering this track should exemplify science and mathematics skills.	Show aptitude in syllogism, number and letter series, visual discrimination and series	Inclinations on life science, medical services, physical science, engineering, and technical writing
▪ Accounting and Business Management Track (ABM)	Students entering this track should exemplify English and mathematics skills with	Show aptitude in verbal analogy, number and letter series, visual discrimination and series	Inclinations on human relations management, personal service, business, management, sales, family activity, finance, and office work
▪ Humanities and Social Science (HUMSS)	Students entering this track should exemplify English and science skills.	Show aptitude in verbal analogy, syllogism, number and letter series	Inclinations on social science, creative arts, human relations management, personal service, politics,

				professional advising, religious activities, social service, teaching family activity, and religious activities
▪ General Academic Strand (GAS)	Students entering this track should exemplify English, math, and science skills.	Show aptitude in verbal analogy, syllogism, number and letter series		Inclinations on social science, creative arts, technical writing, human relations management, business, management, and social science
Sports Track	Students entering this track should exemplify English skills.	Show aptitude in verbal analogy, syllogism, number and letter series		Inclinations on physical or risk taking, athletic coaching, human relations management, personal service, professional advising, and teaching
Arts and Design Track	Students entering this track should exemplify English skills.	Show aptitude in visual discrimination and series, figure and ground perception, object assembly, and surface development		Inclinations on creative arts, creative writing, performing arts, religious activities, and social science
Technical-Vocational Livelihood Track				
▪ Agri-fishery Arts	Students entering this track should exemplify math and science skills.	Show aptitude in visual discrimination and series, object assembly, surface development.		Inclinations on manual labor, outdoor agriculture, physical/risk taking, protective, skilled trades, life science
▪ Home Economics	Students entering this track should exemplify English, math and science skills.	Show aptitude in visual discrimination and series, figure and ground perception, object assembly.		Inclinations on manual labor, outdoor agriculture, physical/risk taking, protective, skilled trades, family activity, human relations management, and personal service
▪ Information and Communications	Students entering this track should	Show aptitude in verbal analogy,		Inclinations on mathematics,

Technology	exemplify English, math, and science skills.	syllogism, number and letter series, visual series and discrimination, and object assembly.	creative arts, information technology, and office work
▪ Industrial Arts	Students entering this track should exemplify math and science skills.	Show aptitude in visual discrimination and series, figure and ground perception, object assembly, and surface development.	Inclinations on manual labor, outdoor agriculture, physical/risk taking, protective, skilled trades, and physical science

Note. Table taken from <http://cladasia2015.wix.com/cladasia#!senior-high-school-placement-test/c1cof>

The entry level assessment in SHS serves the function of assessment “as” learning. In this approach, assessment information is used by the learner to make decision in the improvement of their learning. In the case of the results of the entry level assessment, students use this information in order to decide on the appropriate track given what they have mastered overtime, what they can do in the future, and what they are interested about. This actually prevents students becoming undecided on the track they will take. Having assessment information help students match their strengths on the required tasks in the subjects they will go through in the SHS.

2. Classroom-Based Assessment

The accountability of the classroom-based assessment is directed to the quality of instruction and how it is supervised in the school setting. Instruction greatly accounts for students’ performance when assessment results are reported within the classroom level. Even when accountability of assessment result is high at this level, the impact of intervention is more feasible. Assessment results are easily used inside the classroom in order for the teacher to decide on who needs further scaffolding in the required competencies to be learned following the approach on assessment “for” learning.

The model for the classroom-based assessment is detailed in the memorandum provided by the Department of Education. The framework for the K to 12 describes assessment in two forms: Formative assessment and summative assessment. Formative assessment is emphasized as a way to improve students’ learning

(Hattie & Timperley, 2007). The purpose of formative assessment for the students and the teachers are emphasized. The students need to realize the value of formative assessment on how it helps them learn and at the same time the teacher needs to see the value of formative assessment to improve their instruction. The purposes of formative assessment are also cross-tabulated with the kind of assessment that is done before, during and after instruction. This shows that formative assessment is closely integrated with instruction where it is used as a tool to observe student progress. The study by Magno and Lizada (2015) has presented a model of formative assessment and a set of principles on how it becomes useful in developing teachers' perspectives and classroom practice. They have defined formative assessment in the classroom as a continuous process where interventions are provided based on learning targets and ultimately observing students progress towards this target overtime. There were nine principles of formative assessment provided, namely:

- (a) Formative assessment works along with the perspectives of assessment "for" and "as" learning;
- (b) Formative assessment is embedded with instruction;
- (c) Helping the students focus on the learning goal;
- (d) Diagnostic assessment on the target competency serves the function of formative assessment;
- (e) Formative assessment moves from determining discrete skills to integrated skills;
- (f) Using continuous and multiple forms of assessment;
- (g) Feedback practices using assessment results;
- (h) Working out with students to reach the learning goal; and
- (i) Deciding to move instruction to the next competency.

On the other hand, summative assessment is conducted when the students have mastered the required competencies. The results of the summative assessment are commonly included as part of students marking. Summative assessment provides information on how well the students have achieved the learning competencies that has been shaped through formative assessment and instructional interventions.

The benefit of classroom-based assessment is that both formative and summative assessment can be directly translated into improvement of instruction and eventually observe student progress overtime. Assessment within the classroom setting is largely flexible in

a way that the teacher decides on the appropriate form of assessment given to students.

3. Assessment of Achieved Competencies

The standard competencies learned by students can be assessed using national exams or summative assessments at the end of a school cycle. The basis of the national exam is the standard curriculum that is agreed and common for all schools. It is necessary to have a set of national standards in order to have common direction on what to teach in schools. Assessment of achievement actually refers to what students have acquired within the school year which is usually administered towards the end of the school year. These exams either come in the form of national assessment or assessment from private institutions. These exams, either developed by the government or by private institutions, all follow the same targets that come from the curriculum. These targets that are assessed in the curriculum are considered as standards.

National Assessment. The government of a country usually administers national assessment through its ministry of education. According to Kellaghan and Greaney (2001) that national assessments measure and monitor learning outcomes of a whole education system, or defined part of an education system. Usually the national assessment includes measures of competencies in different subject areas such as English, mathematics, science, social studies, and native language. Usually, the core subjects are covered in the examination if the assessment is a measure of achievement or what students have learned from the instruction of the curriculum. The assessment serves as accountability on students' progress overtime. The government through its public schools use the results of the national assessment to determine how effective is the curricular program, quality of instruction in schools, decisions on school systems improvement, and further training and programs to be implemented in the senior high schools. The national assessment that will be designed for the senior high school will be a good indicator for the implementation of the curriculum for the first time. The batch that will graduate for grade 12 on 2018 will be the first product of the new curriculum for the senior high school. The results of their end of cycle assessment at the end of

grade 12 can provide information on the areas of strengths and improvement about the intentions of the curriculum that was implemented within the span of six years starting from school year 2012-2013. Lessons learned will be derived from the results of the assessment of the first batch of graduates. The results of the assessment will be used to inform policy makers on the improvement that needs to be done with the senior high school curriculum.

Standards-Based Assessments (SBA). The practice involves comparing each student's performance to academic standards that are developed in the national or international setting. The goals are communicated to the school systems, schools, teachers and students that need to be achieved, providing targets for teaching and learning, and shaping the performance of teachers and students. National assessments are commonly standards-based, because the content domains involved in the assessment are based on the competencies include in the curriculum.

According to Magno (2015) that the SBA model requires that assessment practices needs to be closely aligned with the required competencies that students need to learn. The general intention of the SBA is to focus more on the learning of competencies of every student. These learning competencies are focused as the outcomes of instruction and directly assessed. For an assessment task to be standards-based, the task should be directly aligned with the students learning competencies. If the competency in chemistry is for students will explain the relationship between gas temperature and pressure, the assessment should require students to explain their conceptual understanding of the two concepts. If the task is a multiple choice, the stem of the item needs ask about the relationship between the two concepts and the options are all possible explanation. If the competency in English is for students to compose a five line poem, the assessment should be a performance-based task where students will write the required poem and the marking will use a rubric with a set of criteria. Likewise, if the mathematics competency is for students to add two similar fractions, a set of exercises will be provided to students

where they have to determine the sum of $\frac{3}{4} + \frac{2}{4} = x$ in a written work.

Constructive alignment ensures the direct assessment of the competency that is included in the curriculum. The quality of

assessment task (i. e. test, exercises), especially the content validity of the test, is ensured when the items directly measure the competency. Constructive alignment will work well when the competencies are clearly written as measurable standards in the curriculum.

4. Participation in International Benchmarking of Competencies

International benchmarking are those kind of assessment participated by different countries within a specific region that shares some commonalities in the educational standards or learning competencies and specific learning objectives. According to Clark (2011) that aside from looking at commonalities, cross-national variations on student learning, education systems and processes are also explored in international assessments. According to Lockheed (2010) that international assessments involves multiple countries, use standardised instruments, implementation and analyses, sample large student populations that are comparable across participating countries. Results of international assessment provide information on how far is the performance of a country as compared with other countries. The results of international assessment inform policy makers on the specific directions on educational reforms that needs to be made for a particular country. For example in the Philippine setting, when the K to 12 was initiated in 2010, the results of the TIMSS were used in order to justify the reforms needed in the basic education program.

Trends in International Mathematics and Science Survey.

In the previous years, the Philippines have participated in the Trends in International Mathematics and Science Survey (TIMSS) as part of international benchamarking studies in 1999 and 2003. The purpose of the TIMSS is to provide information on mathematics and science achievement of grade 4 and grade 8 students across different countries. TIMSS is sponsored by the International Association for the Evaluation of Educational Achievement (IEA). In 1999, there were 38 countries that participated in the TIMSS and the Philippines ranked 35th on science and mathematics. In 2003, there were 23 countries that participated for fourth graders and the Philippines ranked 23 on both science and mathematics. For the eighth graders, there were 46 countries that participated and the Philippines ranked 41st in Mathematics and 42nd in Science (nces.ed.gov/timss/).

Programme for International Student Assessment (PISA).

The PISA is an international survey that evaluates the educational systems on reading, mathematics and science. The assessment is administered to 15 year old students for more than 70 countries. The PISA is sponsored by the Organisation for Economic Co-operation and Development (OECD), and has been administered in three year cycles since 2000. The Philippines has not participated in the PISA, however, the results of the TIMSS is predicted by the results of PISA. In a study conducted by Care (2010) on the review on large scale assessments, she computed that scores on the TIMSS explains 21% of the variations on the PISA. This indicates that given the Philippines' score on the TIMSS, the predicted score on the PISA can be determined. The score of the Philippines in the TIMSS is relatively low, more likely, scores in the PISA will also be low.

When the Philippines participate in the TIMSS and PISA for the next cycle, information will be provided about the effectiveness on the educational policies implemented for the K to 12 program. The results should be able to provide policy makers with action on specific educational policies to improve the curriculum on K to 12.

5. College Readiness Assessment

The Commission on Higher Education at the onset of the K to 12 educational reform has prepared initiatives to support the smooth transition from basic education to higher education. Part of this transition is the proposed college readiness framework for the Philippines. College readiness is defined by Conley (2007) as “the level of preparation a student needs in order to enrol and succeed, without remediation, in a credit-bearing general education course at a postsecondary institution that offers a baccalaureate degree or transfer to a baccalaureate program” (p. 5). Success was operationally defined by Conley (2007) as “completing entry level courses at a level of understanding and proficiency that makes it possible for the student to consider taking the next course in the sequence or the next level of course in the subject area” (p. 5).

It is important to create a college readiness framework in order to share with basic education the expectations of higher learning systems from pre-university education. If the expectations of college

education are provided, the curriculum in the K to 12 will be able to be well aligned with the advance studies. College applicants will be better prepared for college given the set of expectations developed in basic education. The college readiness framework with a set of specific competencies for different subjects allows basic education to conform with international standards. The specific competencies indicated in the college readiness framework provide capacities for students to be prepared for college life.

The college readiness framework details the topics and competencies that a typical high school graduate need to have in the areas of science (biology, chemistry, physics, earth science), mathematics, English, Filipino, Literature, humanities, and social science. The college readiness provides a set of standards and the specific learning competencies should be used in developing items for assessment tools to assess college readiness. This will ensure that the contents of the test are well aligned with college readiness standards.

Given the competencies stated in the college readiness framework, there are two major implications on assessment in the transition from senior high school to college: (1) HEIs and schools offering grade 12 needs to diagnose readiness based on the given competencies of the college readiness framework; (2) Entrance exams needs to be built using the college readiness framework.

The assessment of college readiness at the end of grade 12 is viewed as an accountability of the school that gives information whether they have produced graduates that are ready for college or advance studies. The assessment results provide information whether students which have achieved the competencies in high school will be able to meet the necessary skill qualifications for college. The assessment of college readiness can be given either at the end of grade 12 or in the college level during the freshmen year. If the assessment of the college readiness is given at the end of grade 12, then the information serves as achievement of the students by high school. Serving as an achievement, it provides information if the students have attained the college readiness competencies across time. If the assessment is given at the start of college during freshmen year, it serves as a diagnostic test on identifying students strengths and weaknesses, what they know and do not know, and what students can do and cannot do so that the curriculum in college can scaffold the

necessary skills that the students are still weak at especially in subjects such as English, science, and mathematics.

Another implication of the college readiness framework is on the contents of the entrance exam. If the entrance exams intend to assess whether students have possessed the necessary characteristics for college, then the contents of the entrance exam should cover the college readiness standards on the different subject areas. It is important to include an assessment of the college readiness standards in the entrance exam since the competencies are prerequisite skills necessary to succeed in the general education subjects in college. The college readiness standards in the entrance should be good predictors of the grades in English, mathematics, and science subjects in college.

6. Career Assessment

Another area of assessment opportunity for the senior high school is on career assessment. Career assessment is conducted as part of career counseling or career guidance before students graduate in the basic education program. Career assessment can be both quantitative and qualitative (McMahon, Patton, & Watson, 2003). Quantitative assessment involves the use of psychometrics and standardized tools where quantitative results are interpreted on the dominant career suitable to the test taker. Likewise, this approach quantifies the characteristics as set of factors. On the other hand, qualitative assessment uses a constructivist approach focuses on the understanding of client's personal meaning where they make sense of their experiences overtime. The transition that occurs from grade school to the senior high school is supported as the students are empowered in the tracks if their choice. Some devises used for qualitative career assessment would be the use of card sorts, genograms, and lifelines (Neimeyer & Neimeyer, 1993).

The standardized tools that are used in career assessment are based on career development theories. Example of these theories that are translated to standardized test are the Holland's hexagonal model and Strong's career development theory. Holland's model features six traits that influences ones career choice: Realistic, investigative, artistic, scientific, enterprising, and conventional. On the other hand, the Strong Interest Inventory measures measures individual interest in six

areas: Occupations, subject areas, activities, leisure activities, people, and characteristics.

The assessment protocols that can be used in qualitative assessment are card sorts, genograms and lifelines. Card sorts are helpful in identifying interest of certain skills in a potential career, selecting a career or expanding career options, and adding insight to improve professional relationships. A genogram is a diagram where standard symbols are used to map at least three generations of the person's family system (Bakshi & Satish, 2015). The genogram can provide patterns of emotions, interpretations, and other family characteristics that can help the client decide on the career. Lifelines show the important events that occurred in the student's life. The progression of the lifeline shows upward and downward movements manifesting the experience of the client. The lifelines provide the direction of the individual's life and insights lead to reflection on the future direction of the client.

Career assessment can be conducted as a school-based assessment through the guidance program or part of a national career assessment program. In the school setting, career guidance is conducted that starts when the child is taught to look forward in their future vision of themselves. This process often involves self-awareness and self-exploration of skills, abilities and interests, understanding the requirements of the program that students intend to enter, and supporting students in their decision making skills (DepEd Order no. 41, 2015). These activities are implemented through the guidance office and integrated with various subject areas. In these activities both qualitative and quantitative assessment are conducted especially in the self-awareness and self-exploration phase. On the other hand, career assessment is also conducted in the national level. In the case of the Philippines the National Career Assessment Examination (NCAE) is administered to public school students. In the previous school years, the purpose of the NCAE was to improve the quality of fourth year high school graduates who will enter college. The results of the NCAE help students' to decide on the programs they will take in higher education institutions matching their aptitude and interests. The NCAE measures student aptitude on the following areas: Reading comprehension, clerical ability, mathematical ability, visual-manipulative skills, verbal ability, scientific ability, logical reasoning, non-verbal ability, and entrepreneurial skills. The private schools in the

Philippines also administer their own career assessment using standardized tools that measures equivalent constructs with the NCAE. In the senior high school, these career assessments supports the students by the end of grade 12, that they are ready to make their career decisions whether they will pursue higher education in universities, colleges and other advance courses, start working and become employed, and initiate their own business.

The aim of the enhanced basic education program (K to 12) is to develop holistic learners with 21st century skills. The addition of two years in high school provides more time and opportunity for students to develop the necessary learning competencies to be ready for their career paths. Assessment takes an important role in delivering students towards 21st century skills. Assessment at the beginning of the senior high school ensures that students select a track of specialization matching their interest, achievement, and aptitude. During grades 11 and 12, classroom-based assessment, assessment of achievement, and international benchmarking help students to acquire the necessary competencies that need to be learned. At this juncture assessment can serve to diagnose student difficulties to provide more appropriate instruction, help students progress overtime through formative assessment ensuring and ensuring proficiency by the time the summative, national, and international assessment are conducted. Assessment also plays an important role to deliver students in their career path after senior high school. Assessment ensures that students acquired the necessary learning to enter college through college readiness assessment. Assessment also ensures their qualifications to work through assessment of employment readiness. The specific programs that they pursue after the senior high school are further clarified through career assessment.

Given the levels of assessment identified in the senior high school level, the following recommendations are derived:

- (1) Provide classroom teachers with the specific results of the assessment that will serve as diagnostic in order to provide instruction based on the students learning needs.
- (2) The periodic assessment conducted in the classroom level should also allow teachers to address immediately students who are at risk of not learning a competency.
- (3) Use the various levels of assessment as indicators in evaluating the senior high school curricular program.

(4) Schools need to consider the competencies covered in the national and international assessments when designing their curriculum. Being able to see the similarities and differences of the schools' curriculum with the national and international competencies allows the school to form students based on important standards.

References

- Akker, V. (2003). Curriculum perspectives: An introduction. In V. Akker, W. Kuiper, & U. Hamyer (Eds). *Curriculum landscapes and Trends* (pp. 1-10). NY: Springer.
- Bakshi, A., & Satish, V. (2015). Qualitative career assessment using a genogram. In M. McMahon, & M Watson (Eds.), *Career Assessment* (pp. 69-80). India: Sense Publisher.
- Barret, J. (2011). *The aptitude test workbook*. London: Kogan Page.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-71.
- Center for Learning and Assessment Development-Asia (2015). The Senior High School Placement Test. [Online Available] cladasia2015.wix.com/cladasia#!senior-high-school-placement-test/c1cof
- Clarke, M (2011). *Framework for Building an Effective Student Assessment System READ/SABER Working Paper*. Washington DC: The World Bank Group.
- Corsi-Bunker, A. (2009). *Guide to the educational system of the United States*. University of Minnesota, Minnesota: International Student and Scholar Services.
- Conley, D. T. (2007). *Redefining college readiness*. Eugene, OR: Educational Policy Improvement Center.
- DepEd Order no. 8, s. 2015. Policy guidelines on classroom assessment for the K to 12 basic education program.
- DepEd Order no. 41, s. 2015. Senior high school career guidance program and early registration.
- Republic Act No. 10533 (2013). Enhanced Basic Education Act of 2013.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Ho, E. S. (2012). *Student learning assessment: Asia-Pacific Education System Review Series No. 5*. Bangkok, Thailand: UNESCO Bangkok.

- Holland, J. L., Fritzsche, B. A., & Powell, A. B. (1997). *The self-directed search technical manual*. Odessa, FL: Psychological Assessment Resources.
- Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*. Sweden: UNESCO, International Institute for Educational Planning.
- Liao, H., Armstrong, P., & Rounds, J. (2008). Development and initial validation of public domain Basic Interest Markers. *Journal of Vocational Behavior*, 73, 159-183.
- Lockheed, M. 2010. The craft of education assessment: Does participating in international and regional assessments build assessment capacity in developing countries? NJ: Princeton.
- Magno, C. (2008). Taxonomy of aptitude test items: A guide for item writers. *The International Journal of Educational and Psychological Assessment*, 2, 39-53.
- Magno, C. (2011). *Analysis of the basic education in the Philippines*. Manila, Philippines: SEAMEO INNOTECH.
- Magno, C. (2009). Taxonomy of aptitude test items: A guide for item writers. *The International Journal of Educational and Psychological Assessment*, 2, 39-53.
- Magno, C. (2015). Editorial Note: On Standards-Based Assessment. *Journal of SBA Research*, 3, 1.
- Magno, C., & Lizada, G.S. (2015). Features of classroom formative assessment. *Educational Measurement and Evaluation Review*, 6, 23-31.
- McMahon, M., Patton, W., & Watson, M. (2003). Developing qualitative career assessment process. *The Career Development Quarterly*, 51(3), 194-202.
- Mikre, F. (2010). The roles of assessment in curriculum practice and enhancement of learning. *Ethiopian Journal of Education and Sciences*, 5(2), 101-114.
- Neimeyer, G., & Neimeyer, R. (1993). Defying the boundaries of constructivist assessment. In G. Neimeyer (Ed.), *Constructivist assessment: A casebook* (pp. 1-30). Newbury Park CA: Sage.
- SEAMEO INNOTECH. (2015). K to 12 toolkit part II: Resource guide for the senior high school program in the Philippines.
- Trends in International Mathematics and Science Survey. Evaluation of Educational Achievement. Retrieved from nces.ed.gov/timss/

Trey, S., Schimitt, V., & Allen, J. P. (2012). Defining authentic classroom assessment. *Practical Assessment, Research, and Evaluation*, 17(2), 1-18.



Assessment Systems in Southeast Asia: Models, Successes and Challenges

Educational Research Unit
SEAMEO INNOTECH

The demand for new knowledge and abilities to provide students the opportunity to develop their 21st century skills have driven nations in the Southeast Asian region to focus on ensuring that students learning can be improved and assessed effectively. The global market signaled education systems of nations to raise further the bar on quality and access to education, prompting them to work together and learn from each other.

According to Darling-Hammond and Wentworth (2010), in this era, students must possess skills which include “the abilities to find and organize information to solve problems, frame and conduct investigations, analyze and synthesize data, apply learning to new situations, self-monitor and improve one’s own learning and performance, communicate well in multiple forms, work in teams and learning independently.” Clearly, it is a challenging task for nations, especially those in the Southeast Asian region given the economic and political hurdles that they equally give importance to assessment of learning. SEAMEO INNOTECH and representatives from ten of the eleven member countries of the Southeast Asian Ministers of Education Organization (SEAMEO) – Brunei Darussalam, Cambodia, Indonesia, Malaysia, Myanmar, the Philippines, Singapore, Thailand, Timor-Leste and Vietnam, convened a regional workshop in 2012 to learn from each other, particularly in the area of assessment of students as a critical tool in maximizing learning and teaching. The workshop revealed that while curricular reforms have been periodically undertaken and are well articulated in all the member countries of SEAMEO, the assessment systems of each country need to be further strengthened and firmly established by instituting reforms by learning

from other another. SEAMEO member countries have taken bold steps, initiating discussion and work groups to share and understand how assessment systems in their own countries could be more relevant and responsive to the demands of the 21st century.

Assessment: Definition, Purpose and Scope

Learners' assessment has been extensively studied within the educational setting. Harlen (2008) and Mertler (2009) defined assessment as the process that teachers use to assign grades to students in particular subject assignments. Assessment can also refer to the standardized testing imposed in schools (Marzano, 2006; Stiggins & Chappus, 2005). Also, assessment has been described by Black and William (1998) as any activity intended to gather information to be used to provide feedback in order to modify teaching and learning activities in the schools or improve instruction and students' performance as discussed by Cohen and Hill (2000).

Assessment and examinations play a very important role in all education systems. Educational institutions put much importance on it as it is a means to gather information to improve learning and instructional practices. High performing nations particularly give special attention to assessment as seen in their consistent high ratings in international achievement tests which give premium to the full potential of student assessment.

From the various definitions shared by the regional workshop participants, SEAMEO member countries generally agree that assessment is a process and procedure of collecting and understanding data or information from students in a range of activities aimed at improving teaching-learning process and making decision and judgment on students' learning outcomes.

As for identifying the purpose of assessment, Gipps and Cummings (2003) argued that no assessment is considered good or bad as it is judged on how it satisfies its intended purpose. Based on the various descriptions provided by the workshop participants, it can be surmised that the purposes of assessment in each country are evidently anchored on its education strategic plan and/or agenda. For example, the Philippines' assessment system touches on the BESRA (Basic Education Sector Reform Agenda) while Singapore and Brunei are anchored on SPN21 (Sistem Pendidikan Negara Abad ke-21 (National

Education System for the 21st Century) and curriculum initiatives, respectively. Hence, the assessment programs of SEAMEO member countries, particularly those at the school level, are meant to improve – processes of learning; processes of instruction; outcomes of learning and outcomes of instruction.

Most SEAMEO countries, excluding Brunei Darussalam and Singapore, conduct assessment using various measures and tools and examinations in all levels of education and school system from pre-primary, primary higher education which covers core, non-core and non-cognitive areas. The scope of assessment is either school-based or national examinations which are also customized based on the curriculum and standards of the respective nation. Brunei Darussalam and Singapore's assessment programs start in the primary level (SEAMEO INNOTECH, 2015).

Assessment for Learning or *A/L* refers to the practices of teachers and education to carry out assessment aimed to determine progress in learning by giving tests and other tools to measure learning while the instructional program or process is going (Murray, 2006; Sparks, 2005). Countries such as Brunei Darussalam, Cambodia, Indonesia, Malaysia, Philippines and Singapore are stressing *A/L* as important purposes of their assessments.

Assessment of Learning (*AoL*) or commonly known as summative assessment pertains to the practices of teachers and school system to conduct assessment to determine the current standing of students' achievement against learning outcomes defined in the curriculum and in some cases, how they are placed in relation to others (Earl, 2005; Harlen, 2008). Countries such as Brunei Darussalam, Indonesia, Malaysia, Myanmar, Philippines, Singapore, Thailand and Vietnam all highlighted this as a purpose of assessment.

Assessment as Learning (*AaL*) is the use of on-going self-assessment by students in order to monitor their learning (WNCP, 2008). The Philippines, Singapore and Thailand are promoting assessment as learning. With the view of the learning process, learners or students are the critical corrections between assessment and learning (Earl & Katz, 2006).

Putting Large Scale and School-Based Assessment Results for Development

As an integral part of the curriculum in each member country, assessment in the Region may be categorized into Large Scale Assessment and School-based Assessment. Most large scale assessments are examinations given to measure completion and/or exit from one level to another and entry to a higher level in the education ladder. An example is the test given from Primary to Secondary and/or from Secondary to Higher Education. The Ministry of Education or Department of Education is usually tasked to oversee the planning, management and supervision of these assessment systems.

To complement the large scale national assessments, Southeast Asian countries also have School-based Assessment (SBA) systems carried out by teachers in schools to their own students or learners. The SBA is formative and diagnostic in nature which is intended to improve students' learning outcomes as well as provide immediate feedback to improve quality of learning, teaching and assessment. It is typically designed and implemented by students' own teachers, making it a highly valid form of assessment since teachers exactly know what and how to measure their students' learning.

During the workshop, a survey of strategies used in conducting assessments in various levels was conducted. The results of the survey showed that the strategies used in the region included the following: anecdotal records; written quizzes; worksheets/seat works; assignments/ projects; oral quizzes/recitations; observation checklists; portfolio assessments; performance assessments/ demonstration; peer assessment; self-assessment; and team assessments.

Each assessment system defines a specific role and use of the results of the assessment process. It also indicates different roles of government, the school, the teachers, the students, as well as the parents. The report underscored that SEAMEO member countries have taken significant strides to make meaningful use of assessment data and communicate these results to various stakeholders. Most of the countries agree that the main use of test data is to determine the level of competencies of learners. Specifically, national examinations are used to gauge students' aptitude and readiness in proceeding to a higher level of schooling. The test results also help educators, policy

makers and teachers design a more appropriate and responsive instructional program.

While assessment of students is a major concern, SEAMEO member countries put much premium on the professional development of teachers who are the key resources of the assessment system.

There is a general consensus of the need to further improve professional development of teachers in the area of assessment in both pre-service and in-service trainings as well as exposing and allowing them to participate in seminars, conferences and workshops. These development programs for teachers give them the opportunity to can share and gain new knowledge on innovations and best practices in learner assessment along with upgrading their skills in using new technologies as a means of providing assessment to students.

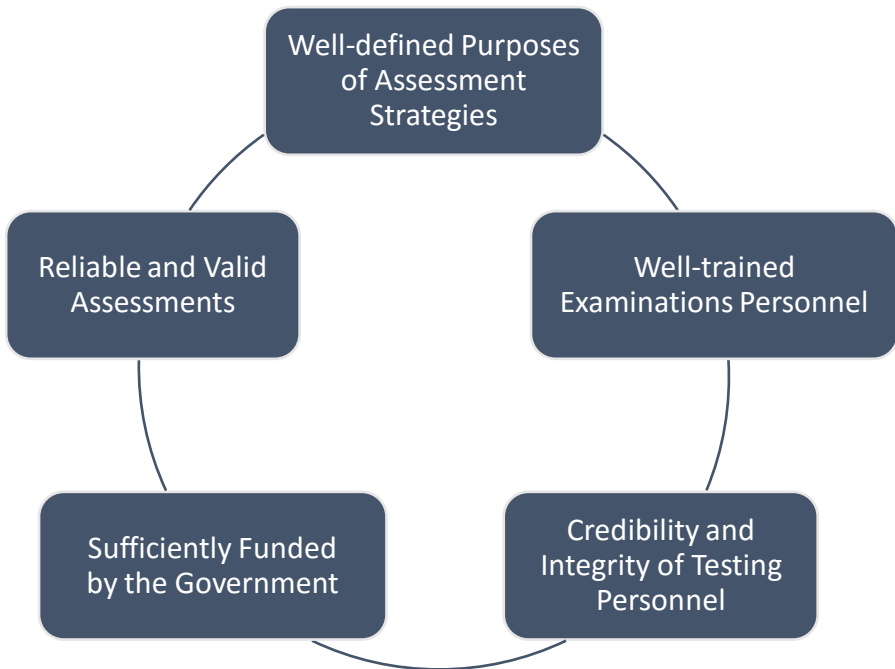


Figure 1. Elements of Good Assessment Practices

SEAMEO member countries identified five elements of good assessment practices (see Figure 1): i) well-defined purposes of

assessment strategies; ii) well-trained examination personnel; iii) reliable and valid assessments; iv) sufficiently funded by the government; and v) credibility and integrity of testing personnel is in place (SEAMEO INNOTECH, 2015).

Taking Steps towards Assessment Reforms

SEAMEO member countries are increasingly focused on embracing a paradigm shift in terms of assessment purposes and approaches. Realizing truly the integration of assessment system into the instructional system, member countries represented at the workshop agreed that shifting to Assessment for Learning (*AfL*) and Assessment as Learning (*AaL*) from Assessment of Learning (*AoL*) is the way to go.

In Brunei Darussalam, an enhanced assessment system has been put into place by strengthening its integration into the teaching and learning process. As for Cambodia, the focus of assessment reform is in supporting the achievement of quality and efficiency in the delivery of education in the country which include, among others, the systematization of integration of assessment results into the grading system. The introduction of internal and external assessment in 2005 marked the reform of Indonesia's assessment system while Malaysia initiated its Holistic Assessment System (HAS) in 2011 which is referred to as National Education Assessment System for Primary and Lower Secondary Students.

On the other hand, Myanmar's assessment reform highlighted quality assurance by enhancing administrative processes aimed at improving quality of teachers and education of students. As for the Philippines, it is establishing a National Assessment and Grading System Framework which underscores the role of national examinations at each strategic stage of schooling, among other reforms, which are all anchored on the new K to 12 curricula. Singapore's Ministry of Education (MOE), initiated the PERI Holistic Assessment in 2010, a significant assessment reform in the country to recognize and sustain the good assessment practices that improve the quality of learning and teaching in schools.

For other countries – Thailand, Timor-Leste and Vietnam, reforms have centered on strengthening existing assessment systems and ensuring that the highest quality of assessment system is in place.

Striding Forward to Success In the Face of Issues and Challenges

So far, assessment systems implemented by countries in the Region have been relatively successful due to various reforms and improvement strategies that have made significant impact on their whole education system. Among the identified success factors include:

1. The importance of stable organizational structure that promotes sustainable programs, including research and development;
2. Supportive government and related agencies to implement the national assessment system; and
3. Well-crafted assessment frameworks and implementation mechanisms supported through legislation and education laws.

In spite of the bright prospects for assessment systems in SEAMEO member countries, issues and challenges continue to arise. As such, the challenge to establish an assessment system of integrity which is accepted by all examination stakeholders is a common challenge among countries like Brunei Darussalam, Indonesia, Myanmar, Philippines and Singapore. Alongside this is the urgent need to address the recurring problem of cheating among some students during examinations.

Another challenge being faced are the doubts by some members of the public on the integrity high stakes examinations and the manner they are undertaken. Moreover, countries oftentimes find it difficult to ensure that an assessment program's purpose is well communicated to all its users. Countries are also sometimes confronted with a poor management of examination activities resulting in wastage of resources, mishandling of test materials and leakage and miscommunication, which undermine the reliability of assessment, as well as a lack of highly trained professionals in the field of education assessment.

Future Directions and Recommendations

The review of the assessment systems in the Region through the workshop initiated by SEAMEO INNOTECH generated the following recommendations for further strengthening of assessment systems:

1. Develop programs to help reduce the examination pressures faced by students which otherwise encourage a focus on passing the test rather than learning.

2. Create an assessment system that is responsive to the diverse needs of students such as inclusive education and student-centered curricular and instructional programs.

3. Craft an assessment system that covers a wider range of curriculum objectives and learning outcomes which promote critical and higher order thinking skills among students.

4. Ensure the balance between summative and formative assessments in the purposes of assessments.

5. Implement a programmed capacity building and professional development program in the region in lieu of the demand for more experts in the field of educational assessment.

6. Explore introduction of technology-supported assessments where appropriate.

7. Develop an assessment policy framework with assured funding support from the government.

SEAMEO member countries are committed to working hand-in-hand to support each country's educational reforms particularly in improving assessment systems. While Western and European counterparts have taken greater strides, Southeast Asian nations are evolving and catching up with the advances in this field notwithstanding the hurdles in streamlining and emulating the best practices in educational assessment.

Author's Note: The earlier version of this article was published as research brief for SEAMEO INNOTECH Research Updates. The full report is available at http://www.seameo-innotech.org/wp-content/uploads/2016/08/SIREP_Assessment-151021.pdf

References

- Black, P., & William, D. (1998). Assessment and classroom learning, *Educational Assessment: Principles, Policy and Practice*, 5(1), 7-74.
- Cohen, D. K. & Hill, H.C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teacher's College Record*, 102, 294-343.

- Darling-Hammond, L., & Wentworth, L. (2010). *Benchmarking learning systems: Student performance assessment in international context*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.
- Earl, L. (2005 Spring). Classroom assessment: Making learning paramount. *Mass journal*, 5(1), 14-15.
- Earl, L., & Katz, S. (2006). *Rethinking classroom assessment with purpose in mind*. Western and Northern Canadian Protocol for Collaboration in Education. Retrieved from www.wncp.ca
- Gipps, C. V., & Cummings, J. J. (2003). *Assessing Literacies*. Paper presented at the International Literacy Conference: Literacy Policies for the Schools We Need, Toronto.
- Harlen, W. (2008). Trusting teachers' judgment. In S. Swaffield, *Unlocking assessment*, (pp. 138-153). Abingdon, OX Routledge.
- Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA: ASCD.
- Mertler, C.A. (1998). *Classroom assessment practices of Ohio teachers*. Paper presented at the 1998 Annual Meeting of the Mid-Western Educational Research Association, Chicago, Illinois, October 14-17, 1998.
- Murray, S. (2006). *The role of feedback and assessment in language learning*. Rhode University, Grahamstown.
- SEAMEO INNIOTECH (2015). *Assessment systems in Southeast Asia: Models, successes, and challenges*. Quezon City: Author.
- Sparks, D. (2005). *Learning for results*. Thousand Oaks, California: Corwin Press.
- Stiggins, R. J., & Chappuis, J. (2005). Using student-involved classroom to close achievement gaps. *Theory into Practice*, 44(1), 11-18.
- WNCP (2008). Classroom Assessment. Retrieved from www.wncp.ca