Volume 11 | October 2010

Philippine Journal of Educational Measurement

A Publication of the Center for Educational Measurement, Inc.

IN THIS ISSUE

Center for Educational Measurement, Inc.

Copyright © 2010 by the Center for Educational Measurement, Inc.

All rights reserved. Except for the quotation of short passages for the purpose of review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means: electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the Center for Educational Measurement, Inc.

Published by
Center for Educational Measurement, Inc.
24th Floor Cityland Pasong Tamo Tower
2210 Chino Roces Avenue, Makati City
www.cem-inc.org.ph

Layout by
Frances F.M.B. Cardona

ISSN: 1656-8575

Table of Contents

Complimentary copy
for the School
President

THE EFFECT OF THE LANGUAGE OF THE TEST ON STUDENTS' PERFORMANCE IN MATHEMATICS

Ong Saw Lan
School of Educational Studies
Universiti Sains Malaysia

Assessment of content areas such as science and mathematics in second language raises the issue of validity. Stansfield (1996) suggested the way to remove the language barriers posed by tests written in the English language is to test the content of the assessment in the students' native language. The aim of this paper is to examine the effect of using Malay (native language) test items on students' mathematics performance. A comparison of item difficulty is also carried out to investigate the effect of language used on test characteristics. Based on the Item Response Theory item difficulty indices, six of the mathematics test items seem to be easier in English. However, three items in Malay were found to be easier than the English mathematics items. The results obtained from the item characteristic curves or ICCs were slightly different from that of the difficulty indices. Six of the mathematics items in Malay were shown to have higher probability of answering correct for all abilities group. Overall comparison using the mean test score also indicate that the Malay version of the mathematics test is slightly better than the English version.

Language Policy in the Malaysian Education System

The Malaysian government has given emphasis and special priority in the improvement of school science and math programs. Realizing the importance of the English language in the advancement of knowledge for science and mathematics, the government announced in 2002 a policy that uses English as the language of instruction in the teaching and learning of science, mathematics and technology subjects for the national education system. This policy was carried out in stages, starting in 2003 for the Year 1 (grade 1) at the primary level, Form One (grade 7) at the secondary level, and Lower Six (grade 11) at the post-secondary level. The other subjects of Humanities and Arts continue to be taught in Malay language for all national secondary schools.

The current national education system comprises six years of primary schooling (grades 1 to 6), five years of secondary schooling (grades 7 to 11), and two years of post secondary schooling (grades 12 & 13). For national unity, Malay language is the national language and medium of instruction. English is a compulsory subject in both primary and secondary schools. To meet the needs of multi-ethnic groups in the population, Mandarin or Tamil is also used as medium of instruction at primary levels. There are three major types of national primary schools in the country to meet the aspirations of the ethnic groups: (1) National (Malay) Primary Schools, (2) National-Type (Chinese) Primary Schools and (3) National-Type (Tamil) Primary Schools. At the secondary level, the primary language of instruction is Malay for the national secondary schools (Kee, 2007).

Language of Instruction and Language of Assessment in Science and Mathematics

The primary and secondary levels follow a national curriculum that prepare Malaysian children for three high-stake, common public examinations: (1) the Primary School Assessment Test, (2) the Lower Secondary School Assessment (LSSA), and (3) the Malaysian Certificate of Education.

To address the effective implementation of the above policy, the English Language Teaching Center (ELTC) developed the "English for the Teaching of Mathematics and Science subjects Workshop" (ETeMS) to facilitate the training of Mathematics and Science teachers. In a study to look at the implementation of this policy, Kee (2007) looked at the teaching-learning processes in the Mathematics and Science classrooms and found that teachers mix Malay or Mandarin into English or alternate between the two languages to facilitate students' comprehension of the concepts. Kee has the opinion that the use of code switching and code mixing may be due to teachers who are incompetent and are unable to speak English or pronounce terminologies correctly. Similar concern has been raised by parents at the initial stage of this policy when incompetent teachers mixed native languages (Malay, Mandarin or Tamil) with English or alternating between native languages and English in the classroom (the New Straight Times, 2003). Teachers' ability to transfer the cognitive demands of the contents is questioned when they are not competent in the English language.

In a survey conducted by Long (2005) regarding the change in language policy for science and mathematics, she found that 60% of Mathematics and Science teachers said they are not fluent in the English language while only 45% are comfortable in using English to teach. As for the students, non-Malay students (Chinese and Indians) in urban areas preferred learning Mathematics and Science in English while their rural Malay counterparts were not enthusiastic in the use of the English language.

From 2003, the mathematics and science subjects were tested in bilingual test booklet to help candidates in understanding the test items during the transition period of the policy. The Malaysian Examination Syndicate (MES) was given the task in the development, production and administration of bilingual assessment (Samad, 2003). The bilingual tests presented the test items in the Malay language side by side with the English version. According to the Ministry of Education, the assessment instruments will be totally in English by 2008.

Content Assessment in Second Language

Assessment of content areas such as science and mathematics in second language raises the issue of validity. To overcome this problem, the MES prepared all paper-and-pencil assessment of mathematics, science, technical and technology subjects for the national public examinations in dual-language since 2003. The MES hoped that by putting the Malay items side by side with the English version, examinees will be able to understand these questions better.

A survey conducted by the MES found that despite learning science and mathematics in English for three years, only 33% of the science candidates and 27% of the mathematics candidates used English when answering the Lower Secondary School Assessment examination questions. Most of the candidates preferred to answer in either Malay or a mixture of both languages (The Star, 2005). Those who answer in Malay however used the mathematics and scientific terms in English as they are familiar with the terminologies after having studied both subjects in English for three years. Students in urban areas indicated that they preferred the use of English in the test. Prior to 2008, students continued to have the option to answer either in English or Malay or a mixture of both. By 2008, the Ministry of Education (MoE) will make a decision regarding the use of English only as the language of assessment for science and mathematics.

Prior to making any decision about this issue, the MoE may need to heed the advice on language issues that are fundamental to exemplary mathematics teaching and assessment. Lager (2006) warned that without a strong command of both everyday language and specialized mathematical language in English, students cannot fully access the mathematics content of the text, lesson, or assessment item.

Literature on Problems and Issues of Testing in Second Language

The purpose of testing should be to obtain information that will be useful for making instructional decisions and determining the extent of student progress in the curriculum of the school (State of Iowa, DoE, 2007). For examinees to demonstrate the knowledge and ability in any subject-matter, the majority of assessment procedures is dependent on the ability to read and answer questions (Goldsmith, 2003). Abedi (2006) noted that the learning and assessment of English Language Learner (ELL) may be inhibited by their weak foundation in English and the linguistic complexity. Language ability is not the intended construct in assessing mathematics and can adversely affect the reliability and validity of such content area assessment.

The impact of the language factor among ELLs was even more acute for mathematical items that were linguistically packed such as mathematical word problem items (Dolan, Murray & Strangman, 2006; Anderson, 2007). As a solution for ELL, Stansfield (1996) suggested the way to remove the language barriers posed by tests written in the English language is to test the content of the assessment in the students' native language.

During the transition period of the language policy, bilingual assessments would be administered. By 2008, the MoE may decide on the use of English as the language of assessment since it is the language of instruction for all levels of education for Mathematics, Science, Technical and Technology subjects. The question arises as to whether English should be considered as the appropriate language of assessment for these subjects, or if there is a need to continue with the use of the Malay language. According to Abedi (2006), translating test items from English to other languages may not help the ELL if the language of instruction is English. Lord (2004) has the same opinion and believed that items translated to the native language may confuse students who had learnt the concepts in the English language as they may not be familiar with the terminologies used.

Test Accommodation Using Dual-Language Test Booklets

Presently, all the students participating in the Malaysian national level mathematics and science tests at grades 6, 9 and 11 are given the dual-language test booklets. No effort of any form has been carried out to identify students as non-English proficient, limited English proficient or English proficient when accommodation is being administered.

Students who are non-English proficient but proficient in the Malay language may use the dual-language test booklet to reduce the effect of English language deficit that would interfere with their achievement in mathematics. In the case of students who are not proficient on both English language as well as the native language, accommodation with the use of dual-language test booklets may not be appropriate in determining students' achievement in mathematics. They should be given the opportunity to engage in the test with alternative accommodations to allow them the opportunity to demonstrate their knowledge and abilities on what is being tested. Decisions about alternative accommodations for assessment may be necessary with the use of data on the student's level of language proficiency in English. Accommodations should only be used to reduce the effect of the student's English language deficits that would interfere with obtaining accurate information about the student's achievement. Alternatives such as multiple measures or formats, (use of objective items, short response items) may be made available to better assess students who lack linguistic proficiency.

In Malaysia, bilingual test booklet has been used to address the language factor in national level mathematics assessment. However, literature is lacking in showing how language proficiency in English and Malay language affects assessments. Due to this, the aim of this paper is to compare the effect of using Malay (the primary language of instruction) and English (the language of instruction) on students' mathematics performance.

Two questions that will be addressed in this study are as follows:

1. Does test language affect performance on mathematics achievement?
2. Does language used affect the difficulty of mathematics test items?

Method

This study uses the single group design; that is the same examinees are administered the mathematics test in both Malay and English.

The 2005 LSSA mathematics test was used for data collection. This paper-and-pencil test is available in dual-language, Malay and English as prepared by the MES. The two language versions of the test utilize similar mathematics items. The MES has prepared these items in one test booklet, with the Malay items on the left-hand side and the English items on the right-hand side of the booklet. All items selected in this study are multiple-choice format. Numerical values in these questions were modified so as to reduce student familiarity with test question as schools have access to the previous year's test questions.

An example of an item that has been modified is as follows:

The original English version is:
Find the lowest common multiple (LCM) of 8, 12 and 15

A 30
B 60
C 120
D 240

The modified item is:
Find the lowest common multiple (LCM) of 3, 8 and 16

A 32
B 48
C 128
D 242

Altogether there were 40 items covering topics on numbers, algebra, measurement, geometry, and statistics. Details of the items according to the topics are as shown in Table 1.

The dual-language mathematics test booklet was separated into two booklets; the first one with only the Malay version and the second booklet consists of the English test items. Adopting the single group design, all

Table 1
Test Item Number(s) in Each Topic of the Mathematics Test

Mathematics Topics	Item Number
Whole Number	5, 37
Factor	1,2, 3,4
Percentage	6
Basic measurement	39
Angles and lines	11, 13, 15
Polygon	7, 10, 14
Perimeter and area	16, 17, 23
Solids and volume	18, 19
Linear equation	29
Geometry construction	8
Pythagoras Theorem	9
Ratio and rate	31, 35
Coordinate	28, 30
Two dimensions Locus	12
Circle	20, 21, 22, 24, 40
Transformation	25
Statistic	26, 32, 34, 36
Linear non-equality	27
Graph Functions	33, 38

students were administered the Malay version of the mathematics test first. Three weeks later, the English version of the mathematics test was administered again to the 505 Form Four (grade 10) students from four schools in the district of Larut Matang and Selama, Perak. All students involved in this study were regarded as proficient in both the Malay and English languages as identified by the District Education Office.

Data Analysis

Students' performance in the mathematics achievement test was determined by computing their raw score obtained in answering the items correctly. For each student, two scores were obtained. One for the Malay mathematics test and another score for the English mathematics test. The difference in performance was compared using the score distribution for the two versions of the test.

Item difficulties were computed on each item for the different language versions. This helped to determine the effect of language on the difficulty of test items. The difficulty of test items were compared using the difficulty index of the test item calibrated using the computer program WINSTEPS which is based on the Rasch model. The difference in difficulty index between the Malay item and that of the English item is significant if t 1.96 (p=.05). In addition, test item difficulty using percentage correct was also computed.

The item difficulties were also compared by drawing each item characteristic curve (ICC) which represents the probability of answering each item correctly in a graphical form. ICCs for the English and Malay versions were plotted on the same axis to compare the differential probability of giving the right response for all ability groups. The area between these two ICCs differ by more than 0.20 units showing significant difference in the two language versions of test items (Raju, 1990).

Results

Table 2 shows the item difficulty, *b*, analysed using WINSTEPS, for the mathematics tests in English and Malay. Six of the mathematics test items (Item 6, 9, 18, 20, 23, and 25) seem to be easier in English compared to the Malay version of the mathematics test. Only three items (Item 13, 32, and 33) are easier when tested in Malay. For the rest of the 31 items,

Table 2
Item Difficulty Using Rasch Model for English Version and Malay Version

Item No.	b_R English	b_F Malay	T	df	Favor
6	.74	1.12	-2.32*	927	E
9	-2.28	-1.59	-2.10*	934	E
13	1.43	.96	2.99*	938	M
18	.07	.55	-2.78*	935	E
20	.12	.58	-2.72*	938	E
23	-.18	.30	-2.65*	935	E
25	1.13	1.70	-3.67*	936	E
32	.33	-.61	4.94*	933	M
33	2.49	2.12	2.32*	929	M

* p = .05

there are no significant differences whether the items are administered in Malay or in English (Appendix A).

As shown in Table 3, analysis of the percentage correct using classical test theory again showed that items 13, 32, and 33 appear to be easier in Malay. However, only item 25 seemed to be easier in English.

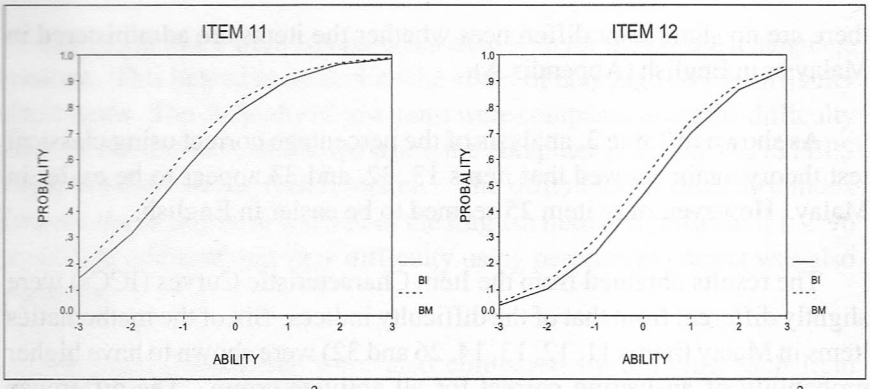
The results obtained from the Item Characteristic Curves (ICCs) were slightly different from that of the difficulty indices. Six of the mathematics items in Malay (items 11, 12, 13, 14, 26 and 32) were shown to have higher probability of answering correct for all abilities group. The difference between the areas of these two characteristics were especially noticeable for item 32. Six items (items 6, 9, 18, 20, 23 and 25) were also found to have higher probabilities of success when they are presented in the English language.

All six items that were identified easier based on the difficulty indices were also shown to have higher probability of being correctly answered by

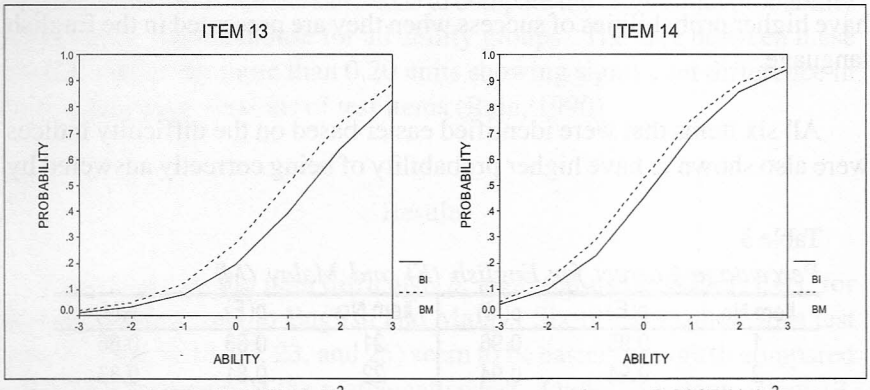
Table 3

Percentage Correct for English (E) and Malay (M)

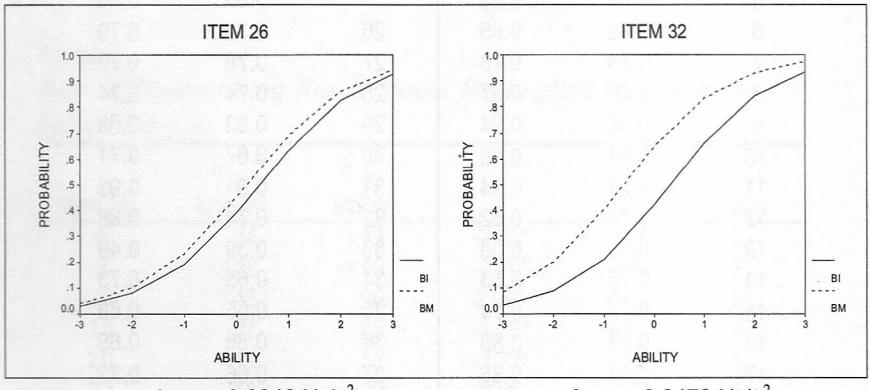
Item No.	p(E)	p(M)	Item No.	p(E)	p(M)
1	0.95	0.96	21	0.63	0.66
2	0.94	0.94	22	0.81	0.83
3	0.96	0.97	23	0.81	0.78
4	0.75	0.76	24	0.72	0.72
5	0.97	0.98	25	0.63	0.56
6	0.68	0.65	26	0.73	0.79
7	0.78	0.78	27	0.78	0.79
8	0.86	0.87	28	0.74	0.74
9	0.96	0.94	29	0.63	0.68
10	0.94	0.96	30	0.67	0.71
11	0.91	0.94	31	0.91	0.93
12	0.77	0.82	32	0.74	0.88
13	0.58	0.68	33	0.39	0.49
14	0.76	0.83	34	0.65	0.73
15	0.72	0.74	35	0.65	0.68
16	0.88	0.89	36	0.86	0.89
17	0.83	0.85	37	0.66	0.72
18	0.78	0.74	38	0.71	0.73
19	0.67	0.69	39	0.71	0.77
20	0.78	0.74	40	0.71	0.72



Area = 0.2228 Unit ^{2} Area = 0.22625 Unit ^{2}

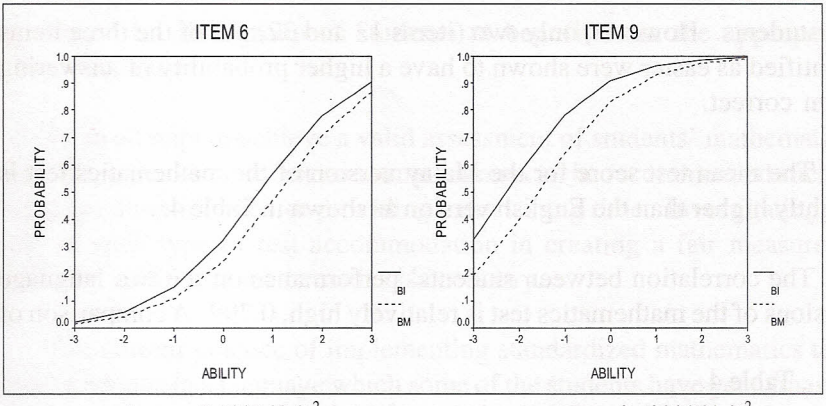


Area = 0.3962 Unit ^{2} Area = 0.2805 Unit ^{2}

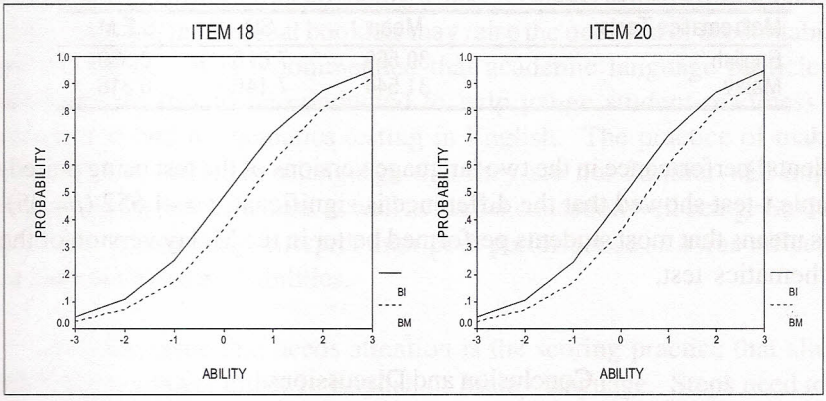


Area = 0.2342 Unit ^{2} Area = 0.8472 Unit ^{2}

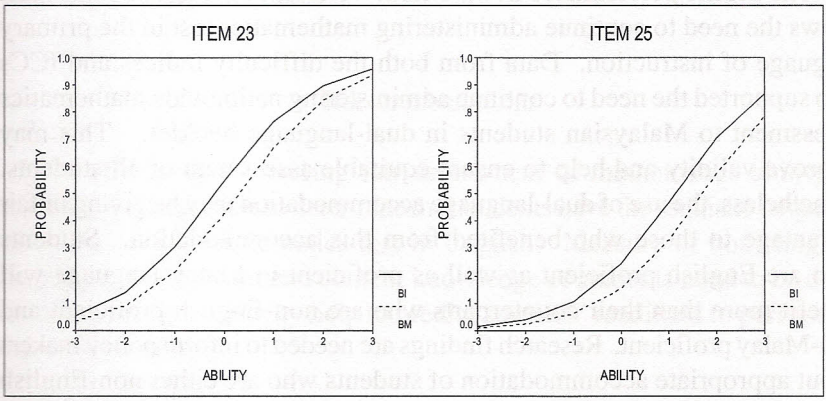
Figure 1. ICC difference with Malay items easier than English (area between ICC > 0.2 unit ^{2})



Area = 0.3299 Unit ^{2} Area = 0.50686 Unit ^{2}



Area = 0.443214 Unit ^{2} Area = 0.4137 Unit ^{2}



Area = 0.434 Unit ^{2} Area = 0.46545 Unit ^{2}

Figure 2. English items ICCs with higher probability compare with B.M. items (Area between ICCs > 0.2 unit ^{2})

the students. However, only two (items 13 and 32) out of the three items identified as easier were shown to have a higher probability of answering them correct.

The mean test score for the Malay version of the mathematics test is slightly higher than the English version as shown in Table 4.

The correlation between students' performance on the two language versions of the mathematics test is relatively high, 0.799. A comparison of

Table 4
Mean Score for the English and Malay Version of Mathematics Tests

Mathematics Tests	Mean	SD	S.E.M.
English	30.569	7.616	0.339
Malay	31.544	7.146	0.318

students' performance in the two language versions of the test using paired-sample t-test showed that the difference is significant, $t = -4.652$ (p=.05). This means that most students performed better in the Malay version of the mathematics test.

Conclusion and Discussions

The higher performance in the Malay version of the mathematics test shows the need to continue administering mathematics test in the primary language of instruction. Data from both the difficulty indices and ICCs also supported the need to continue administering nationwide mathematics assessment to Malaysian students in dual-language booklet. This may improve validity and help to ensure equitable assessment of all students. Nonetheless, the use of dual-language accommodation may be giving unfair advantage to those who benefited from this accommodation. Students who are English proficient as well as proficient in Malay language will benefit more than their counterparts who are non-English proficient and non-Malay proficient. Research findings are needed to inform policy makers about appropriate accommodation of students who are either non-English proficient or non-native language proficient or both on assessing mathematics. A change of approach for future accommodation strategies

is needed for this group of students to benefit from the appropriate accommodations.

In an attempt to achieve a valid assessment of students' mathematical abilities, other suitable test accommodations can be recommended such as language simplification or providing extra testing time. Investigating the role of each type of test accommodation in creating a fair measure of students' abilities in mathematics is necessary.

The current practice of implementing standardized mathematics tests to all students, in a language which some of the students have yet to master, may produce invalid results and do not provide information on the mathematics abilities that are being tested. The waste of public monies to pay for testing in bilingual booklets may raise the question of accountability by the MES. It is recommended that academic language proficiency assessments should be conducted to help gauge student readiness for instruction and mathematics testing in English. The practice of making students take a test in a language that they are not familiar will impact negatively on their emotions as well as his/her education well being. Students and their parents may interpret their poor performance as a true reflection of their mathematical abilities.

Another issue that needs attention is the scoring practice that allows students to answer either in English or Malay language. Steps need to be taken to enhance inter-rater reliability in scoring responses of students in different language versions.

Recommendations

On the issue of ensuring fair assessment of mathematics ability of students, it is important to determine that students have the requisite language skills for taking mathematics test in English. Measure of both English language development and content knowledge of students need to be done simultaneously. This may help to predict students' readiness to perform in the content area.

References

Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14, 219-234.

Ainan Abdul Samad (2003). *English as A Tool for Science and Technology knowledge Acquisition: Bilingual Assessment Instruments as A Transitional Measure.* Paper presented at the International Association of Educational Assessment. 5-10 October 2003 Manchester, United Kingdom.

Anderson, J. (22/3/2007). *The Mathematics of language.* Harvard Graduate School of Education.

Bailey, A.L. & Butler, F.A. (2004) Ethical considerations in the assessment of the language and content knowledge of U.S. school-age English learners. *Language Assessment Quarterly*, 1(2&3), 177-193.

Callosum, C. (April 13, 2005). Do You Need to Speak in English to do Science? [Electronic Version]. *Linguistics- General and Professionals*. Retrieved 29/4/2007 from http:// semanticcompositions.typepad.com/index/2005/04/do you need to. html.

Dolan, R. P., Murray, E. A., & Strangnab, N. (April 2006). Mathematics Instruction and assessment for Middle School Students in the Nargins: Students with learning disabilities, Students with mild mental retardation and students who are English Language Learners [Electronic Version].

Goldsmith, S. M. (2003). *Lost In Translation: Issues in Translating Tests for Non-English Speaking, Limited English Proficient and Bilingual Students*.

Kee, G.H. (2007) *What are the Impacts of the English-Medium Instruction Policy of 2002 on the School Experience of Malaysian Students*?

Lager, C.A. (2006). Types off mathematics-reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology, 27(2-3)*, 165-204.

Shaffel, J. & Belton-Kocher, E. (2006). The impact of language characteristics in Mathematics test items on the performance of English Language learners and students with disabilities. *Educational Assessment, 11*(2), 105-126.

Stansfield, C. W. (1996). Content Assessment in Native Language: Eric Digest [Electronic Version]. Retrieved 24/5/2007 from http://www.ericdigest.org/1998-1/content.htm.

State of Iowa, Department of Education (2007). Guidelines for the Inclusion of English Language Learners (ELLs) in K-12 Assessments.

Ong Saw Lan is currently a senior lecturer in the School of Educational Studies of Malaysia Science University. She teaches courses in educational measurement, psychometric theory, and curriculum and program evaluation. She has published and co-published research findings in numerous local and international journals. She is actively involved in several national and university level research projects on assessment issues, testing of bilingual students, and post-graduate students research and statistical competency. A number of her works have been presented at local and international conferences. She currently serves as a member of the editorial board for Journal of Educators and Education and International Journal of Testing.

Appendix A

Comparing item difficulty in English version and Malay version

Item No.	b_R English	b_F Malay	t	df	Favor
6	0.74	1.12	-2.32*	927	E
9	-2.28	-1.59	-2.10*	934	E
13	1.43	0.96	2.99*	938	M
18	0.07	0.55	-2.78*	935	E
20	0.12	0.58	-2.72*	938	E
23	-0.18	0.3	-2.65*	935	E
25	1.13	1.7	-3.67*	936	E
32	0.33	-0.61	4.94*	933	M
33	2.49	2.12	2.32*	929	M
1	-1.86	-1.99	0.39	939	
2	-1.65	-1.57	-0.3	941	
3	-2.23	-2.35	0.31	941	
4	0.31	0.46	-0.86	938	
5	-2.58	-2.61	0.05	941	
7	0.14	0.26	-0.67	938	
8	-0.54	-0.61	-0.34	936	
10	-1.73	-1.85	0.38	940	
11	-.1.12	-1.49	1.02	941	
12	0.16	-0.09	1.39	935	
14	0.21	-0.1	1.77	935	
15	0.52	0.6	-0.49	937	
16	-0.77	-0.69	-0.4	940	
17	-0.3	-0.27	-0.13	940	
19	0.84	0.84	-0.02	931	
21	1.08	1.12	-0.22	936	
22	-0.16	0.16	0	939	
24	0.56	0.73	-1.02	937	
26	0.45	0.19	1.55	932	
27	0.13	0.21	-0.51	936	
28	0.35	0.55	-1.2	935	
29	.1.04	0.98	0.39	926	
30	0.81	0.75	0.36	930	
31	-1.24	-1.45	0.8	934	
34	0.93	0.62	1.86	933	
35	0.9	0.95	-0.32	930	
36	-0.58	-0.81	1.06	936	
37	0.91	0.72	1.17	933	
38	0.59	0.65	-0.41	935	
39	0.55	0.32	1.36	932	
40	0.51	0.7	-1.14	930	

READING COMPETENCIES IN THE PRIMARY AND INTERMEDIATE LEVELS AS PREDICTORS OF APTITUDE FOR HIGH SCHOOL

Jerome A. Ouano
De La Salle University
Manila, Philippines

Reading is a fundamental literacy that affects elementary students' readiness for succeeding academic tasks. Based on the assumption that reading ability makes gainful contributions to students' academic performance, as well as its relevant thrust to the Basic Education Sector Reform Agenda (BESRA) of the Department of Education, a predictive-retrospective quantitative design was employed in determining the predictive utility of reading competencies in each grade level from Grade 3 to Grade 6 on students' reading performance in the same grade level. The effects of these competencies as well as of general reading performance on their scholastic aptitude for high school work was also determined. Standard regression reported that elementary students' competence in vocabulary, comprehension, and study aids each predicted their overall reading performance. Only their reading performance in Grade 4 and Grade 6 predicted their readiness for high school work. In Grade 4, only comprehension and study aids predicted their aptitude. In Grade 6 only vocabulary and comprehension predicted aptitude. Implications were discussed in terms of exploring effective pedagogical framework for reading instruction in the hope of maximizing the benefits elementary students can gain from their reading competencies as they go through the higher levels of basic education.

An ever increasing breadth of literature on reading has provided us with relevant explanations as well as comprehensive understanding of this complex but fundamental literacy. Both educators and psychologists express interest in further exploring this literacy in the hope of providing an even more contextual and practical account of the development of skills in reading

as well as its association with academic performance. This pursuit is revealing because, to begin with, reading is a purposeful activity requiring the orchestration of a wide variety of cognitive skills to decode, comprehend, and learn from text (Cross & Paris, 1988). Our increased understanding about students' reading development and competencies makes us better see how reading skills serve as resources for learning.

As a complex but fundamental literacy, reading is an interface of various but related skills (Vellutino, Tunmer, Jaccard, & Chen, 2007), that can be developed through experience or training. Some of the most potentially powerful of these skills are vocabulary (Roskos, Ergul, Bryan, Burstein, Christie, & Han, 2008; Ricketts, Nation, & Bishop, 2007) and comprehension (Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008; Keith, 1999).

Although more credit could be due for the primary underlying skill known as phonological awareness (Gottardo, Collins, Baciu, & Gebotys, 2008), vocabulary is one of the important subskills of reading in that it is associated with other important reading skills (Ricketts, Nation, & Bishop, 2007). It undergirds other skills that affect general reading competence. For example, word knowledge is related to later reading (Lindsey, Manis, Bailey, 2003). This indicates how vocabulary affects reading. It is even partly responsible for later difficulties in learning (Beitchman, Jiang, Koyama, Johnson, Escobar, Atkinson, & Brownlie, 2008), when its development is curtailed, because students' knowledge of words and their meanings form the building blocks of their literacy development (Neuman, 2006).

Comprehension is also among those subskills that have been substantially explained as a powerful foundation of general reading ability. In previous studies, comprehension provided substantial effects on the development of reading fluency and proficiency (Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008; Keith, 1999). There is reasonable explanation as to why this particular reading subskill has a significant effect on students' reading performance. In a multivariate study on the evidence for a convergence of different skills on which reading comprehension builds, it was explained that various reading subskills and reading-related cognitive skills contribute to reading comprehension performance (Vellutino, Tunmer, Jaccard, & Chen, 2007). In other studies, specific reading subskills, such as vocabulary (Berry, 2008), verbal working memory, and cumulative linguistic knowledge (Jincho, Namiki, Mazuka, 2008), syntactic awareness (Mokhtari & Thompson, 2006), and specific reading strategies (van der Schoot, Vasbinder, Horsley, van Lieshout, 2008), were found to be of vital contribution to reading

comprehension proficiency. Clearly, these pieces of evidence support the general assumption that reading comprehension influences a student's general reading performance.

How these subskills bring about change in the reading competence of students is essential to understanding the relationship of reading subskills and students' future reading competencies, and even their future academic performance (Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov, Pagani, Feinstein, Engel, Brooks-Gunn, Sexton, Duckworth, & Japel, 2007; Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008; Keith, 1999). A good number of literature on reading explains that reading leads to more efficient and adaptive cognitive processing of text (Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008). Specifically, vocabulary (Roskos, Ergul, Bryan, Burstein, Christie, & Han, 2008; Ricketts, Nation, & Bishop, 2007) and comprehension, are associated with more effective creation of richly knitted mental models of the text (Van der Schoot, Vasbinder, Horsley, Van Lieshout, 2008). It is their undeniable that the ability to create appropriate mental models in some efficient way is a resource for successful academic engagement.

Considering the role of different reading subskills, such as vocabulary, word reading, and comprehension (Ricketts, Nation, & Bishop, 2007), the ability to read is a basic requisite for learning in various content area subjects, among these are science and math. For example, reading is considered as an essential tool in science literacy (Creech & Hale, 2006); and reading is generally an indicator of changes in problem solving skills (Grimm, 2007). In other studies, integrating reading with learning in social studies, mathematics (Robb, 2003) and science (Holliday, 2003; Robb, 2003) has reported students' success in learning in these subject areas. These findings suggest that reading could assist the acquisition of content area knowledge and subsequently lead to improved academic performance, particularly in later academic level (Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov, Pagani, Feinstein, Engel, Brooks-Gunn, Sexton, Duckworth, & Japel, 2007).

After accounting for the benefits of reading performance on students' cognitive processes, it can be assumed that students' competence in the reading subskills as well as their general reading proficiency prepare them for the rigorous cognitive processes required in their later general academic work and subject-area-specific tasks. In a study with pre-school children, reading status was found to be one of the indicators of school readiness

(West, 2000). Similarly among high school students who entered university, reading inadequacy was found to inhibit their learning efficiency (Dele, 1985). These two studies suggested the essential impact of reading ability on students' readiness to pursue the next level of their schooling, whether from nursery to elementary or from high school to college.

The works cited earlier strongly point to reading competencies as important sources of students learning in various subject areas. The same works may encourage Philippine education's reform efforts by sustaining the implementation of its new policy reforms known as the Basic Education Sector Reform Agenda (BESRA). This package of policy reforms (Ocampo, 2006) has also put an emphasis on satisfactory achievement levels of students at every grade in the whole of basic education schooling, among other objectives. The Key Reform Thrust (KRT) # 3 of BESRA expresses the need to formulate effective national learning strategies for the English and Filipino languages (Ocampo, 2006). Our increased understanding of the important contribution of reading competence to the academic achievement of students could be of help in the national initiative of identifying available resources for mastery of English language; enhancing and universalizing Filipino students' learning in English language.

Informed by substantial literature on the relationship of reading and academic performance, reading ability that builds on a variety of competencies, as well as the policy reforms of the Department of Education, the study operationalized three conceptual models. The first model was based on the assumption that reading performance, especially in the elementary years, is a function of not only one but many reading subskills (Vellutino, Tunmer, Jaccard, & Chen, 2007), where these subskills influence a reader's ability to understand the literal and inferential meanings of written texts. Another model was drawn from the assumption that students' reading ability offers many potential benefits to academic performance, and therefore this ability can have a beneficial effect on elementary students' readiness for future academic engagement. The third model was a follow up of the second based on the assumption that not only the general reading achievement but also the specific reading competencies could more clearly account for elementary students' scholastic aptitude.

In view of the three models, the study sought to answer three specific questions in the context of Philippine elementary classroom; *first*, whether elementary students' competence in vocabulary, comprehension, and study

aids forecast changes in their reading achievement; *second*, if these elementary students' reading achievement in each grade level contribute to their readiness for high school academic work; and *third*, whether all or only some reading competencies in all grade levels could better predict students' scholastic aptitude.

Method

The study employed a predictive retrospective quantitative design (Johnson, 2001). This was fitting for testing the effects of the predictor variables on the criterion variables included in this study. It sought to find out the earlier predictors of reading performance across four grade levels of elementary studies and of the scholastic aptitude for high school.

Participants

A total of 203 elementary students from selected schools participated in this study. The mean age for Grade 3 was nine and forr Grade 6, 12. Selection of participants was based on the following criteria. *First*, each of the participants must have taken the Level-1Reading Test in his/her primary years (Grades 3 and 4). *Second*, the same students took the Level-2 Reading Test in his/her intermediate years (Grades 5 and 6). *Finally*, the same students took the Elementary Level Competency Measure (ELCOM) in their final year in elementary. In all, participants included 122 females (60%) and 81 males (40%).

Measure

Students' reading ability was measured through their scores in the *Reading Test* (Levels 1 and 2). This test was developed, normed, and used by the Center for Educational Measurement (CEM). Level-1 Reading Test was designed for primary-level students, with 3 subtests, including *vocabulary* (i.e., synonyms, polysemantics, and affixes), *comprehension* (i.e., literal, inferential, and reorganizational), and *study aids* (i.e., alphabetizing and interpreting tables). It was administered to the participants during their Grade 3 and Grade 4 elementary years. The Level-2 Reading Test, designed for the intermediate levels containing *vocabulary* (i.e., synonyms, idioms, and polysemantics), *comprehension* (i.e., literal, inferential, and reorganizational), *study aids* (i.e., use of references and

interpreting tables), and *scanning*, was administered to the same students in Grade 5, then in Grade 6. The measures of students' reading competencies per grade level were their percent-correct scores in the subtest (*i.e., vocabulary, comprehension, study aids, and scanning*). For their general reading performance, the measures were students' percent-correct scores of the Reading Tests per grade level.

For scholastic aptitude, the measures were students' percent-correct general scholastic aptitude (GSA) scores in the *Elementary Level Competency Measure* (ELCOM). The test, which was developed by the CEM to assess elementary graduates' readiness for high school academic work, consists of science, mathematics, verbal English, verbal Filipino, and nonverbal reasoning. Participants responded to this test during their final year in the elementary.

Procedure

Primary data were accessed from the databank of the CEM. The selection of data strictly conformed to the criteria for selection of participants based on the research design. The first to be retrieved were the data of students who had ELCOM scores, then dataset was adjusted based on the students with ELCOM scores who also had Reading Test Level-2 scores. Data selection was further adjusted to students with ELCOM and Reading Test Level-2 scores who also had Reading Test Level-1 scores.

Data analysis

Consistent with the questions posed in this study, data were analyzed using standard multiple regression. This statistical analysis was instrumental in measuring the effects of reading competencies on reading performance and scholastic aptitude, as well as the effects of reading performance on scholastic aptitude.

In testing the first hypothesis, standard regression analyzed the predictive power of each reading competency score on the reading achievement scores. This analysis was done for each grade level. The second hypothesis used the same statistical tool to analyze the effects of reading achievement scores in each grade level on the GSA. The third hypothesis was tested by taking each score of the reading competencies in all grade levels as predictor variables for general scholastic aptitude. Finally, interpretations were drawn from the results in view of the research questions.

Results

Table 1 shows the descriptive statistics covering the participants' profile as well as the average percent-correct scores and other descriptive values on each reading competency, the reading performance in each grade level, and the GSA. Noticeable in the mean percent-correct scores for *vocabulary*, *comprehension*, *study aids*, and *scanning* is a trend of increase from the lower grade level to the final year of the primary as well as of the intermediate levels.

Results revealed significant correlations among general reading scores in the four grade levels. Cases of high correlations were also noted among the four scores of each subtest. The three reading subtests measured in Grade3 were also correlated. The same case was observed among all the subtests in the succeeding grade levels. Part of Table 1 shows these correlations.

In terms of the first research question whether reading competencies predict reading achievement, it may be recalled that it was hypothesized that reading achievement in a particular grade level is undergirded by various subskills or competencies, such as *vocabulary*, *comprehension*, and *study aids* (plus *scanning* in the intermediate levels). Standard multiple regression analysis showed that in each grade level, each competency predicts reading achievement, and each regression model explained 100% of the total variances as shown in Table 2.

Grade 3 students' general reading performance, as assessed through a reading achievement test, was significantly influenced by their vocabulary skills, comprehension, and study aids [$F(3, 202) = 2.643$, $p < .001$, $R^2 = 1.000$]. Effects of the same competencies in Grade 4 on their 4th grade reading achievement were similar [$F(3, 202) = 2.484$, $p < .001$, $R^2 = 1.000$]. In the intermediate years, students' skills in vocabulary, comprehension, study aids, and scanning all significantly explained 5th graders' [$F(4, 202) = 1.401$, $p < .001$, $R^2 = 1.000$] as well as 6th graders' [$F(4, 202) = 1.659$, $p < .001$, $R^2 = 1.000$] reading achievement.

Among all the predictors of reading achievement in Grades 3 and 4, *comprehension* showed greater effects as indicated in their standardized beta (β), followed by *vocabulary*, then *study aids*. This suggests that primary graders' degree or extent of understanding of a written text greatly influenced their achievement scores in reading. Their ability in the literal,

Table 1

Means, Standard Deviations, and Correlations of Age, Reading Competencies, General Reading Scores, and Scholastic Aptitude

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Gender	1.60	0.50	------	-0.08	-0.14	-0.13	-0.13	0.1	0.11	**0.19**	**0.25**	**0.17**	**0.27**	**0.19**	**0.25**	**0.16**	**0.2**	0.13	**0.31**	**0.22**	**0.27**	**0.17**	**0.23**	**0.21**	**0.29**	**0.23**
2. Age in Grade 3	8.85	1.06		------	**0.55**	**0.55**	**0.3**	**-0.15**	-0.13	**-0.15**	**-0.21**	-0.08	-0.07	-0.12	**-0.18**	0	-0.04	**-0.14**	-0.1	0	**-0.16**	-0.09	-0.09	-0.13	**-0.19**	**-0.16**
3. Age in Grade 4	9.91	0.56			------	**0.95**	**0.51**	-0.09	-0.1	-0.09	**-0.17**	-0.08	**-0.18**	-0.08	-0.12	-0.05	-0.07	-0.11	**-0.18**	-0.03	**-0.18**	-0.09	**-0.14**	-0.09	**-0.17**	-0.13
4. Age in Grade 5	10.91	0.57				------	**0.53**	-0.1	-0.1	-0.1	**-0.16**	-0.08	**-0.19**	-0.08	-0.13	-0.05	-0.08	-0.11	**-0.2**	-0.03	**-0.2**	-0.09	**-0.15**	-0.09	**-0.18**	-0.13
5. Age in Grade 6	11.84	1.01					------	-0.03	-0.02	-0.1	-0.13	0	-0.08	0	-0.1	0.01	0	-0.03	-0.12	-0.11	-0.11	-0.01	-0.05	-0.06	-0.13	-0.07
6. Vocabulary 3	48.22	18.39						------	**0.75**	**0.8**	**0.77**	**0.73**	**0.64**	**0.72**	**0.71**	**0.59**	**0.6**	**0.55**	**0.58**	**0.55**	**0.6**	**0.85**	**0.73**	**0.78**	**0.76**	**0.68**
7. Vocabulary 4	57.66	19.84							------	**0.76**	**0.77**	**0.69**	**0.74**	**0.69**	**0.71**	**0.59**	**0.68**	**0.57**	**0.61**	**0.51**	**0.66**	**0.76**	**0.88**	**0.76**	**0.78**	**0.73**
8. Vocabulary 5	48.54	19.15								------	**0.87**	**0.75**	**0.73**	**0.82**	**0.8**	**0.6**	**0.69**	**0.6**	**0.71**	**0.61**	**0.69**	**0.81**	**0.8**	**0.92**	**0.88**	**0.8**
9. Vocabulary 6	56.25	19.80									------	**0.72**	**0.76**	**0.78**	**0.83**	**0.57**	**0.67**	**0.55**	**0.68**	**0.58**	**0.7**	**0.78**	**0.82**	**0.84**	**0.92**	**0.83**
10. Comprehension 3	55.37	23.23										------	**0.76**	**0.74**	**0.74**	**0.71**	**0.67**	**0.59**	**0.66**	**0.54**	**0.66**	**0.95**	**0.79**	**0.78**	**0.78**	**0.76**
11. Comprehension 4	66.58	23.26											------	**0.73**	**0.75**	**0.59**	**0.72**	**0.59**	**0.74**	**0.56**	**0.67**	**0.77**	**0.94**	**0.77**	**0.82**	**0.82**
12. Comprehension 5	46.39	19.70												------	**0.81**	**0.6**	**0.7**	**0.67**	**0.69**	**0.64**	**0.7**	**0.78**	**0.78**	**0.96**	**0.84**	**0.8**
13. Comprehension 6	56.12	20.45													------	**0.58**	**0.67**	**0.6**	**0.71**	**0.66**	**0.72**	**0.77**	**0.79**	**0.85**	**0.95**	**0.83**
14. Study Aids 3	51.77	21.40														------	**0.65**	**0.52**	**0.56**	**0.47**	**0.53**	**0.84**	**0.67**	**0.64**	**0.63**	**0.61**
15. Study Aids 4	57.56	24.11															------	**0.58**	**0.67**	**0.47**	**0.59**	**0.72**	**0.87**	**0.73**	**0.73**	**0.73**
16. Study Aids 5	46.08	21.37																------	**0.62**	**0.45**	**0.53**	**0.63**	**0.64**	**0.75**	**0.64**	**0.62**
17. Study Aids 6	62.49	22.93																	------	**0.56**	**0.65**	**0.68**	**0.75**	**0.75**	**0.82**	**0.72**
18. Scanning 5	57.54	19.83																		------	**0.66**	**0.59**	**0.58**	**0.75**	**0.68**	**0.59**
19. Scanning 6	68.55	18.89																			------	**0.68**	**0.71**	**0.75**	**0.82**	**0.7**
20. Reading 3	52.37	19.07																				------	**0.83**	**0.83**	**0.82**	**0.78**
21. Reading 4	62.91	20.06																					------	**0.83**	**0.86**	**0.85**
22. Reading 5	58.48	18.55																						------	**0.9**	**0.84**
23. Reading 6	48.93	17.42																							------	**0.88**
24. Aptitude	59.80	16.44																								------

Note: Figures in boldface are significant coefficients ($p<.05$)

Table 2

Regression Analysis of Primary- and Intermediate-Level Reading Subskills as Predictors of Reading Achievement

Reading Subskills	Grade 3				Grade 4				Grade 5				Grade 6			
	B	SE	β	t	B	SE	β	t	B	SE	β	t	B	SE	β	t
Vocabulary	.314	.002	.303	183.247***	.312	.002	.309	170.048***	.304	.002	.334	138.606***	.300	.002	.321	138.420***
Comprehension	.437	.002	.532	277.193***	.437	.002	.507	264.183***	.400	.002	.453	169.840***	.401	.002	.442	180.570***
Study Aids	.249	.001	.279	172.363***	.252	.002	.288	162.691***	.133	.001	.163	89.820***	.135	.001	.175	94.507***
Scanning	----	----	----	----	----	----	----	----	.163	.002	.185	104.527***	.165	.002	.168	89.767***
	$F(3, 202) = 2.643, p < .001,$ $R^2 = 1.000$***				$F(3, 202) = 2.484, p < .001,$ $R^2 = 1.000$***				$F(4, 202) = 1.401, p < .001,$ $R^2 = 1.000$***				$F(4, 202) = 1.659, p < .001,$ $R^2 = 1.000$***			

***$p<.001$

inferential, and reorganizational content areas of comprehension affects their reading achievement more strongly than their word knowledge and competence in alphabetizing and interpreting graphs, tables or maps. Analogous results were also observed in their intermediate levels (Grades 5 and 6), in which *comprehension* contributed the most effect on reading achievement, followed by *vocabulary*, although the effects of *study aids* and *scanning* varied across these two year levels.

Another interesting result of the study was the evidence of trend in the effect sizes of the different reading competencies within and across the curriculum levels, i.e., *primary* and *intermediate*. In their primary level, Grade 4 students' skills in arranging words alphabetically and using and interpreting maps, tables, and graphs to gather information improved their reading achievement scores. Slight increase in effect sizes was also observed in vocabulary. The effect of comprehension on their reading achievement score decreased in Grade 4. On their intermediate level, only study skills had an increasing effect. The effects of vocabulary, comprehension, and scanning on students' reading achievement scores decreased.

The increasing effects of study skills and vocabulary on reading students' achievement scores in the primary grades could be due to mastery of the strategies to increase organization of information and of explicit rules that govern the language to facilitate word knowledge. The decreasing effects of comprehension in the primary grades could be a function of the inclusion of more lengthy items, that actually assess the same competency standards. In the intermediate level, only the effect sizes of study skills increased, implying students' mastery of strategies for using explicit rules of alphabetizing, referencing, and labelling. The effects of comprehension and scanning decreased generally because of the inclusion of more lengthy and more complex items measuring the same competency standards. Decrease in the effect sizes of vocabulary could be due to the inclusion of more uncommon words.

In terms of the difference in effect sizes of reading competencies between primary and intermediate levels, that is between Grade 4 and Grade 5, vocabulary showed an increasing effect on Grade 5 reading achievement score. The inclusion of idioms in the test may have provided students with more contexts for the application of polysemantic knowledge. All other reading competencies had decreasing effects, which may be explained in terms of the inclusion of more content areas in the subtests on comprehension and study aids.

The second research question sought to find out the effects of students' reading ability on their aptitude for high school academic work. Results showed that students' reading ability in the final year of each curriculum level predicted their scholastic aptitude for high school studies. Table 3 indicates that regression analysis accounted for 81% of the total variance, which reflects a highly acceptable explanation by the model [$F(4, 202) = 205.368$, $p<.001$, $R^2 = .806$]. In the primary levels, only the reading achievement scores in Grade 4 predicted students' aptitude for high school studies measured in their Grade 6 year. In the intermediate levels, only their Grade 6 reading achievement scores predicted their scholastic aptitude.

Considering the incremental nature of the development of reading ability, it may be expected that some levels of mastery have been achieved only during the final years of primary and intermediate levels. This expected

Table 3

Regression Analysis of General Reading Achievement across Elementary Levels as Predictors of Scholastic Aptitude

	Scholastic Aptitude for High School			
	B	*SE*	*β*	*T*
Primary Level				
Grade 3 General Reading Achievement	0.018	0.055	0.020	0.319
Grade 4 General Reading Achievement	0.275	0.056	0.336	4.871***
Intermediate Level				
Grade 5 General Reading Achievement	0.132	0.075	0.139	1.751
Grade 6 General Reading Achievement	0.392	0.074	0.443	5.294***

$F(4, 202) = 205.368$, $p<.001$, $R^2 = .806$

***$p<.001$

mastery could show in their performance because the curriculum standards across the grades in the primary level are similar. The same applies to the intermediate levels, whose curriculum standards are similar between Grade 5 and Grade 6, but more complex as compared to the primary-level curriculum. The greater effect size of Grade-6 reading achievement score on scholastic aptitude suggests the incremental benefit of reading ability on students' academic readiness.

Upon the assumption that not only does the general reading ability of students predict their scholastic aptitude for high school studies, but also the various specific competencies of reading, regression analysis posted a

significantly convincing 82% explanation of the variance that, indeed, certain reading competencies could explain elementary students' readiness for high school [$F(4, 202) = 60.461$, $p<.001$, $R^2 = .818$]. Table 4 shows the specific reading competencies measured at different levels from Grade 3 to Grade 6 that predicted students' scholastic aptitude.

Table 4
Regression Analysis of Reading Competencies across Elementary Levels as Predictors of Scholastic Aptitude

	Scholastic Aptitude for High School			
	B	*SE*	*β*	*T*
Grade 3				
Vocabulary	-0.06	0.053	-0.067	-1.132
Comprehension	0.052	0.045	0.074	1.168
Study Aids	-0.005	0.037	-0.006	-0.132
Grade 4				
Vocabulary	0.005	0.049	0.006	0.1
Comprehension	0.157	0.045	0.222	3.499
Study Aids	0.08	0.038	0.112	2.103
Grade 5				
Vocabulary	0.048	0.067	0.056	0.714
Comprehension	0.076	0.056	0.091	1.364
Study Aids	0.04	-0.035	0.053	1.163
Scanning	-0.011	0.038	-0.013	-0.291
Grade 6				
Vocabulary	0.182	0.063	0.219	2.883
Comprehension	0.19	0.055	0.236	3.467
Study Aids	-0.004	0.037	0.006	0.114
Scanning	0.025	0.045	0.029	0.55

$F(4, 202) = 60.461$, $p<.001$, $R^2 = .818$

$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

Among the reading competencies measured in each of the grade levels across the primary and intermediate curriculum levels, only Grade 4 *comprehension*, Grade 4 *study aids*, Grade 6 *vocabulary*, and Grade 6 *comprehension* predicted scholastic aptitude. Noticeable in the results was the predictive utility of some reading competencies in the final year of the primary level as well as of the intermediate level, which is consistent with the predictive power of Grade 4 and Grade 6 reading ability discussed earlier. Results further indicated that Grade 4 *comprehension* and *study aids*, but not *vocabulary*, made stronger contributions to the predictive

power of Grade 4 reading performance on students' scholastic aptitude. Grade 6 *vocabulary* and *comprehension* did a similar impact on the effects of Grade 6 reading performance on their scholastic aptitude.

Discussion

From the results of this study, it was notable that a number of competencies undergird students' ability to read. Indeed, reading is not a function of only one skill but of a combination of skills that interact with each other. Elementary students' ability to read written texts is explained by their ability to understand word meanings in context and apply specific rules that govern the language. It is also explained in terms of their ability to identify significant details, including specific facts and the main idea of a written text, and infer both explicit and implicit ideas. The extent of their knowledge and use of reference materials, graphs, and table to gather and interpret information also influenced their reading ability.

The greater effects of students' comprehension skills on their general reading performance across all grade levels were also notable. These results are consistent with literature on reading development and instruction that explains that comprehension is a powerful set of subskills of reading (Vellutino, Tunmer, Jaccard, & Chen, 2007) because it involves many different cognitive processes and resources that are linked to reading performance. Among these are syntactic awareness (Mokhtari & Thompson, 2006), cumulative linguistic knowledge (Jincho, Namiki, Mazuka, 2008), vocabulary (Berry, 2008), and specific reading strategies (van der Schoot, Vasbinder, Horsley, van Lieshout, 2008). Clearly, it shows that comprehension benefits from many other subskills, where vocabulary skills, study aids, and scanning could be no exception as the results imply a sharing of variances among these competencies.

Taken together, students' skills in vocabulary, comprehension, study aids, and scanning are powerful skill areas that accounted for the elementary students' performance in reading from Grade 3 to Grade 6. Consequently, their general reading ability in Grade 4 and Grade 6 predicted their aptitude for high school academic work. The final year of each level (primary and intermediate) marked a significantly meaningful effect of general reading achievement on scholastic aptitude, indicating that significant amount of explicit changes in reading ability happen at these times when curriculum-

aided mastery may be achieved. Also shown in Table 1, students' scores on each reading competency increased from Grade 3 to Grade 4 and also for the intermediate levels from Grade 5 to Grade 6. This change suggests the importance of reading instruction during these years, i.e., Grade 4 and Grade 6, to focus on those competencies that provide a scaffold for students' reading ability. Reading instruction at this time may be more gainful if it strengthens students' vocabulary, comprehension, study aids, and scanning. Reading instruction in Grade 3 and Grade 5 is no less important because it serves as an opportunity for the enhancement of those skills or competencies that could account for students' reading ability. In view of BESRA's reform thrust on learning strategies for English language (Ocampo, 2006), it might be important to figure out how reading instruction must be planned strategically from Grade 3 to Grade 4, then from Grade 5 to Grade 6 so that the incremental nature of reading ability (Snow, Porche, Tabors, & Harris, 2007) will be effectively maximized.

Students' general reading performance in Grade-4 and Grade-6 predicted their scholastic aptitude; the same is true for the competencies assessed in these grade levels. The predictive effects of comprehension in Grade 4 and Grade 6 on their readiness for high school academic work were more likely due to the high utility of comprehension skills in many subject areas included in the aptitude test. The measure of scholastic aptitude included subject-matter tests in science, mathematics, English, and Filipino, all of which called for comprehension skills. On this note, it is important to acknowledge the fact that various reading subskills as well as reading-related cognitive skills contribute to reading comprehension (Vellutino, Tunmer, Jaccard, & Chen, 2007). For this reason, students' fundamental capabilities in comprehension lend more benefit to their readiness for high school studies by virtue of the multifacetedness of this competency.

It must also be noted that the English and Filipino components of the aptitude test also included vocabulary, which could explain that students' vocabulary skills in Grade 6 had a greater effect on their scholastic aptitude. The significant prediction made by Grade 4 study aids on their scholastic aptitude may be due to the fact that compared to study aids measured in Grade 6, the former included alphabetizing, which is a very basic organizational skill, in addition to the use of tables, maps, and graphs to gather and interpret information, while the latter did not include alphabetizing, but instead added the use of reference materials, which may not be as functional in their measured aptitude as the other study aid strategies.

Conclusions

True to its definition, reading requires orchestration of various skills (Cross & Paris, 1988). Elementary students' skills in *vocabulary*, *comprehension*, *study aids*, and *scanning* all accounted for their ability to read during each grade level. Enhancement of these skills is a key to their reading achievement. Students' reading achievement in the final years of primary and intermediate schooling predicted their readiness for high school academic work. During these grade levels (Grades 4 & 6), students' improvement in reading competencies became more explicit. In the primary levels, Grade 4 *comprehension* and *study aids* were the only reading competencies that explained students' scholastic aptitude for high school. In the intermediate levels, only Grade 6 *vocabulary* and *comprehension* predicted scholastic aptitude. The utility of these reading competencies was more conspicuous because the measure of scholastic aptitude called for these skills.

In the light of the findings, reading instruction must focus on these competencies not only during Grade 4 and Grade 6 but also in the other grade levels.

Based on the results and with reasonable concurrence to existing literature, this study has put forward the predictive utility of elementary students' reading competence to their readiness to embark on the cognitive rigors of later academic work. Noting that reading ability of elementary students as well as their skills in those competencies embedded in their general reading ability lend some advantage to their readiness to pursue in the next levels of basic education, i.e., high school, greater challenge is posed on the adequacy and effectiveness of curricular and instructional focus on the development of reading capabilities among young Filipino learners. The challenge rests on the way reading instruction must be planned strategically across the elementary year levels in order to increase students' resources for better achievement, and all teaching efforts consequently respond to the relevant thrust of the Basic Education Sector Reform Agenda (BESRA) of the Department of Education (DepEd). There is now a pressing need to find out what pedagogical framework for teaching reading will work better with young Filipino learners, or perhaps explore into what specific strategies might be more productive for Filipino learners. More research must be undertaken to find out what approaches to reading instruction will have more benefits on young learners academic achievement.

If we are able to identify these approaches, then we can move on to institutionalizing instructional reforms in our basic education by building more capabilities in the teachers to effectively deliver relevant reading instruction. All reform efforts on education must consider teacher empowerment in terms of access to appropriate professional development and all its needed resources (Ouano, Dela Torre, & Chavez, 2007), because their readiness to perform the needed change is key to a successful move for reform. With adequate teacher-training opportunities, we hope to see improved ways of teaching reading in Philippine schools as measured in terms of higher future academic benefits.

References

Beitchman, J.H., Jiang, H., Koyama, E., Johnson, C.J., Escobar, M., Atkinson, L., Brownlie, E.B., & Vida, R. (2008). Models and determinants of vocabulary growth from kindergarten to adulthood. *Journal of Child Psychology and Psychiatry, 49*, 626-634.

Berry, M.F. (2008). The influence of specific aspects of preschool children's language on growth in reading comprehension across elementary school. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 68(7-A)*.

Creech, J., & Hale. G. (2006). Literacy in science: A natural fit. *The Science Teacher, 73*, 22-27.

Cross, D.R. & Paris, S.G. (1988). Developmental and Instructional Analyses of Children's Metacognition and Reading Comprehension. *Journal of Educational Psychology, 80,* 131-142.

Dele, O. (1985). An investigation of reading readiness for higher education in the medium of English as a second language. *Journal of Teacher Education, 1*, 203-213.

Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology, 43*, 1428.

Gottardo, A., Collins, P., Baciu, I., & Gebotys, R. (2008). Predictors of grade 2 word reading and vocabulary learning from grade 1 variables in Spanish. *Learning Disabilities Research & Practice, 23*, 11-24.

Grimm, K.J. (2007). A longitudinal dynamic analysis of the impacts of reading on mathematical ability in children and adolescents. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67(9-B).*

Holliday, W. G. (2003). Teaming up for science & reading success. *Science and Children, 40,* 38-40.

Jincho, N., Namiki, H., & Mazuka, R. (2008). Effects of verbal working memory and cumulative linguistic knowledge on reading comprehension. *Japanese Psychological Research, 50,* 12-23.

Johnson, B. (2001). Toward a new classification of non-experimental quantitative research. *Educational Researcher, 30,* 3-13.

Keith, T.Z. (1999). Effects of General and Specific Abilities on Student Achievement: Similarities and Differences across Ethnic Groups. *School Psychology Quarterly, 14,* 239-262.

Lindsey, K.A., Manis, F.R. & Bailey, C.E. (2003). Prediction of First-Grade Reading in Spanish-Speaking English-Language Learners. *Journal of Educational Psychology, 95,* 482–494.

Mokhtari, K, & Thompson, H. B. (2006). How problems of reading fluency and comprehension are related to difficulties in syntactic awareness skills among firth grades. *Reading Research and Instruction, 46,* 73-94.

Neuman, S. B. (2006). Building vocabulary to build literacy. *Scholastic Early Childhood Today, 21,* 9.

Ocampo, Dina (2006). The Full Report for the Policy Recommendations for the Basic Education Sector Reform Agenda for the National Language and Literacy Learning Strategies for the Filipino and English Languages. Retrieved February 26, 2010 from http:// dina.ocampo.googlepages.com

Ouano, J. A., dela Torre, L. G., & Chavez, J. C. (2007). Experiences and prospects in teachers' professional development: A continuing journey of teacher empowerment. In B. Atweh, I. Macpherson, A. B. I. Bernardo, M. Ferido, & M. Balagtas (Eds.), *Ripples of Change: A Journey of Preservice Teacher Education Reform in the Philippines.* Philippines: Commission on Higher Education.

Ricketts, J., Nation, K., & Bishop, D.V.M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11*, 235-257.

Robb, L. (2003). Teaching reading in social studies, science, and math. New York: Scholastic Professional Books.

Roskos, K., Ergul, C., Bryan, T., Burstein, K., Christie, J., & Han, M. (2008). Who's learning what words and how fast? Preschoolers' vocabulary growth in an early literacy program. *Journal of Research in Childhood Education, 22*, 275-290.

Savolainen, H., Ahonen, T.,Aro, M., Tolvanen, A., & Holopainen, L. (2008). Readingcomprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction, 18,* 201-210.

Snow, C. E., Porche, M. V., Tabors, P. O., & Harris, S. R. (2007). Is literacy enough? Pathways to academic success for adolescents. Baltimore, MD, Paul H. Brookes Publishing Co., Inc.

Van der Schoot, M., Vasbinder, A., Horsley, T.M., & Van Lieshout, E.C.D.M. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading, 31*, 203-223.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., and Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11*, 3-32.

West, J. (2000). New indicators of school readiness from ELCS-K. *The Child Indicator, 2* (4).

Jerome A. Ouano is a faculty member of the Counseling and Educational Psychology Department of De La Salle University where he teaches courses in educational psychology and assessment. He has written textbooks for Assessment and Field Study courses, and co-authored in scientific papers published for local and international readership. He is actively engaged in quantitative and qualitative research projects, and a number of his works have been presented in local and international conferences. For many years he has been sharing his expertise in pedagogy and assessment through seminar-workshops. Jerome is finishing his Ph.D. in Educational Psychology, major in Learning and Human Development at De La Salle University.

UTILIZING TEST RESULTS TO DESIGN INSTRUCTIONAL MODULES: THE CEM PROFESSIONAL EDUCATION PROGRAM SERIES (PEPS)[1]

Esperanza C. Buen
Center for Educational Measurement, Inc
Philippines

This paper discusses the potential of utilizing test data to design instructional modules for the in-service training of teachers in the core subjects of the school curriculum. It documents a major effort of the Center for Educational Measurement (CEM) in developing and instituting a series of professional education programs, the PEPS, for teachers of English, Mathematics, Science, and Filipino, across ten grade levels of basic education in the Philippines.

The design of each module was based on patterns of academic strengths and deficiencies culled from an analysis of the results of diagnostic tests administered within five years to students of schools that participated in CEM's nationwide testing program. Results showed that areas with learning deficiencies in the lower grades tend to appear in the higher grades. The data served in prioritizing the contents and skills for which the first twenty modules were drafted. Other factors affecting performance were considered in the choice of strategies employed. Highlights of module development, as well as the initial implementation and evaluation of the training series are presented.

Why the PEPS?

One important lesson from the reengineering efforts to improve the system of education in the Philippines is the recurrent finding that teachers are the fulcrum that determines whether any school initiative tips towards success or failure. The quality of instruction in a school is the single most

[1]Paper read at the International Association for Educational Assessment (IAEA) Annual Conference at Cabridge, UK, September 2008.

important factor affecting the quality of learning that takes place, the achievement of standards, the delivery of the curriculum, and the assessment of student progress.

The Center for Educational Measurement (CEM), a local private nonprofit assessment agency, has for many years led the private sector in developing standardized diagnostic instruments for schools nationwide. It has done this with the firm belief that educational assessment should focus on the effects of instruction and that data derived from reliable instruments could support and improve teaching and learning. The CEM has also conducted seminar-workshops with school test users to interpret their students' results and discuss action plans. In the process it has become increasingly aware of teaching-learning problems, the inadequacy of teachers' pre-service training and the need for continuing development programs in schools.

The CEM decided to look more closely into student performance in the core academic areas as a basis for the development of school improvement programs, particularly for teachers' in-service training. It is convinced that such programs should be guided by identified instructional needs and should provide opportunities for teachers to improve their strategies. In the year 2003 CEM organized an enhanced program for instructional seminars, the Professional Educational Program Series (PEPS), which is intended to respond to the need of teachers to increase their knowledge of the subjects and improve their teaching skills. The PEPS was meant to be more directed in focus, standards related, innovative and stimulating, and possessing the essentials of effective professional development (Hawley and Valli, 1999).

I. Analysis of Student Performance in Core Subjects

The first step undertaken in the development of the PEPS was to study the past performance of students in the four major academic subjects emphasized in the local curriculum, in order to provide a basis for the PEPS design.

Instruments. The CEM Diagnostic Tests (DT) are standardized tests in four subject areas, namely, English, Mathematics, Science, and Filipino, for each grade level in basic education (Elementary Grades 1-6, Secondary Years 1-4). In multiple-choice format, they measure performance in

competencies taught in the local curriculum, have a unique form per level, test length, and time allotment. Each content area consists of item clusters, the number of which vary according to level. Developed under the classical test model, the instruments were normed on representative samples of students for each grade level subject (TDD-CEM, 2002).

Sample. The DT's are used as external objective measures in many private schools in the country. Test data were derived from student examinees across five years, from 1999 to 2003. These are students from schools that subscribe to the tests voluntarily, thus the total sample sizes per level test vary, ranging from 7,606 to 19,167 in English, 5,638 to 21,383 in Mathematics, 6,302 to 17,505 in Science, and 695 to 3,928 in Filipino.

Scoring and Analysis. Each examinee's competency scores on the tests are summarized into percent correct scores on the specific content and skill areas and an overall percent correct score with its standard score equivalent. The mean score of the standardization sample or norm group served as the cut score for identifying weaknesses or areas of concern. Content scores that fell below this point would indicate areas where the examinees performed at a level below that of the average student in the norm group. These are to be identified as areas of concern, which would receive attention in preparing teaching modules.

Results. From the detailed analysis of the five-year data on the four subjects, the identified contents with means below the norm average were tabulated per subject and grade level. The mean scores for these content areas as well as the cut score for the norm group per level are given in each table.

1. English. An illustration of the outcomes in the English subject is given in Appendix A1. The initially measured contents run across the grade levels, but only those with scores lower than the norm are reported here.

It is noted that some contents in the Elementary English curriculum were poorly learned across the grade levels. The most commonly unmastered topic was that of *Verb*, from Grades 1 to 5, followed by *Adjective* and *Preposition*. As an apparent consequence, *Sentence*, also figured as difficult in the four higher levels, Grades 3-6, and *Paragraph Development* at Grade 6. Likewise, the *Vocabulary* difficulties in the higher elementary grades might have contributed to the weakness in *Understanding a Selection*.

The secondary level students showed weakness in *Tenses and Agreement* between *Subject and Predicate*, see Appendix A2. This reflects on the upper levels' not doing well enough in *Grammar and Usage* as well as in *Sentence* and *Paragraph Development*. It is also interesting to note the lower years' difficulties in *Vocabulary*, *Recognition of Main Idea and Details*, as well as *Inferring and Reorganization* and the seniors' similar inadequacy in *Reading Comprehension*.

Table 1
Areas of Concern across Elementary and Secondary Levels

Elementary	Secondary
Noun	Simple tenses-verb
Adjective	Modal auxilliaries
Verb	Modification
Preposition	
Sentence	Sentence
Paragraph development	Subject predicate agreement
Structural analysis	Grammar and usage
	Reported speech
	Paragraph development
	Brevity and parallelism
Vocabulary	Recognition of main idea and details
Understanding a selection	Inferring
Sequencing events	Reorganization
	Reading Comprehension

The summary comparison in Table 1 shows how areas of concern in the elementary level tends to resurface in the weaknesses of high school students.

2. Filipino. A similar process was done for the Filipino subject, where weaknesses were also observed across the grades (Appendix B1). Pupils were deficient in learning the parts of speech in Filipino from *Pangngalan* to *Pangatnig*, the ability to handle *Pangungusap (Sentences)* and *Pagsusuring Pangkayarian (Structural Analysis)*, likewise in *Talasalitaan (Vocabulary)* and *Pag-unawa sa Binasa (Reading Comprehension)*. The secondary school students found difficulty in adverbs, phonology, sentence and paragraph development, and understanding of reading material in Filipino.

Compared to English, however, the carryover to the secondary level was observed in the first year students who showed difficulty in a number of competencies relevant to development of *Pangungusap (Sentence)* and *Talata (Paragraph)*, as well as *Talasalitaan (Vocabulary)* and *Pang-unawa sa Binasa (Reading Comprehension)*, and in two areas of second year, having to do with *Pangunawa sa Binasa (Reading Comprehension)* and *Kasanayan sa Pag-aaral (Study Skills)*. Refer to Appendix B2.

3. Science. Some areas of concern in elementary science were on similar topics developed across the grades (Appendix C1). Lower graders found difficulty in learning about *Plants*, *Human Beings*, and *Matter*, as well as *Light*, *Magnets*, and *Electricity*. The upper grades were weak in *Force and Energy* and *Weather and Climate*. But the most consistent and greatest difficulty in all levels was concerning *Earth and Space*.

The high school science subjects are specialized for each year level. However, the inadequacies in elementary science continued through the first year of secondary, with *Nature of Matter* and *Earth and Universe* not being mastered. Second year Biology, had a good number of inadequately learned topics, with the lowest scores in *Diversity and Adaptive Forms*, as well as *Heredity, Variation, and Population*. Third year Chemistry also showed many areas of concern, with the lowest scores in *Chemical Kinetics*, *Electrochemistry*, *Phases of Matter*, and *Symbols and Formulas*. In Physics the students needed more help in the topics on *Energy and Communication*, *Electromagnetic Energy*, *Force and Energy*, and *Waves and Energy*. Refer to Appendix C2.

4. Mathematics. The Mathematics curriculum appeared to have been designed with a clearer map of learning competencies developed with continuity but increasing degrees of complexity along the basic education ladder.

A good number of competencies in Mathematics were far from being mastered by the elementary pupils (Appendix D1). *Numeration* and *Fundamental Operations* posed a problem in the lower grades, while *Geometry* and *Graphs and Scales* figured in the higher grades. But *Decimals* and *Fractions* remained consistently difficult across the levels.

The high school Mathematics competencies showed increasing complexity of some topics which started in the elementary level (Appendix D2).

Number Concepts reviewed *Factors and Primes* and included *Exponents and Radicals* as well as *Complex Numbers*. These the students found difficult. *Geometric Concepts*, problems on *Quadrilaterals*, *Triangles* were not learned well either, and so with *Language of Algebra*, *Algebraic Expressions*, *Quadratic and Circular Functions*. The integrated curriculum included *Statistics*, where the students needed improvement. Understandably, they stumbled in learning the concepts of *Quartiles, Deciles and Percentiles*, *Sequence and Variations*, and *Linear Correlation* as well.

II. Development of PEPS Modules

Work Flow. Based on the findings of the analysis of core subjects, the CEM proceeded with the development of the modules for faculty in-service training, guided by the general flow of activities in Fig ure 1. The instructional design itself followed the basic process suggested by Vinzon (2002) but included more details.

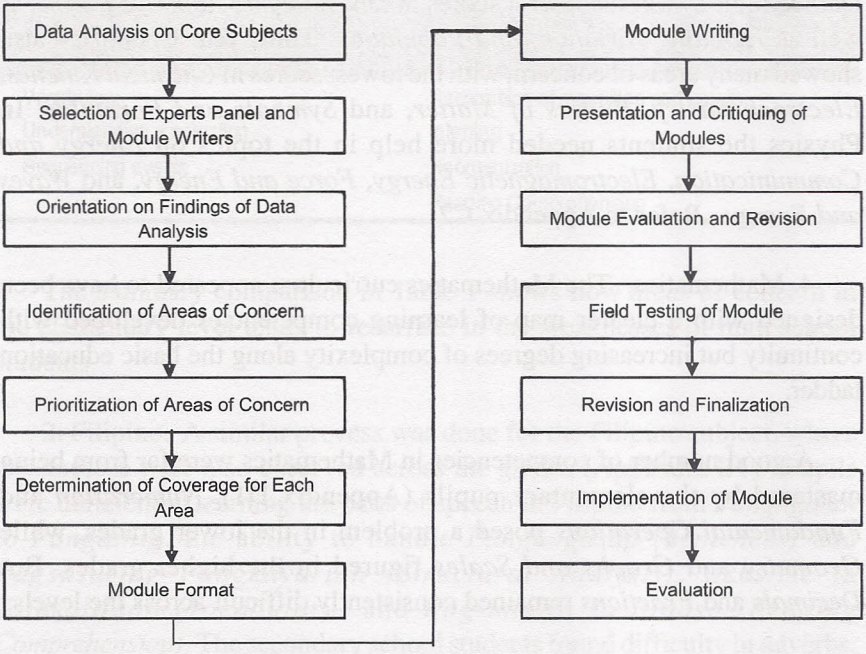


Figure 1. PEPS Module Development Flow Chart

Experts Panel and Module Writers. Educators noted for their experience in actual teaching, curriculum development, writing instructional materials, and conducting staff development programs for teachers, acted as consultants in the development of the PEPS. The experts panel, which included an education and curriculum supervisor, a faculty development director, and a measurement expert, provided guidelines for module development, monitored the production, and evaluated the outputs. The twenty (20) module writers, among whom were university professors, subject coordinators in basic education departments, and textbook writers, wrote the modules on specific topics. They were experienced teachers of the core subjects, conversant with current research, assessment, and the use of technology for instructional purposes. The experts panel and module writers worked in coordination with CEM, which provided them with the student performance data analysis.

Priority Topics. The discussion on analysis results posed a serious challenge to the PEPS consultants. The number of contents needing attention appeared formidable, so efforts had to be exerted in order to limit the topics to the most important as well as the most doable, within the resources available in most of the schools. They dissected the areas of concern for every core subject and their specific components in the syllabi, the score values in relation to other topics, the frequency of occurrence across levels, albeit in varying complexity, and their possible impact on other areas within the same subject and across disciplines. There was also the attempt to put together some areas in one module as long as they are closely related, not only in learning the subject but in application to real life situations.

An example of this discussion centered on *Reading Comprehension.* This is a larger strand which covers the areas of concern on *Vocabulary Skills*, *Scanning*, *Understanding Selections*, *Inferring and Reorganization.* A module could be built around this generic area, with specific time frame devoted to the components and variations done for different levels. The importance of *Reading Comprehension* is its impact on the other subjects in the curriculum. It is a skill that helps students hurdle materials they need to understand as they master the competencies in science, mathematics and other subjects.

The PEPS consultants worked on the priority topics for the first year of the program. They tried to hit an equal representation per core subject, with each of the twenty module writers doing a topic of his choice. They finally arrived at the following 20 modules with the intended level of

Table 2
List of Modules per Subject and Intended Participants

Subject	Title	Teaching level
English	Beginning Reading	Pre-school, Lower Elementary
	Teaching Grammar and Vocabulary	Elementary, Secondary
	Integrated Language Arts	Elementary, Secondary
	Reading Into Writing	Secondary
	Reading Notes and Strategies	Elementary
	LIteracy and Literature	Secondary
Filipino	Pagbasa sa Elementarya	Elementary
	Wika sa Elementarya	Elementary
	Wika sa Sekundarya	Secondary
	Panitikan sa Pilipino	Secondary
Science	The Physical Earth, its Weather and Climate	Elementary
	The Earth, A Special Place in the Cosmos	Secondary-Year 1
	Genetics	Secondary- Year 2
	How Molecules Behave; Laws Governing Gases	Secondary- Year 3
	Waves and Wave Properties	Secondary- Year 4
Mathematics	Number Theory and Rational Numbers	Elementary
	Elementary Algebra	Secondary- Year 1, Year 2
	Geometry and Measurement	Secondary- Year 3
	Congruence of Triangles and Properties of Quadrilaterals	Secondary- Year 3
	Special Functions Geometric Sequences	Secondary- Year 3 & 4

participants (see Table 2), ensuring that for each subject there were modules for the elementary and secondary levels.

The Module Format. During the workshop the experts panel and module writers arrived at the format which the PEPS modules would take and this would be applicable to all the core subjects they were working on. A module was set to run for five-days to a maximum of 25 participants. The allocation of specific content coverage and activities within the time frame should allow a more intensive treatment of the topic. The small group size was intended to maximize involvement of the teacher-learners. The basic elements of the module would be as listed in Figure 2.

Elements of PEPS Modules
1. Introduction
Conceptual framework
Goals/Objective
Uses
Intended Audience
2. Module Text
Expected learning outcomes
Content
Instructional Techniques
Materials and equipment
Evaluation procedures
3. Appendices
Suggested readings
Sample cases or lessons
Glossary of terms
4. Facilitator's/Teacher's Guide
5. Handouts/Worksheets for Participants

Figure 2. Elements of PEPS Modules

Module Writing. It is evident from the required module elements that the writers wanted the PEPS to be an instrument in effective teaching. This would help the teachers deepen their knowledge of the subject, the student, and teaching practice. Effective educators should have clear expectations of learners' ability to achieve and use strategies to help them do this at the highest retention level.

Local teaching practice at the time of the study has been reported to be largely teacher-dominated, with high dependence on text guides, emphasizing recall rather than higher order thinking skills, and passive pupil behavior. It had little use of problem-solving techniques and group methods for cooperative learning and managing students' varied learning styles. (Taguiwalo, 1993).

Schools' inadequacy in facilities and management support have been cited as deterring factors to high student performance, together with often limited and fragmented instructional programs. On the teacher level, it was not rare to see a mismatch between teacher and assignment, lack of preparation in pre-service training, and more often, inefficient use of time, strategies and materials. As a consequence, the accumulated deficiencies of students rendered them unprepared for higher levels of study.

Aware of the situation of schools, teachers and students, the PEPS module writers endeavored to produce work that embodied desired approaches, theory and practice. The strategies employed in the modules were carefully selected for more effective learning of the competencies and adapted to the teachers as adult learners. The resulting modules would be characterized as:

Research-based. Aside from being needs-based, derived from results of student performance, the modules also encourage teachers to try strategies that have been found to work, contents that are in line with current theories and research in the field. They also employ technology in sourcing and delivery.

Constructivist/developmental in approach. The participants would be actively engaged in the learning process and integrate their new experiences with what they already know or are currently practicing in their work. There is emphasis on multiple sources of information, peer collaboration and student-generated questions (Edelstein, 1992).

Learner-centered. The needs of participants are assessed through a pre-seminar survey and the activities ensure individual participation and feedback, constant reflection on learnings. The activities are carefully chosen to stimulate interest and enjoyment, and to meet different learning styles. The limited number of participants allows for greater focus on individual needs.

Focus on metacognitive skills. The design exposes the participants to activities that develop higher order thinking skills. Even in the review of basic algorithms and language rules, concept mapping, identifying patterns, analysis, and creativity are emphasized.

Assessment-oriented. Pre-seminar and post-seminar assessment, written, oral, and performance types are employed to help participants gauge their knowledge of the topics. Student evaluation practices for the core subjects are also given attention.

Transfer of learning. Discussions on application to work, lesson planning, presentation, and critiquing help the participants improve their skills and be ready to use them in the classroom situation.

Presentation and Critiquing of Draft Modules. All the writers' products were presented to the experts' panel and were given critical evaluation as to whether they complied with the required elements and contained suitable methods and techniques for the topic. The module outlines were presented and discussed with the total group of writers. Portions of the module, such as activities crucial to learning particularly difficult contents, were demonstrated.

The modules were revised according to the evaluation and suggestions of the expert's panel and colleagues. The CEM coordinating staff assisted in monitoring the completion of the revised product. Equipment and materials were identified and prepared for the field testing of the modules.

Field Testing of Modules. Each of the revised modules was tried by the writer on a faculty group representing the target clients. One or two teachers in the same subject, with training experience, acted as critic-observer(s) during the trial run. The quality of the written module and its effectiveness on the participants were evaluated. The pilot run afforded a chance to test the exercises in a school setting. Wherever possible, some activities were tried by the trainer on actual students. The materials and facilities of the school were also used. Content and strategies, theories and research as well as use of materials, equipment and technology were discussed with the participants. This way, the teachers were provided the chance to deepen understanding of theory and research behind the knowledge and skills being taught to collaborate in solving problems and connect all these to the focus on improvement of student learning. All these would contribute towards a more effective program design. (Hawley & Valli, 1999)

The ensuing critiques were discussed with the experts panel once again and the writer made revisions anew, where necessary. Some modules experienced a number of alterations in techniques; in other cases, new activities were introduced, with notes on variations for certain topics and teaching levels. The pilot run also served as an opportunity to identify new recruits for the pool of facilitators. The observers in some sessions evolved as trainers in later implementation of the PEPS.

Production of final modules. Each module was subjected to close review as to compliance with the required format and completeness of elements. The introductory note included the purpose, the target client, the instructional objectives and prerequisite skills. The text included the syllabus

and the details of content and activities. The contents were checked by experts for accuracy, theoretical and pedagogical support. The lessons should contain relevant examples from learners' environment and previous knowledge, illustrations that are suitable, and self-learning devices and techniques that arouse critical thinking, all organized in logical and coherent sequence. They should also include research, readings and assessment devices. Technology and media support are indicated where applicable and facilitative of learning.

The CEM coordinating staff secured all materials relevant to the production of the final modules. Every page of the text, appendices, worksheets, handouts, trainer and learner guides were reviewed and edited, visual aids, CD's were screened for quality. The final form of the module was produced in hard and soft file copies. Trainer's guides and participant's handouts and worksheets were reproduced and bound into sets for the desired number per training session.

III. Implementation of PEPS

The PEPS development project was launched in September 2003 during a national conference of the CEM to its test subscribers and a survey was made of the schools' interest in particular topics. The resulting demand profile was used to prepare for the first schedule of seminars in the series. The field test of the modules proceeded till the end of Year 2004. The final implementation of the series started in 2005.

The PEPS seminar-workshops were offered at subsidized rates for the CEM test user schools; nonetheless, the response to the PEPS was slow. Besides the cost, the 5-day format rendered it difficult for some schools to release their faculty. This led to the revision of modules to adapt to a 4-day and a 3-day format, with corresponding reduction of coverage. Some strategies were also tried in terms of discounted rates and subsidies for participants.

Over the years, from 2005 to 2007, a total of 23 seminars were conducted and in the four subjects, participated in by 138 schools and 232 teachers. (See Figure 3 and Table 3). The most popular seminars were those in Mathematics and English. This may indicate the priority that schools themselves give to these core subjects in their efforts at improvement.

Table 3

Summary of Seminars, Participants, and Schools per Year across Subjects

	English			Mathematics			Science			Filipino			Total		
	Seminars	Participants	Schools	Seminars	Participants	Schools	Seminars	Participants	Schools	Seminars	Participants	Schools	Seminars	Participants	Schools*
2005	2	32	21	2	25	22	1	12	12	1	6	6	6	75	61
2006	3	27	10	5	45	25	2	13	6	2	22	10	12	107	51
2007	1	13	9	3	26	9				1	1	8	5	50	26
Total	6	72	40	10	96	56	3	25	18	4	39	24	23	232	138

*Schools may have been counted more than once if they have sent participants to more than one seminar

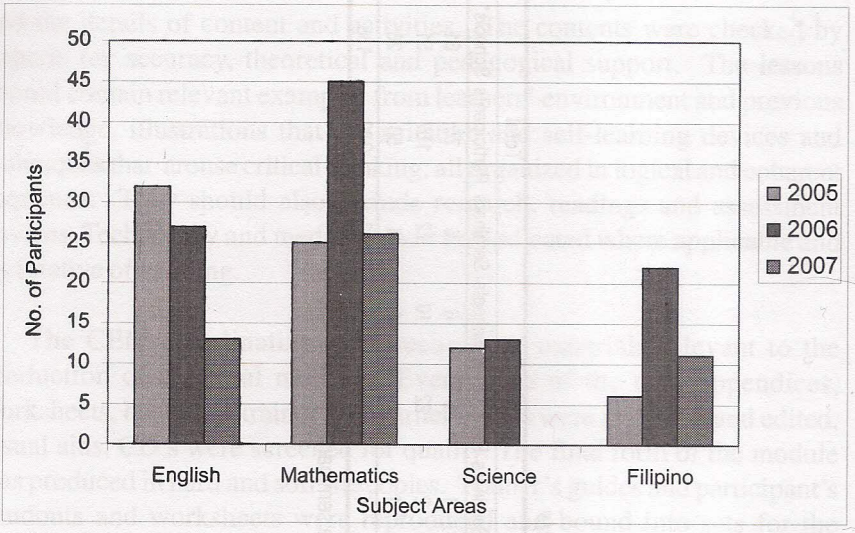


Figure 3. Distribution of Participants per Subject from 2005 to 2007

Before each seminar, the participants respond to a training needs assessment survey and a diagnostic pretest on sample competencies to be covered. Some tests are of the objective multiple-choice type. The facilitator also starts the seminar with a leveling of expectations. The results of all these provide him with baseline information about the teachers and enable him to alter some specific objectives or add supplementary ones, as needed, and conduct the session in ways that would benefit the participants most.

The trainers kept to the scheduled activities as closely as needed to attain the objectives. Daily feedback is done and discussions about the content, strategies, and participants' skills and behavior serve to deepen their learning, correct misconceptions and lead to a greater appreciation of the subject. Through the posttest and the sample lesson presentations, the participants demonstrate knowledge and skills acquired as well as their ability to transfer learning to students upon return to their schools.

IV. Evaluation of the PEPS

Results of Pretest and Posttest. As earlier mentioned, some trainers used an objective test on the coverage of the content areas to be learned during the seminar. Some used this as diagnostic pretest, others as posttest to gauge what was retained, and still others in a pretest-posttest design.

Table 4

Results of Pretest and Posttest in Three Modules

Seminars	No. of Items	Mean Percent Correct Pretest	Posttest	Difference	Proportion of participants with increase in scores
Teaching and Testing Reading, Grammar, Vocabulary	75	72%	72%	0	50%
Physical Earth, Its Weather and Climate	40	35%	46%	11%	80%
Number Theory and Rational Numbers	35	48%	62%	14%	100%

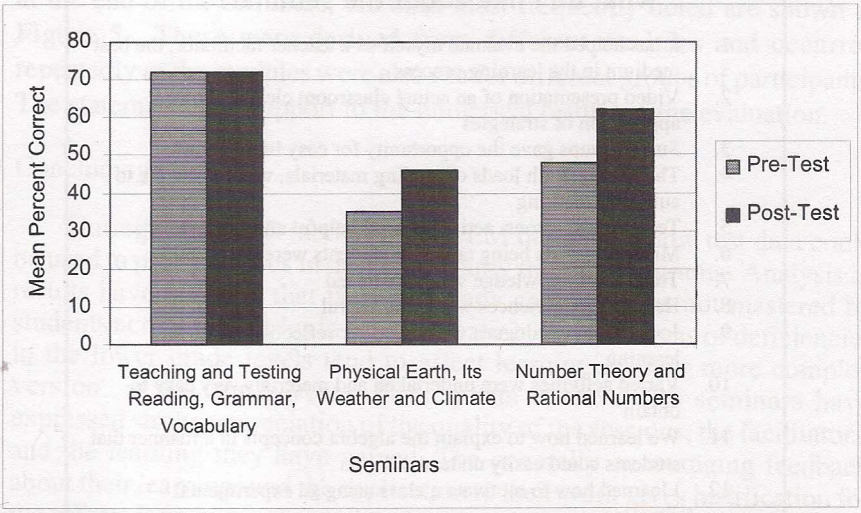


Figure 4. Comparative results of pretest and posttest in three modules

Table 4 shows the comparative pretest-posttest results for three seminars, in English, Science and Mathematics.

For the English module, the participants' performance in both pretest and posttest remained at the same score level (72%) while the science and mathematics groups appear to have benefited from the sessions, gaining a few points beyond their low pretest scores. It is also interesting to note the proportion of teachers who showed improvement, with all the Mathematics participants registering an increase in performance.

Results of Seminar Evaluation. Aside from the daily feedback form, the critiquing after each activity, the content discussions, and the journals,

Table 5

Evaluation Summary Results of Sample Modules in Core Subjects

Subject Areas	Quality of Sessions	Effectiveness of Strategies	Applicability to work	Resource Person	Training Management	Overall Rating
Mathematics	3.8	3.8	3.9	3.9	3.8	3.8
English	3.8	4.0	4.0	3.9	3.8	3.8
Science	3.9	3.9	3.9	3.9	3.6	3.8
Filipino	3.7	3.6	3.8	4	3.8	3.9
Average	3.8	3.8	3.9	3.9	3.8	3.9

Note: 4 = excellent; 3 = good; 2 = fair; 1 = poor

What do I like about the seminar?

1. It has helped me evaluate myself as a teacher facilitator, the best medium in the learning process
2. Video presentation of an actual classroom clearly showed application of strategies
3. Small groups gave the opportunity for easy interaction
4. The setting, with loads of reading materials, was conducive to authentic learning
5. Techniques, games, activities were helpful and interesting
6. Misconceptions being taught to students were corrected
7. Transfer of knowledge was maximized
8. Hands-on experiences were very useful
9. Journals and dialogues were good for reflecting on one's learning
10. Varied activities were undertaken and materials very easy to obtain
11. We learned how to explain the algebra concepts in a manner that students could easily understand
12. I learned how to motivate a class using an experimental situation
13. We were able to try different ways of solving word problems

Figure 5. Some Participants' Feedback on the PEPS Seminars

an evaluation form was accomplished by the participants at the end of the seminar.

The following results are derived from the responses of participants to the evaluation form. Each specific content covered in the area module was judged as to quality of delivery, strategies involved, and usefulness to one's situation in school. The resource person or trainer was also evaluated in various aspects, such as mastery of subject, communication skills, teaching enthusiasm. Training management had to do with the smoothness of running

the activities, materials and visuals, venue and time management. A summary evaluation of a sample module from each of the core subjects is shown in Table 5. The Mathematics module was on Number Theory and Rational Numbers, for English – Reading Notes and Strategies, for Science- Genetics, and for Filipino – Panitikan (Literature). The ratings given by participants were on a 4-point scale, and the mean score for each specific content was obtained. The overall mean for each subject module and for each criteria are all close to Excellent. This is evidence of the generally positive regard the participants had of the PEPS seminars.

From the written comments and suggestions given by the participants at the end of the seminars, the ones most frequently noted are shown in Figure 5. These were derived from different modules and occurred repeatedly as the modules were also run for different groups of participants. The statements add support to the numerical results of the evaluation.

Concluding Remarks

Through the PEPS modules, the CEM has shown how test data could be used to provide focus in desiging faculty training programs. Analysis of results have revealed that there are many competencies not mastered by students across the grades in basic education, and that areas of deficiencies in the lower grade levels tend to affect learning of their more complex version in the upper years. Participants to the PEPS seminars have expressed strong appreciation of the quality of the sessions, the facilitators, and the learning they have gained. The generally encouraging feedback about their learnings and the application to their work gives justification for the efforts invested in the project.

The PEPS program, however, is intended for a long-term impact on student learning. It is to train teachers who will in turn implement new ways of helping their own students. There is need to proceed with a more powerful evaluation that focuses on the effect on student achievement. (Killion, 2002). This may be achieved by involving the participants in monitoring activities after the training. Lessons under the areas studied could be scheduled, and peer or mentor observation could be done during the actual lessons. A pretest-posttest design for target competencies may also be used for assessing the impact of the training on the students. The assessment could include measures of knowledge, skills and attitudes related to the subject matter.

To extend the benefits of the program, it would be well to analyze more recent results of students' tests on the core subjects and proceed to devise in-service training to match the new areas of concern. Offering seminars

on topics that meet their students' needs may be more appreciated by the schools. The continuous efforts at utilizing test data will lead to more focused and improved teaching interventions. It will not hurt to increase the library of modules and revise existing ones with versions adapted to actual needs in the school situation.

The PEPS could be part of a more comprehensive school improvement program which includes staff development, academic supervision, instruction, assessment, and research. To this end, support could be obtained from public or private sectors for the conduct of the seminars in schools where teachers could benefit from more intensive and stimulating training.

The country has been challenged with discouraging outcomes of student learning in both local and international measures, prompting top education leaders to call on all sectors to help bring back quality to basic education (Lapus, 2007). The education department has set in motion a basic education reform agenda which highlights a national strategy in support of learning in the core subjects, and the structural support that goes with this (DepEd, 2005).

Amidst these efforts, policy makers must remember that the kinds of change that really matter are those that build teacher capacity and professional culture. There are no shortcuts to educational improvement (Ingvarson, 2005). Evidence-based decision making in professional development, as discussed in this paper, is certainly a move in the right direction.

References

DepEd (2005). Basic Education Sector Reform Agenda (2006-2010). Department of Education Republic of the Philippines. Pasig City.

Edelstein, W. (2002). Development as the aim of education – revisited. In Oster,F.,Dick, A., Patry, J. (Eds.) *Effective and Responsible Teaching, the New Synthesis*. Joey Bass Pubs. San Francisco.

Hawley, W. and Balli, L. (1999) The essentials of effective professional development : a new consensus. In Darling-Hammond, Linda and Sykes, Gary. Editors. *Teaching as the Learning Profession, Handbook of Policy and Practice*. Josey-Bass Publishers, San Francisco.

Ingvarson, L. (2005). Getting professional development right. In Getting Data to Support Learning Conference.(2005) ACER , Melbourne, Australia.

Killion, Joellen. (2002). For Good Measure. National Staff Development Council.

Lapus, Jesli (2005). Save our schools for quality education. *Education News.* Manila.

TDD–CEM (2001). The Diagnostic Tests. Test Development Division. Center for Educational Measurement, Inc., Makati City.

Vinson, Cindy, (2002). Instructional design, educational strategies, and hybrid instruction. De la Salle University Press. Manila.

Ms. Esperanza C. Buen was formerly the Vice President for Operations at the Center for Educational Measurement, Inc.

Appendix A1

Areas of Concern in English across Elementary Grade Levels

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Noun (48)					
Adjective (52)	Adjective (43)	Adjective (51)			
Verb (44)	Verb (50)	Verb (45)	Verb (46)	Verb (50)	Verb (46)
Preposition (54)		Preposition (48)		Preposition (42)	
		Sentence (38)	Sentence (48)	Sentence (52)	Sentence (47)
					Paragraph Dev't. (51)
		Structural Analysis (49)		Structural Analysis(58)	
			Vocabulary (40)	Vocabulary (52)	Vocabulary (49)
			Understanding a selection (45)	Understanding a selection (51)	
Sequencing Events (53)					
Norm (58)	**Norm (52)**	**Norm (52)**	**Norm (51)**	**Norm (59)**	**Norm (54)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix A2

Areas of Concern in English across Secondary Year Levels

Year 1	Year 2	Year 3	Year 4
Simple Tenses (48)	Tenses and voice (51)		Grammar and Usage (36)
	Modal Auxiliaries (48)		
	Modification (51)		
Subject & Predicate Agreement (53)	Special Rules on Agreement (47)		
Sentence (49)	Sentence (52)	Reported Speech (51)	
		Paragraph Development (45)	
		Brevity & Parallelism (50)	
Vocabulary (48)			Reading Comprehension (54)
	Recognition of Main Idea & Details (45)	Recognition of Main Idea & Detail (43)	
Inferring (47)	Inferring (43)	Inferring (50)	
Reorganization (50)	Reorganization (42)	Reorganization (51)	
Norm (54)	**Norm (55)**	**Norm (52)**	**Norm (55)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix B1

Areas of Concern in Filipino across Elementary Grade Levels

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Pagkilala ng Salita (66)					
Pangkat ng mga Salita (78)					
Balarila (75)		Pangngalan (55)	Pangngalan (48)		
	Pandiwa (64)			Pandiwa (57)	Pandiwa (53)
	Pang-uri (64)	Pang-uri (58)		Pang-uri (46)	
		Pang-abay (45)		Pang-abay (54)	
					Pangatnig (42)
	Pangungusap (61)	Pangungusap (55)	Pangungusap (51)	Pangungusap (48)	Pangungusap (50)
		Pagsusuring Pangkayarian (45)		Pagsusuring Pangkayarian (56)	
					Pagsulat (43)
	Talasalitaan (58)	Talasalitaan (51)	Talasalitaan (51)	Talasalitaan (52)	Talasalitaan (49)
	Pag-unawa sa Babasahin (62)	Pag-unawa sa Babasahin (42)	Pag-unawa sa Binasa (42)	Pag-unawa sa Binasa (43)	
	Pagkasunud-sunod ng mga Pangyayari (57)				
Norm (82)	**Norm (69)**	**Norm (60)**	**Norm (58)**	**Norm (60)**	**Norm (55)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix B2

Areas of Concern in Filipino across Secondary Year Levels

Year 1	Year 2
Pang-abay (43)	
Pangungusap (60)	
Pagbuo ng Talata (43)	
Talasalitaan (51)	
Kahulugan ng Idyoma at Salawikain (62)	
Pag-unawa sa Babasahin (55)	Pag-unawa sa Babasahin (66)
	Kasanayan sa Pag-aaral (70)
Norm (67)	**Norm (71)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix C1

Areas of Concern in Science across Elementary Grade Levels

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Plants (57)	Plants (45)		Plants(41)	Plants (50)	
Animals (58)	Human Being (53)			Animals (47)	
Matter (60)		Matter (50)	Matter (43)	Matter (48)	Matter (55)
				Energy (42)	Energy (43) Heat (46)
				Force & Motion (38)	Force & Motion (47)
		Light & Shadow (51)			
		Magnets (44)			Magnets (47)
			Electricity (46)		
				Weather & Climate (41)	Weather & Climate (48)
					Ecosystem (49)
	Earth & Space (51)	Earth (50)	Earth (46)	Earth (40)	The Earth (48)
		Solar System (41)	Space (39)		Solar System & Beyond (45)
Norm (61)	**Norm (56)**	**Norm (55)**	**Norm (50)**	**Norm (50)**	**Norm (59)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix C2

Areas of Concern in Science across Secondary Year Levels

Year 1 (General Science)	Year 2 (Biology)	Year 3 (Chemistry)	Year 4 (Physics)
Nature of Matter (42)	Nature of Biology (47)	Atom (43)	Modern Physics (47)
	Chemical basis for life (50)	Symbols, Formulas, & Equations (37)	Force & Energy (44)
	Energy transformation (50)	Phases of Matter (38)	
Earth & the Universe (45)	Organ systems (45)	Chemical Bonding (41)	Electromagnetic Energy (42)
	Reproduction (49)	Types of Chemical Reactions (42)	Waves & Energy (47)
	Heredity, Variation, & Population (43)	Chemical Kinetics (30)	Energy & Communication (31)
	Evolution (49)	Electrochemistry (37)	
	Diversity & Adaptive Forms of Living Things (43)	Carbon and its compounds (48)	
Norm (52)	**Norm (51)**	**Norm (50)**	**Norm (49)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix D1

Areas of Concern in Mathematics across Elementary Grade Levels

Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
	Numeration system (56)	Numeration system (43)	Numeration system (50)		
Sets(65)					
Money (60)		Decimals(32)	Decimals & Ratio (43)	Decimal & Percentage (45)	Decimal, Ratio, & Percentage (48)
	Parts of Whole Objects (56)		Fractions (51)	Fractions (34)	Fractions (42)
Subtraction (61)	Fundamental Operations (55)	Fundamental Operations on Whole Numbers (50)			
			Geometry (48)	Geometry (38)	
			Graphs, Maps, & Scales (47)		Graphs, Maps, & Scales (47)
Norm (72)	**Norm (59)**	**Norm (51)**	**Norm (51)**	**Norm (62)**	**Norm (57)**

Note: Mean Percent Correct Score enclosed in parenthesis

Appendix D2

Areas of Concern in Mathematics across Secondary Year Levels

Year 1	Year 2	Year 3	Year 4
Factors & Primes (43)		Exponents and radicals (46)	Complex Number (38)
Rational Numbers (44)		Rational Expressions (35)	
Non-metric Geometry (46)	Geometry concepts (46)		
	Quadrilaterals (42)		
	Triangle Congruence (41)	Similarity (47)	
Language of Algebra (33)	Algebraic Expressions (46)		
Equations & Inequalities (34)		Quadratic Equations (44)	Circular Functions (40)
	Statistics (40)	Quartiles, Percentiles & Deciles (37)	
		Sequence (46)	
		Variations (43)	Linear Correlation (38)
Measurement (30)			
Norm (55)	**Norm (51)**	**Norm (50)**	**Norm (49)**

Note: Mean Percent Correct Score enclosed in parenthesis

REVISITING THE PHILIPPINE APTITUDE CLASSIFICATION TEST: ANALYSIS OF POTENTIALLY BIASED TEST ITEMS[1]

Cecile Gatchalian and Armi Lantano
Center for Educational Measurement, Inc.
Philippines

Differential item functioning (DIF) indicates that an item is potentially biased. To gain a better understanding of the behavior of the subgroups on potentially biased items, the distractor response analysis can be conducted. In this procedure, the incorrect alternatives are examined for differences in patterns of responses among different subgroups of a population. Bias is inferred to be present in an item when the subgroups are differentially attracted to the distractors of the item.

This follow-up study on the fairness of the Philippine Aptitude Classification Test (PACT), a battery of tests measuring different kinds of aptitudes or abilities, aims to gain more substantive evidence of item bias in the test. It further investigates the two PACT subtests identified in the previous study as having the most number of items with DIF, namely, the Perceptual Acuity and the Verbal Filipino. The Rasch model technique was applied on another sample of examinees, and distractor response analysis was used to examine the response pattern across the incorrect options of items with DIF among the gender and the geographic subgroups of the PACT takers. The analysis of response patterns provided insights on group differences, and implications to test construction were drawn.

The consideration of fairness is important in sound testing practice. A test should be constructed such that differences between groups' test scores are due to differences in what the test measures. Test developers have to ensure that items possessing unfair biases will be detected and replaced by

[1]Paper read at the International Association for Educational Assessment (IAEA) Annual Conference at Bangkok, Thailand, August 2010.

items that are less likely to contain such biases. Developing tests that are as fair as possible contributes to the societal goal of equal opportunities for all.

A well-made test addresses the concern on fairness from test design to testing outcomes. It is subjected to careful review and empirical checks to minimize bias. Bias happens when the psychometric properties of a test or when the manner in which it is used results in different meanings for scores earned by members of different subgroups (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999). Evidence of bias may be found by studying how the test scores are used or by examining the test items themselves.

Differential item functioning (DIF) is an indicator that an item is potentially biased. To gain a better understanding of the behavior of subgroups on potentially biased test items, Osterlind (1983) recommends that, after an initial investigation of DIF in all of the test items, a distractor response analysis be conducted on the items with DIF. In this procedure, the incorrect alternatives are examined for differences in patterns of responses among the different subgroups. Bias is inferred to be present in an item when the subgroups are differentially attracted to the distractors of the item.

Background

The Philippine Aptitude Classification Test (PACT) is a battery of tests that measure different kinds of aptitudes or abilities important in college and vocational work. The preliminary investigation of its fairness (Franco & Lantano, 2009) resulted in the detection of items with gender and geographic DIF. Gender and geographic location were chosen as subgroupings as local studies (Caoli-Rodriguez, 2007; Hicap, 2006 as cited in Franco & Lantano, 2009) showed significant achievement differences among Filipino students across these variables.

Among the three DIF-detection procedures used, the Rasch model technique was the most sensitive. Out of 210 items, it identified 85 items with gender DIF and 55 items with geographic DIF. The subtest Perceptual Acuity contained the most items with gender DIF, and the subtest Verbal Filipino contained the most items with geographic DIF.

The present study aims to gain more substantive evidence of item bias in the PACT. It further investigates the two PACT subtests identified in the previous study as having the most number of items with DIF, namely, the Perceptual Acuity and the Verbal Filipino. Applying statistical procedures, it seeks to find out how the test items behave for the different subgroups. Specifically, it describes the items that are found to be easier for the different subgroups and determines if the difficulty of an item for a subgroup stems from an internal characteristic of the item. Based on technical criteria, potentially biased items will be detected and corrected or replaced by items that are less likely to contain bias.

Methods

Subjects

A sample of 2,008 third year high school students was randomly chosen from the total PACT takers of school year 2008-2009. The sample has an approximately 6:4 ratio of both male – female and National Capital Region (NCR) – other regions (non-NCR) subgroups.

The NCR is one of the 17 regions of the Philippines. The division of the Philippine archipelago into regions is based on the geographical, cultural, and ethnological characteristics of the provinces (Philippine Travel Photos, n.d.). The NCR is mainly Metropolitan Manila. It is the center of politics, economics, society, and culture of the country.

Instrument

The PACT is a battery of tests that measure different kinds of aptitudes or abilities important in college and vocational work. The subtest Perceptual Acuity measures the ability to visually detect the pattern of change or the differences in given figural or semantic stimuli. It consists of two item types: Figure Series and Proofreading, each with 15 items. The subtest Verbal Filipino measures the ability to understand Filipino, the national language of the Philippines. It consists of two item types: Talasalitaan (vocabulary) and Mga Salitang Magkaugnay (verbal analogies). Talasalitaan is further subdivided into Kasingkahulugan (synonym) and Kasalungat (antonym). Verbal Filipino consists of 19 items: 6 Kasingkahulugan, 4 Kasalungat, and 9 Mga Salitang Magkaugnay.

Procedure

The Rasch model technique was used to identify the items that exhibit DIF. The subtests Perceptual Acuity and Verbal Filipino were analyzed for gender and geographic DIF, respectively. With the Rasch model, the item difficulty of each item was calibrated separately for each of the gender (male vs. female) and geographic location (NCR vs. non-NCR) subgroups. The difficulty indices were tested for statistical significance using the chi-square. The items that showed a significant difference between the subgroups were flagged as potential DIF items.

Moreover, the items that consistently exhibited DIF in the initial study and in this study were analyzed by examining the response pattern across the incorrect options among the gender and the geographic location subgroups of the PACT takers. Employing the distractor response analysis, it determined significant difference between two groups. It involved the following steps: (a) preparing a matrix of choice response alternatives for the test items under study, (b) placing the data in a series of 2x2 contingency tables for significance testing, and (c) hypothesis testing. In this study, the Mantel-Haenszel procedure was used to test if there existed significant difference in the proportions selecting distractors on a test item between identified subgroups of the population.

Results

DIF Analysis

The analysis on Perceptual Acuity for the gender subgroups resulted in the identification of 14 items with DIF. (Refer to Table 1) The results of 13 of these items were consistent with those in the initial study. Six (6) were Figure Series items, while seven (7) were Proofreading items. All of the Figure Series and one Proofreading items were consistently easier for males. Whereas, most of the Proofreading items (6 out of 7) were consistently easier for females. These results indicate that Figure Series items were more difficult for female examinees, whereas Proofreading items were more difficult for male examinees.

On the other hand, the analysis on Verbal Filipino for the geographic subgroups resulted in the identification of 11 items with DIF. (Refer to Table 2.) The results of 10 of these items were consistent with those in the

initial study. Both Talasalitaan (vocabulary) and Mga Salitang Magkaugnay (verbal analogies) have 5 items with DIF. Majority of these items (8 out of 10) were consistently easier for the NCR subgroup.

Table 1
Summary of Gender DIF Analysis

Item No.	Subtest	Item Difficulty			Easier for	Comparison with Previous
		Male	Female	Difference		
3	Figure Series	-1.08	-0.63	-0.45*	Male	Consistent
5	Figure Series	-0.18	0.23	-0.40*	Male	Consistent
6	Figure Series	-0.12	0.22	-0.34*	Male	Consistent
7	Figure Series	-0.23	0.05	-0.28*	Male	Consistent
8	Figure Series	-0.49	-0.13	-0.35*	Male	Consistent
9	Figure Series	-0.27	0.10	-0.37*	Male	Consistent
16	Proofreading	-1.42	-1.81	0.39*	Female	Consistent
18	Proofreading	-1.07	-1.50	0.43*	Female	Consistent
19	Proofreading	-0.60	-0.93	0.33*	Female	Consistent
21	Proofreading	-0.57	-0.99	0.42*	Female	Consistent
22	Proofreading	-0.48	-0.95	0.46*	Female	Consistent
25	Proofreading	-0.50	-0.86	0.37*	Female	Consistent
29	Proofreading	0.98	1.28	-0.30*	Male	Consistent
30	Proofreading	1.41	1.74	-0.32*	Male	

* Significant at 0.01 alpha level

Table 2
Summary of Geographic DIF Analysis

Item No.	Subtest	Item Difficulty			Easier for	Comparison with Previous
		NCR	non-NCR	Difference		
1	Talasalitaan	-1.79	-1.34	-0.45*	NCR	Consistent
4	Talasalitaan	0.18	1.54	-1.36*	NCR	Consistent
9	Talasalitaan	-0.93	-0.36	-0.58*	NCR	Consistent
10	Talasalitaan	-0.91	-0.54	-0.37*	NCR	Consistent
11	Talasalitaan	-1.09	-1.62	0.53*	non-NCR	Consistent
19	Mga Salitang Magkaugnay	-1.06	-0.19	-0.87*	NCR	Consistent
20	Mga Salitang Magkaugnay	-0.15	0.74	-0.89*	NCR	Consistent
21	Mga Salitang Magkaugnay	-0.09	0.27	-0.36*	NCR	Consistent
22	Mga Salitang Magkaugnay	-0.67	-0.22	-0.45*	NCR	Consistent
23	Mga Salitang Magkaugnay	-0.24	-0.57	0.32*	non-NCR	Consistent
25	Mga Salitang Magkaugnay	0.70	0.41	0.29*	non-NCR	

* Significant at 0.01 alpha level

Distractor Response Analysis

Results of the distractor response analysis showed that 11 of the 13 items with gender DIF have significantly different response patterns for male and female examinees. (Refer to Table 3.) All of the Figure Series items and one Proofreading item were found biased against female examinees. In these items, a significantly large number of female examinees were attracted to the incorrect options, indicating unfamiliarity with the concept reflected by the items. Conversely, most of the Proofreading items (6 out of 7) were found biased against male examinees. This indicates that male examinees have a significant tendency to favor other options over the correct options in these items. Hence, male examinees exhibit less familiarity with the concept reflected in the said items.

Table 3

Summary of Distractor Response Analysis for Gender

Item No.	Subtest	Correct Option	Easier for	Attractive Distractor(s) to the Other Group
3	Figure Series	3	Male	4,5
5	Figure Series	4	Male	2
6	Figure Series	3	Male	Not significant
7	Figure Series	1	Male	Not significant
8	Figure Series	5	Male	2
9	Figure Series	2	Male	5
16	Proofreading	1	Female	2,3,5
18	Proofreading	1	Female	2,3
19	Proofreading	3	Female	1,2
21	Proofreading	2	Female	1,3
22	Proofreading	2	Female	1,3
25	Proofreading	2	Female	1,3
29	Proofreading	3	Male	1

Based on the result of distractor response analysis on Verbal Filipino, only those items with geographic DIF favoring the NCR group were found significant. (Refer to Table 4.) Both items found to be easier for the non-NCR subgroup did not have significantly different response pattern when compared with their NCR counterpart. Hence, Verbal Filipino significantly favored the NCR subgroup.

The distractor response analysis on Verbal Filipino further revealed 8 items significantly showing bias between the NCR and the non-NCR

Table 4
Summary of Distractor Response Analysis for Geographical Location

Item No.	Subtest	Correct Option	Easier for	Attractive Distractor(s) to the Other Group
1	Tala salitaan	3	NCR	2,4
4	Tala salitaan	1	NCR	3,4,5
9	Tala salitaan	1	NCR	4
10	Tala salitaan	3	NCR	2,4
11	Tala salitaan	4	non-NCR	Not significant
19	Mga Salitang Magkaugnay	5	NCR	3
20	Mga Salitang Magkaugnay	2	NCR	3
21	Mga Salitang Magkaugnay	5	NCR	4
22	Mga Salitang Magkaugnay	3	NCR	1,5
23	Mga Salitang Magkaugnay	4	non-NCR	Not significant

subgroups. In these items, one to three distractors obtained significantly large number of responses from the non-NCR subgroup. Thus, these results indicate that the NCR is more familiar with the concepts reflected in the said items.

Three examples of test items with gender or geographic bias are found in the Appendix. The choice-response matrix of each item is also presented..

Discussion

Gender DIF in Perceptual Acuity

The Perceptual Acuity subtest contains two item types: Figure Series and Proofreading. The Figure Series items consist of visual stimuli, while Proofreading items consist of verbal stimuli. The Figure Series items were easier for males, while the Proofreading items were easier for the females.

The behavior of the gender groups on the two different item types is consistent with the findings of scientific research on the differences between the genders in regard to the brain and cognitive functioning (Bush, n.d.; Weiman, 2004). Males are found to be right-hemisphere dependent, specializing in visual tasks, such as the Figure Series items. On the other hand, females have been shown to be left-hemisphere dependent, performing better in verbal tasks. Thus, Proofreading is easier for them.

The difference in the performance of the genders on the Perceptual Acuity subtest is illustrated in the sample analyses of a Figure Series and a Proofreading items in the Appendix. In the Figure Series item, significantly more males chose the correct answer, while significantly more females chose two distractors. In the Proofreading item, the behavior of the genders reversed. Significantly more females correctly answered the item, while significantly more males chose three distractors.

Geographic DIF in Verbal Filipino

The Verbal Filipino subtest was easier for the NCR examinees than for the non-NCR examinees. This result may be attributed to the diversity in the linguistic background of the Filipino people. There are 175 individual languages used in the different regions of the country (Philippines, n.d.). Filipino, the national language, is primarily based on Tagalog, which is just one of the eight major languages of the Philippines (Santiago & Tiangco, 1991). The facility of NCR examinees in Verbal Filipino can be explained by the fact that Tagalog is mainly spoken in the region. Filipino is less familiar in non-NCR as there are other languages dominant in these regions. Non-NCR examinees may not have sufficiently acquired the Filipino language, resulting in their low performance on the test.

The unfamiliarity of the non-NCR examinees with some of the words in the Verbal Filipino subtest is illustrated in the sample analysis of a Mga Salitang Magkaugnay (verbal analogies) in the Appendix. A significantly smaller percentage of non-NCR examinees correctly answered the item. Moreover, two distractors were significantly chosen by more non-NCR examinees. These two words are also related to the third word in the stem, but only the correct option is related to the third word in the same way the second word is related to the first word in the stem. The non-NCR examinees seem to have difficulty in determining the fine distinctions among the correct answer and the two distractors.

Conclusion

In this study, differentially functioning items in Perceptual Acuity and Verbal Filipino were further analyzed. The results illustrated gender differences in cognitive functioning and diversity in the linguistic background of the Filipino people.

Through the distractor response analysis, bias in several items was confirmed. In some items, a single distractor is significantly more attractive to a subgroup. These distractors have to be replaced in order to correct the bias. In a couple of cases, an item has two or more significantly more attractive distractors. These items will have to be replaced.

In addition, the following implications to test construction can be drawn:

(1) In developing test items, the inclusion of visual stimuli, even in achievement tests, should be considered so as not to disadvantage male examinees.

(2) In developing tests in Filipino, the use of words familiar nationwide has to be seriously considered.

Test fairness is one of the standards for educational and psychological tests. The selection of items based on discrimination and difficulty indices alone is not sufficient to ensure this. DIF and distractor analysis provide additional information on which items are appropriate to the target examinees.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: Author.

Bush, T. (n.d.). Gender differences in cognitive functioning: a literature review. *Cyber-Journal of Sport Marketing*. Retrieved June 7, 2010, from http://fulltext.ausport.gov.au/fulltext/1997/cjism/v1n1/tbush1.htm

Franco, L. M. & Lantano, A. S. (2009). Study of fairness in the Philippine Aptitude Classification Test. Paper presented at the 35th IAEA Annual Conference. Retrieved March 15, 2010 from http://static.icms.com.au/iaea2009/papers/323.doc

Osterlind, S. J. (1983). Test item bias. Newbury Park, CA: Sage.

Philippine Travel Photos. (n.d.). Retrieved June 19, 2010, from http://www.philippine-portal.com/regional-information/regions.html

Philippines. (n.d.). Retrieved June 24, 2010, from http://en.wikipedia.org//wiki/Philippines

Santiago, A. O. & Tiangco, N. G. (1991). *Makabagong balarilang Filipino*. Manila, Philippines: Rex Book Store.

Weiman, H. (2004, May 10). Gender differences in cognitive functioning. Retrieved June 7, 2010, from http://homepages.luc.edu//~hweiman/Gnederdiffs.html

---

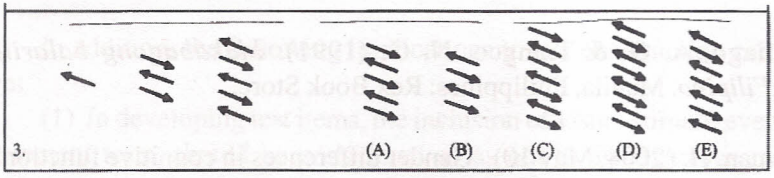
Ms. Cecile Gatchalian works in the Aptitude & Special Tests Unit of the CEM Test Development Section which currently maintains 14 aptitude tests. She has been in service for over 21 years. She holds a Master of Science degree in Psychology, major in Psychological Measurement which she acquired from the De La Salle University-Manila.

Ms. Armi Lantano holds the position of Test Specialist III in the Statistical Processes and Item Banking Unit of the CEM Test Development Section. She has been with CEM for more than five years. She obtained her undergraduate degree in Bachelor of Science in Mathematics from the University of the Philippines–Diliman and is currently taking up Masters of Statistics in the same university.

Appendix A

Example of Biased Items with Corresponding Choice-Response Matrix

FIGURE SERIES ITEM



Choice-Response Matrix

	Response Choice					
Groups	A	B	C*	D	E	Omits
male	3.44%	5.40%	70.33%	9.02%	11.07%	0.74%
female	3.78%	3.55%	64.26%	14.09%	13.52%	0.80%

PROOFREADING ITEM

16. Security Pacific Bank
    333 South Hope Street
    Los Angeles, California

    (A) Security Pacific Bank
    333 South Hope Street
    Los Angeles, California

    (B) Security Pacific Bank
    333 South Cope Street
    Los Angeles, California

    (C) Security Pacific Bankers
    333 South Hope Street
    Los Angeles, California

    (D) Security Pacific Bank
    333 South Hope Street
    Los Afigeles, California

    (E) Securities Pacific Bank
    333 South Hope Street
    Los Angeles, California

Choice-Response Matrix

	Response Choice					
Groups	A*	B	C	D	E	Omits
male	76.24%	8.49%	7.16%	3.44%	3.91%	0.76%
female	83.29%	6.16%	4.86%	2.49%	2.13%	1.07%

MGA SALITANG MAGKAUGNAY ITEM

22. BAYANI : MAGITING ::
PULUBI :

(A) mangmang
(B) mahiyain
(C) gusgusin
(D) baliw
(E) gala

Choice-Response Matrix

	Response Choice					
Groups	A	B	C*	D	E	Omits
NCR	21.44%	3.47%	67.97%	2.12%	4.66%	0.34%
non-NCR	25.46%	5.51%	48.96%	3.18%	16.03%	0.86%

SURVEY OF READING HABITS AND COMPETENCE OF FILIPINO CHILDREN[1]

Janet T. Evasco and Analyn P. Capalihan
Center for Educational Measurement, Inc.
Philippines

This study provides a profile of Filipino primary pupils in terms of reading competence and reading habits. Locally developed standardized reading tests were used to measure reading competence. A questionnaire was also fielded to collect data on reading habits. The instruments were administered towards the end of the school year to a cohort group of 899 pupils when they were in their fourth and sixth grades of elementary education. The reading competencies that were measured include scanning, vocabulary, comprehension, and study aids. The reading habits that were surveyed were number of hours spent daily on reading, reading companion, kinds of reading materials, and type of reading strategies used. The results showed that majority of the sample pupils demonstrated relatively competent reading abilities when they were in grades 4 and 6. Likewise, the patterns of reading habits of the cohort are relatively consistent as they move from grade 4 to grade 6. The results also implied that reading habits tend to influence reading competence.

The assessment of reading competence in grade school is important because it is at this stage where students are expected to have acquired adequate reading skills in order to learn effectively. In the Philippines, grade 4 and grade 6 are considered exit points of the primary and the intermediate levels of basic education, respectively. As one moves to higher educational levels, with more subject areas covered, the demand for good reading skills is imperative. Equally important is the practice of good reading habits which plays a crucial role in learning. Positive reading habits sustain the acquisition of new vocabulary; improve spelling, grammar, and word analysis skills; and enhance higher levels of understanding and reasoning.

[1]Paper read at the International Association for Educational Assessment (IAEA) Annual Conference at Bangkok, Thailand, August 2010.

Not much is known about the reading habits and preferences of children in the Philippines. This study thus intends to profile the reading habits of a cohort group of pupils and explore the influence of these habits to reading competency as the group progresses from grade 4 to grade 6. This is deemed important in the light of very little improvement in the achievement levels of elementary students. The results of 2009 National Achievement Test described the current achievement levels of elementary students as "near mastery level" only, albeit an upward trend of 66% Mean Percentage Score from the previous 2006 results of 55%. Similarly, the data from the 2003 Functional Literacy, Education and Mass Media Survey (FLEMMS) and the Philippine-Informal Reading Inventory scores show that only one-sixth of Filipino pupils can read independently at the desired grade level (Literacy Coordinating Council, 2003). This low achievement level may be attributed to various factors including socio-economic factors, teacher-related factors, inadequate learning materials, short and congested school curriculum, and poor reading habits of learners.

Thus, it is expected that the findings of this study will inform teachers and school administrators on how to design reading programs and help students develop a more positive attitude towards reading. Identifying the reading habits of pupils and its influence on reading achievement may also prompt parents, teachers, librarian, and other stakeholders in education to look into books and other services that would accommodate the reading habits of pupils.

Methodology

Respondents

A cohort group of 899 elementary pupils who took the CEM Reading Tests when they were in grade 4 and in grade 6 constitutes the subject of this study. In terms of gender distribution, 58% of the group are females whereas 42% are males.

Data Sources

The reading competencies of the respondents were measured through locally developed standardized multiple-choice Reading Tests. These tests measure reading competencies under four content areas, namely: scanning, vocabulary, comprehension, and study aids. The Reading Tests are also

offered in three levels. The Reading Test Level 1 is designed for Grade 3 and Grade 4 students; Reading Test Level 2 is for Grade 5 and Grade 6; and Reading Test Level 3 is for students in the secondary level and for college freshmen as well. For this study, the 899 respondents took the CEM Reading Test Level 1 in SY2006-2007 when they were in the Grade 4 and the CEM Reading Test Level 2 in SY2008-2009 when they reached the Grade 6 level.

The overall score in the entire test and the scores in the different content areas are reported in terms of percent of items answered correctly. For norm-referenced interpretation of reading ability, the overall score is also expressed in standard score (that ranges from 200 to 800, with a mean of 500 and a standard deviation of 100), percentile rank, stanine, and quality index.

On the other hand, the reading habits of the students were identified using the CEM Student Information Questionnaire (SIQ). The SIQ is a 19-item checklist administered before the Reading Test. It consists of items pertaining to the respondent's demographic characteristics (i.e., number of siblings in the family, birth order, parent's occupation and educational attainment), his/her reading habits (i.e., time spent daily on reading, reading companion, kind of reading materials available at home, borrowed from the library, and enjoyed reading, and reading strategies used when a word, sentence, paragraph, or page is difficult to understand), and to the respondent's perceived self-efficacy (i.e., educational goals and self-reported academic standing). The students were made to select a single answer for some items and multiple answers for other items. For instance, only single answer is required for the time spent daily on reading and reading companion. On the other hand, only five reading materials are selected from among the reading materials available at home, borrowed from the library, or enjoyed reading. For the kind of reading strategies used, respondents are to indicate whether they use each of the listed reading strategy.

Data Analysis

Descriptive statistics including frequency counts, percentages, means, and standard deviations were computed from the scores and responses of the cohort group in the Reading Tests and Student Information Questionnaire. Since reading habits cover four variables that require different responses, two statistical tests were used to ascertain their influence on reading competence. One-way analysis of variance (ANOVA) was used

for the time spent daily on reading and reading companion. The General Linear Model univariate procedure was used to analyze the influence of reading materials and kind of reading strategies on reading competence.

Results and Discussion

Reading competence

Descriptive indices of individual overall scores in the Reading Test ranged from Low Average to Superior reading abilities when the cohort group was in fourth grade. Relatively no significant changes in the range of reading abilities were observed when the cohort group was tested again in sixth grade. The cohort group's mean overall standard score in grade 4 (mean SS=533) may have increased by 19 standard score points to 552 in grade 6 but both mean standard scores remained in the high average range.

On the other hand, their scores on the specific reading subtests show that on the average, the cohort group performed above the typical grade 4 and grade 6 pupils. Although they exhibit relative weakness in Vocabulary when they were in grade 4, all subtest scores are still above national averages. On the other hand, when they were in grade 6, Scanning is a relative strength as compared to other reading skills.

Reading habits

This section presents the pattern of reading habits of the cohort in terms of the time spent on reading, their reading companion, the kind of reading materials they have at home, borrowed from the library, and enjoyed reading, and the kind of reading strategies they used when encountering difficulties in reading. Since the Reading test performance of the cohort has been described earlier, the influence of their reading habits on their Reading scores was the focus of analysis.

Time spent daily on reading. Table 1 shows that about one-third of the cohort spends 1 to 2 hours reading daily when they were in grade 4. However, this is not at all encouraging since majority of the children read even less than an hour per day. Although a relatively similar pattern was observed when they were in grade 6, majority of them tend to spend more hours in reading. This may be an indication that either the cohort enjoyed reading or the demands of reading for academic work have increased at this grade.

Table 1 also shows that the average Reading scores vary depending on the time spent by the cohort in reading when they were in grade 4 and in grade 6. Generally, the Reading scores increase as more hours is spent on reading. In grade 4, however, spending more than four hours in reading had yielded lower mean score than spending 3 to 4 hours. This may be attributed to small sample size and large spread of scores as compared to other groups. Statistically comparing the mean scores obtained, the analysis of variance reveals that mean differences exist significantly ($F = 4.95$ $p<.05$) when the cohort is in grade 4. That is, children who spend 3 to 4 hours daily on reading have significantly higher Reading scores than children who spend less than one hour in reading. On the other hand, the variation in Reading scores are also significant when the cohort is in grade 6 ($F = 43.22$, $p<.05$). Children who spend at least one hour daily on reading had significantly higher Reading scores in grade 6.

Table 1
Mean percent correct scores of the cohort group according to time spent daily on reading (N = 899)

	Grade 4			Grade 6		
Number of hours	%	Mean	SD	%	Mean	SD
Less than 1 hour	51	65	18	47	58	16
1 to 2 hours	34	69	18	37	67	16
3 to 4 hours	10	71	19	11	72	16
More than 4 hours	5	70	22	5	76	14

Reading companion. Table 2 shows that the cohort tend to spend reading most of the time with different significant others around them when they were in grade 4. Although about 39% of them tend to read independently, majority of them spend reading either with their immediate relatives or friends.

When they were in grade 6, however, there is a marked increase in their tendency to read on their own. This may be quite expected, since at this stage of development, boys and girls tend to be independent in doing certain tasks. Peer influences likewise tend to dominate at this grade because when they have a reading companion, they tend to read with friends.

Looking at how reading companion influence the overall Reading score, the analysis of variance reveals that the average Reading scores of the cohort significantly vary with their reading companion both in grade 4 ($F = 5.93$, $p<.05$) and in grade 6 ($F = 17.94$, $p<.05$). In both grades, children

who read independently have significantly higher overall Reading scores than those who either read with relatives or friends.

Table 2
Mean percent correct scores of the cohort group according to reading companion (N = 899)

	Grade 4			Grade 6		
Reading companion	%	Mean	SD	%	Mean	SD
Nobody	39	71	18	62	67	16
Mother	19	63	18	8	55	16
Father	4	61	21	1	50	27
Brother or sister	12	66	19	8	61	16
Friend(s)	19	66	17	18	57	16
Others	7	63	19	3	56	16

Kinds of reading materials. There is a wide variety of reading materials available in the homes of the cohort (Figure 1).

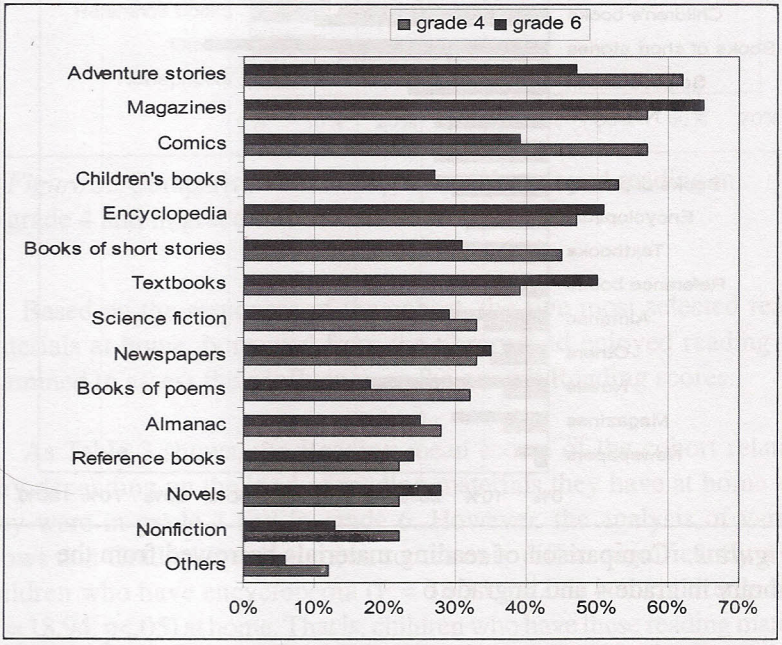


Figure 1. Comparison of reading materials available at home in grade 4 and in grade 6

Data likewise shows that when the cohort is in grade 4, most of the reading materials are for literary experience including adventure stories, magazines, and comics. However, when they were in grade 6, the reading materials available included textbooks, encyclopedia, and newspapers. Since these reading materials are generally read for information, this may indicate that at this grade, the purpose of reading has relatively changed. This is quite expected since children are towards the exit point of intermediate level of elementary education.

As Figure 2 shows, the cohort group does not generally borrow books from the library when they were in grade 4. Although about 60% of them check out adventure stories, less than half of them borrow other reading materials. One contributing factor is that most of these reading materials are already available at their homes. When they were in grade 6, however, borrowing tendencies of the cohort has slightly improved. Among the top choices included adventure stories, books of short stories, and science fiction.

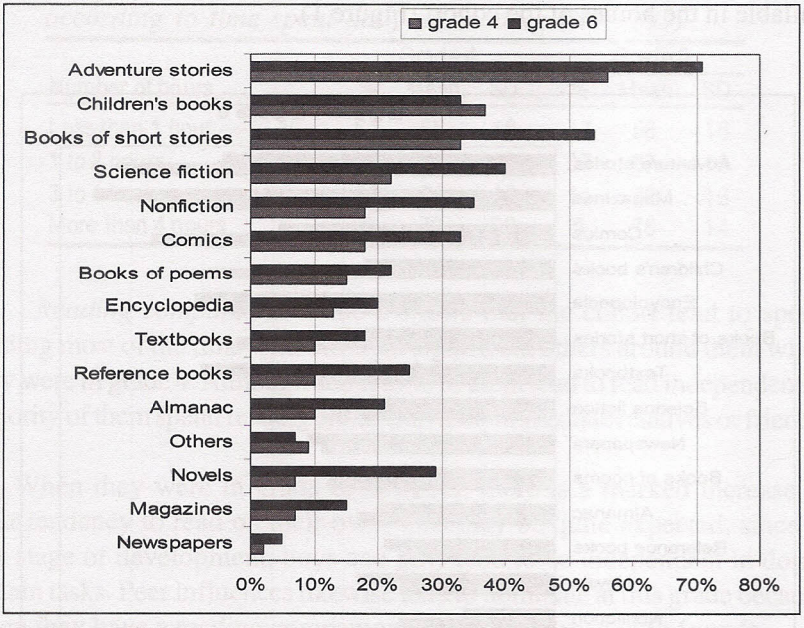


Figure 2. Comparison of reading materials borrowed from the library in grade 4 and in grade 6

As to the kind of reading materials enjoyed reading, Figure 3 shows that reading adventure stories, children's books, comics, and magazines are still

among the top choices in grade 4. When they were in grade 6, a relatively consistent pattern of reading materials enjoyed was observed. Preference to read children's books likewise declines in grade 6. Although most of the available materials at home are for information, findings suggest that children tend to prefer reading materials for literary experience.

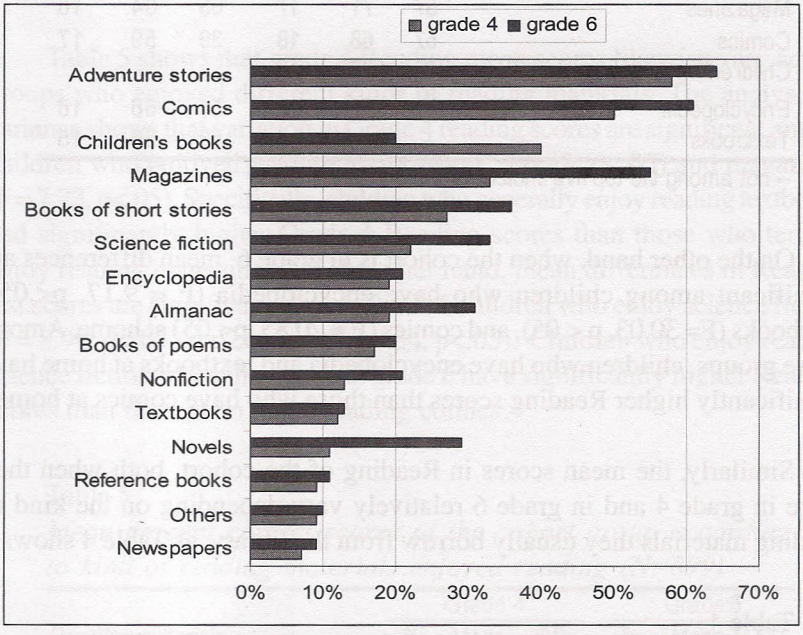


Figure 3. Comparison of reading materials enjoyed reading in grade 4 and in grade 6

Based on the responses of the cohort, the five most selected reading materials at home, borrowed from the library, and enjoyed reading were examined to assess their influence on the overall Reading scores.

As Table 3 shows, the Reading mean scores of the cohort relatively vary depending on the kind of reading materials they have at home when they were in grade 4 and in grade 6. However, the analysis of variance shows that Grade 4 Reading mean scores only differs significantly among children who have encyclopedia ($F = 38.35$, $p<.05$) and children's books ($F = 18.94$, $p<.05$) at home. That is, children who have these reading materials at home have significantly higher Reading scores than those who have other reading materials.

Table 3

Mean percent correct scores of the cohort group according to kind of reading materials at home (N = 899)

	Grade 4			Grade 6		
Reading materials	%	Mean	SD	%	Mean	SD
Adventure stories	62	70	18	47	63	17
Magazines	57	71	17	65	64	16
Comics	57	68	18	39	59	17
Children's books	53	72	17		--	
Encyclopedia	47	72	16	51	66	16
Textbooks		--		50	66	15

-- not among the top five choices of the cohort

On the other hand, when the cohort is in grade 6, mean differences are significant among children who have encyclopedia ($F = 9.17$, $p<.05$), textbooks ($F= 30.03$, $p <.05$), and comics ($F = 20.85$, $p<.05$) at home. Among these groups, children who have encyclopedia and textbooks at home have significantly higher Reading scores than those who have comics at home.

Similarly, the mean scores in Reading of the cohort, both when they were in grade 4 and in grade 6 relatively vary depending on the kind of reading materials they usually borrow from the library, as Table 4 shows.

Table 4

Mean percent correct scores of the cohort group according to reading materials borrowed from library (N = 899)

	Grade 4			Grade 6		
Reading materials	%	Mean	SD	%	Mean	SD
Adventure stories	56	69	17	71	64	17
Children's books	37	70	17	33	65	16
Books of short stories	33	71	16	54	65	16
Science fiction	22	73	17	40	67	16
Comics	18	63	18		--	
Nonfiction		--		35	69	15

-- not among the top five choices of the cohort

However, the analysis of variance shows that grade 4 Reading mean scores only differs significantly among children who borrow science fiction ($F = 7.20$, $p<.05$) and comics ($F = 5.60$, $p<.05$). Children who borrow

science fiction have significantly higher Reading scores than those who usually borrow comics from the library. On the other hand, when the cohort is in grade 6, mean differences are significant among children who borrow science fiction ($F = 9.83$, $p<.05$) and nonfiction ($F = 28.67$, $p<.05$). Children who generally borrow nonfiction have significantly higher Reading scores than those who borrow science fiction.

Table 5 shows that grade 4 Reading mean scores likewise vary across groups who enjoyed different kinds of reading materials. The analysis of variance shows that variation in Grade 4 reading scores are significant among children who enjoyed reading textbooks ($F = 6.98$, $p<.05$) and magazines ($F = 7.23$, $p<.05$). Specifically, children who generally enjoy reading textbooks had significantly higher Grade 4 Reading scores than those who tend to enjoy reading magazines. On the other hand, mean differences in Reading test scores are significant among Grade 6 children who enjoy science fiction ($F = 8.99$, $p<.05$) and comics ($F = 8.54$, $p<.05$). Children who enjoy reading science fiction when they were in grade 6 have significantly higher Reading scores than those who enjoy reading comics.

Table 5

Mean percent correct scores of the cohort group according to kind of reading materials enjoyed reading (N=899)

	Grade 4			Grade 6		
Reading materials	%	Mean	SD	%	Mean	SD
Adventure stories	58	70	18	64	64	17
Comics	50	68	18	61	62	17
Magazines	33	69	18	55	64	16
Encyclopedia	19	72	18		--	
Textbooks	12	73	20		--	
Books of short stories		--		36	64	16
Science fiction		--		33	67	16

-- not among the top five choices of the cohort

The findings above show that reading textbooks and encyclopedia had significantly yielded high Reading scores. This is relatively expected since these materials are generally read for information. However, it shows that reading comics had consistently yielded lower Reading scores as compared to other reading materials read for literary experience such as science fiction.

Kinds of reading strategies used. As Table 6 shows, when a word is not understood, pupils in the cohort sample use two reading strategies when they were both in grade 4 and in grade 6. The first is the common practice of looking in the dictionary for the meaning of the word while the second is asking help from others. Looking at how the kind of reading strategy used influence reading scores, the analysis of variance shows that children who tend to ask for help when they do not understand a word have higher Reading scores in grade 4 ($F = 12.51$, $p<.05$). When the cohort is in grade 6, using context clues to understand a word yielded significantly higher Reading scores ($F = 26.54$, $p<.05$).

Table 6

Mean percent correct scores of the cohort group according to reading strategy used when a word is difficult to understand (N = 899)

	Grade 4			Grade 6		
Reading strategy	%	Mean	SD	%	Mean	SD
Use a dictionary	52	70	17	61	64	16
Ask for help	47	70	17	57	65	17
Use context clues	20	71	19	38	72	15
Say the word	16	66	20	17	66	16
Skip over the word	8	67	18	17	66	15

On the other hand, when a sentence, a paragraph, or page is difficult to understand, children still maintain a positive attitude towards reading. Instead of putting away the reading material (which has the least percentage), the cohort sample tend to ask someone to explain or re-read the part carefully as their reading strategies.

Table 7

Mean percent correct scores of the cohort group according to reading strategy used (N = 899)

Reading strategy when a sentence,	Grade 4			Grade 6		
paragraph, or page is difficult to understand	%	Mean	SD	%	Mean	SD
Ask someone to explain	50	70	17	52	65	15
Reread the part carefully	47	73	17	66	68	16
Pause and try to think of the meaning	30	75	16	40	69	15
Continue reading	24	66	19	24	63	16
Look at the title, pictures, or headings	17	73	17	23	68	14
Put important ideas into your own words	12	69	19	10	69	17
Put away the reading material	2	59	21	2	60	21

As Table 7 shows, there is a variation in the mean scores depending on the kind of reading strategy used when the cohort is in grade 4. Looking at how these reading strategies influence reading scores, the analysis of variance shows children who pause and try to think of the meaning the word had significantly higher Reading scores ($F = 7.64$, $p<.05$). In grade 6, on the other hand, mean scores do not differ significantly among groups. This might suggests that no specific reading strategy is most effective at this grade level.

Implications of the Findings

Reading has always been a crucial part of the elementary school program. Thus, the goal of this study is to describe, on the average, the reading competency of elementary school pupils when they are in grade 4 and then in grade 6 and explore how their reading habits influence their reading achievement.

Findings of this study show that majority of the pupils are reading above norm-level expectations in grade 4 and in grade 6, the exit points in the Philippine basic education system. However, looking at their pattern of reading habits, the findings suggest that the number of hours they spent on reading daily is not encouraging at all. This poses a challenge for many educators, parents, librarians, and other stakeholders to motivate and inspire students to pick up a book or any reading material in the first place. No matter what instructional methods we employ, students must spend substantial time applying the reading skills and strategies we teach before they develop reading proficiency. To become good readers, students must read and practice reading more.

Furthermore, while it is important to provide an inclusive and wide range of reading materials at home, in the classroom, and in the library, it is even more important to recognize the different kinds of demands that the reading materials make on readers and the different kinds of satisfaction they provide. Thus, reading intervention programs should be designed in such a way that would make reading more meaningful among readers. Among the factors to be considered would include motivation and purpose of reading. As literature suggests, if readers are highly motivated to comprehend a text, they will have the drive to push themselves despite complexity. Teachers should also know how to select and prioritize reading materials to ensure

balance between the purpose of reading and the interest and ability of the readers.

Lastly, it is important to keep in mind that the patterns found in this or any other study cannot be pegged on an individual whose interests are always unique and specific. What this study and other similar researches can do, however, is point to trends that educators, librarians, and child psychologists can generally use to respond to elementary children.

References

Department of Education (2010). Education Reforms on Track to Meet EFA Goals: Press Release from the Office of the Secretary. Retrieved June 23, 2010 from http://www.deped.gov.ph/cpanel/uploads/issuanceImg/jan21-efa.pdf

Literacy Coordinating Council (2003). 2003 Functional literacy, education and mass media survey, Retrieved April 29, 2010 from http://lcc.deped.gov.ph/lcc index.php? option com_content&view category&layout blog&id=38&Itemid=67

Test Development Section (2007). *Reading Test Levels 1-3 Score Interpretation Guide*. Philippines: Center for Educational Measurement, Inc

---

Ms. Analyn Capalihan holds the position of a Test Specialist in the Aptitude and Special Tests Unit of the CEM Test Development Section. She has been with CEM for more than five years. She is currently taking up her Masters of Education with specialization in Educational Measurement and Evaluation at the Philippine Normal University.

Ms. Janet Evasco works in the Aptitude and Special Tests Unit of the CEM Test Development Section as a Test Specialist. She has been with CEM for almost five years. She holds a Master of Arts degree in Psychology which she acquired from the University of the Philippines–Diliman.

PUBLICATION GUIDELINES

The Center for Educational Measurement, Inc, (CEM) has always recognized the efforts undertaken, individually or collectively, by the teachers, principals or guidance counselors, particularly in the areas of educational measurement, assessment, and evaluation.

It is the intent of CEM to support these efforts by inviting schools to submit articles or research papers to our official publications, the *Philippine Journal of Educational Measurement* (PJEM) and *The CEM Standard.* However, to facilitate publication of papers in printed form, we request the authors to adhere to the guidelines detailed below.

A. The Philippine Journal of Educational Measurement

Description

The *Philippine Journal of Educational Measurement* (PJEM), a refereed journal, is published annually by the Center for Educational Measurement, Inc. The journal aims to contribute to a better understanding of the system of measurement in the field of education across all levels basic to higher education in the Philippines. As such, the journal contains empirical and nonempirical reports such as theoretical studies, research studies, evaluation studies, specialized reviews, essays, reflective inquiry, critical book reviews, commentaries related to educational testing, measurement, assessment, evaluation, and research on substantive, innovative, and methodological issues.

Article Content

The PJEM welcomes contributions from teachers, researchers, measurement theorists, school administrators, policy-makers, and other key stakeholders across all levels. All articles are accepted on the basis that they are original materials, have not been previously published, and are not currently under consideration for publication elsewhere. Articles may fall into the following categories:

1. *Theoretical studies*. These are studies concerned with the development of theory in the analysis of measurement in order to enhance the understanding of measurement processes.

2. *Evaluation Studies*. These are studies that assess the extent of implementation and impact of a specific program or project and usually emphasize needs assessment and/or ongoing feedback to program implementers.

3. *Research Studies*. These are studies that typically use data derived from qualitative or quantitative methods or both, including but not limited to, experiments, case studies, surveys, philosophical investigations, and historical studies, to yield new information, to focus on specific projects or settings, or to synthesize emerging patterns.

4. *Specialized reviews*. These are articles aimed at critically relating issues, comparisons, and analyses to the application of educational measurement and models in the educational process, but well founded in the existing literature. Reviews focused on research, theory, methodology, theme, theoretical contributions, critiques, and instructional techniques are accepted.

5. *Book reviews*. These are articles critically evaluating the strengths and weaknesses of a book relevant to the scope of PJEM. Books reviewed must be within two (2) years of its publication date.

6. *Commentaries*. These are articles that critically analyze any of the following: (1) policy trends related to educational measurement, (2) relationship between research and evaluation, and (3) connections between research, policy, and practice.

7. *Critiques of articles*. Constructive comments on articles previously published in PJEM are accepted. It is encouraged that these should stimulate discussions and present ideas or alternatives in print. Authors will be invited to respond to the critique made on their article before publication. When possible, the critique and the response will be published at the same time.

Preparation of Manuscripts

Authors are encouraged to prepare manuscripts following the Publication Manual of the American Psychological Association (APA) 5th edition. Manuscripts, including the Abstract and References, should be typed double-spaced on clean short 8½ by 11 inch white bond papers with a margin of one inch on all four sides, using 12-point Times New Roman font and justified to the left margins only. Page numbers should be placed at the center-bottom of the page. Notes, if applicable, are grouped in one section at the end of the article.

If the manuscript is based on a thesis or dissertation, a funded research project, or a paper presented at a conference (whether local or international), a footnote on the cover sheet should provide the relevant facts, including the thesis or dissertation adviser or the organization sponsoring the project or conference.

1. *Author Identification.* Since the review process is blind, the first page should indicate the title of the article, full name(s) of author(s), particulars of present position(s)/institutional affiliation(s), e-mail address(es), phone and fax number(s), and any information of special relevance. Identity of the author(s) should appear only on this page.

2. *Abstracts and Keywords.* All manuscripts must include abstracts of 120 words or less. An abstract is a complete, condensed summary of the article and not a description nor an introduction to the article. In addition, supply 4-5 keywords/phrases that characterize the content of the paper, which can be used for indexing purposes.

3. *Length of Articles.* Manuscripts for theoretical studies, evaluation studies, research studies, and specialized reviews should not exceed 50 pages. Book reviews, commentaries, and critiques of articles should not exceed five pages.

4. *Language.* Manuscripts written in English and Filipino are accepted. However, an English title should be submitted if the manuscript is written in Filipino.

5. *Tables.* Each table should be presented on separate pages and not in the body of the text. It should include a caption and presented in the order in which they appear in the text. As such, they should be given sequential Arabic numbers (i.e., Table 1, etc.), and should be in Microsoft Word (.doc) format. When preparing tables in Microsoft Word, be sure to use the table feature of the program. Furthermore, equations should be generated directly in the text file using the equation feature of Microsoft Word. Importing equations into the text file from a different word processing or graphic applications is discouraged.

6. *Figures.* Illustrations, such as diagrams, drawings, graphs, maps, and photographs, are considered as figures and should be designated as 'Figure 1,' 'Figure 2,' etc., and in sequential Arabic numbers. The text document should not contain the figures but rather are submitted on separate pages at the end of the manuscript and should

be in Tagged Image File Format (TIFF) at a minimum of 600 dpi resolutions. Figure captions should be typed on a separate page, not on the figures, and presented in the order in which they appear in the text.

Since authors know best what they want to show in a figure, they should crop their own figures, leaving only essential materials. If necessary, figures can be rotated 90 degrees and printed sideways. Photocopies of figures are not acceptable.

7. *References*. References cited in the text document must appear in the reference list at the end of the article. Authors must check that reference details are correct, complete, and in accord with the citations in the text. Authors are responsible for all information in their reference list. The Publication Manual of the American Psychological Association (APA) should be consulted in preparing the reference list. In particular, use of its Fifth edition is highly preferred.

Submission of Manuscripts

Interested contributors must submit the following: (1) one clear printed version of the manuscript, and (2) an electronic copy which may be sent as e-mail attachment to cemresearch@cem-inc.org.ph or on a CD-RW containing the appropriate files.

The electronic copy should follow the style guidelines indicated above (with appropriately placed notes on where to insert the tables and figures in the text). In particular, the following sections of the manuscripts should be submitted as separate files: (1) the title of the article and the abstract, (2) the main text, (3) tables, (4) figures, and (5) references. Appendices are not encouraged but may be allowed if considered necessary to facilitate understanding of the manuscript content. Moreover, authors should submit their figures and tables as camera-ready copies, using a laser-printed computer output, at the maximum dimensions given below.

	Figure Dimension (in cm)	
Paper Orientation	Width	Height
Portrait	6.75	9.00
Landscape	9.00	6.75

The minimum acceptable height for a letter or number in a camera-ready figure is about 3 mm or about 9 points. However, it is always better to use larger characters since the Editor may decide to reduce a figure.

Author(s) will be notified in case there is a problem with the electronic files. The printed version of the manuscript, CD-RW copy, and a cover letter should be addressed to:

Corporate Communication Section
Center for Educational Measurement, Inc.
24th Floor, Cityland Pasong Tamo Tower
2210 Chino Roces Avenue, Makati City

The Editor will acknowledge receipt of all contributions. The review process is greatly facilitated when manuscripts are submitted in the proper form.

Editorial Procedures

All manuscripts undergo preliminary screening by the Editorial Staff. They determine whether or not to reject the submission outright. If manuscript fails to meet the Journal's technical and stylistic requirements, it is returned to the author for revision before forwarding to at least two Editorial Consultants for their review. Since the review process is blind, articles sent for review are anonymous. Review is normally completed within three months of the submission, and a comment sheet is provided to facilitate the review. Reviewed manuscripts are generally returned to the authors with specific comments from the Editorial Consultants. Authors may be advised to resubmit their manuscripts to include editorial changes.

The Editor reserves the right to make editorial changes of nonsubstantive nature. Once the final version of the paper has been accepted, authors cannot make further changes to the text. Proofs will be sent to the author(s) if there is sufficient time to do so. These should be corrected and returned to the Editor within seven (7) days to facilitate printing.

Copyright and Permissions

Obtaining necessary permissions for the use of published materials (e.g., figures, tables, etc.) is the responsibility of the authors. These should be properly cited in the list of references.

Papers accepted become the copyright of the Journal, unless otherwise specifically agreed, and may not be reproduced without the permission of the Center for Educational Measurement, Inc.

Copies and Honoraria

Authors whose manuscripts are chosen for publication receive a modest honorarium after the article is printed. Additionally, authors receive two complimentary copies of the PJEM issue in which their article appears.

B. The CEM Standard

Description

The CEM Standard is a newsletter that aims to provide a forum for documentation of experiences and best practices at the school and classroom levels, which have direct application for teachers, guidance counselors, and school administrators. The newsletter also addresses issues, questions, and concerns about educational testing and assessment.

Article Content

Articles submitted in *The CEM Standard* may run from 15 to 50 pages and may fall into the following categories:

1. *Reflective Documentation.* These are narratives or documentation of experiences and/or best practices in the school setting, which reflect the assumptions, importance, and relationship of educational measurement in the teaching-learning process or in school administration.
2. *Essays.* These are short articles that discuss new knowledge or provide insight in the field of educational measurement.

Submission of Manuscripts and Editorial Procedures

The preparation, submission, and editorial procedures followed for processing and reviewing a manuscript submitt for consideration to *The CEM Standard* is essentially the same as for the PJEM. Thus, the review process may also take up to three (3) months.

Philippine Journal of Educational Measurement

OCTOBER 2010 • **VOLUME 11**

The Philippine Journal of Educational Measurement is published annually by the Center for Educational Measurement, Inc., and contains empirical researches and non-empirical reports such as theoretical notes, research studies, evaluation studies, specialized reviews, essays, reflective inquiry, critical book reviews, commentaries related to educational testing, measurement, assessment, and evaluation, and research on substantive, innovative, and methodological issues.

Advisory Board
Lenore Ll. Decenteceo
Maria Cynthia A. Alcantara

Joy J. Macalintal
Kristhine M. Marasigan
Frances Faye M.B. Cardona
Editorial Assistants

Center for Educational Measurement, Inc.
24th Floor Cityland Pasong Tamo Tower
2210 Chino Roces Avenue, Makati City
Tel. Nos: 813-3691, 813-3694 to 95
Fax Nos: 894-1392, 894-5536
www.cem-inc.org.ph