Using Bloom's Revised Taxonomy
to Analyze a Reading Comprehension Instrument

Lorraine Dagostino[1], James Carifio[2], and Qing Zhao[4]
University of Massachusetts at Lowell

Jennifer D. C. Bauer[3]
Middlesex Community College

Nor Hashimah Hashim[5]
Universiti Sains Malaysia

Though the designers may claim otherwise, the vast majority of reading comprehension instruments measure a student's mastery of specific skills, such as decoding or phonological awareness, rather than their ability to extract meaning from text. This article is the third in a series in which the researchers examine two specific Malaysian reading comprehension instruments developed by a team of researchers at the Universiti of Sains Malaysia. These tests were developed for the purpose of evaluating reading comprehension abilities of students in the primary grades (Test I for grade 1-3, Test II for grades 4-6) in Malaysia (Hashim, 2006). In the previous studies, we established that the English version of the test was comparable to the Malay version, and that Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001), could be used to classify the test items. In this study, we extend our prior work by comparing the Malay classification scheme with Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001). They found that rarely did questions classified by the Malay structure as testing higher-order reading skills, such as evaluation, actually do so. Instead, most questions merely identified students' reading skills, rather than comprehension. Thus, more work is needed in order to develop an instrument that actually measures the evaluative abilities of students' as they interact with texts.

Keywords: reading comprehension; cultural background; cognitive process; Bloom's Revised Taxomony

For decades researchers in the field of education (and beyond) have attempted to define the term "reading comprehension," as well as isolate the skills needed to extract meaning from text. In conjunction with this research, models of reading comprehension were (and continue to be) developed, and subsequently tests were designed to assess the reading comprehension skills of students of all ages. However, do these tests actually measure what they purport and reflect models of reading comprehension? With the amorphous definitions of reading comprehension found within reading research literature, is it possible to develop a test that accurately measures a student's ability to extract meaning from text?

Previous Work

Prior to this study, we completed two studies using the same series of reading comprehension tests

developed by a team of researchers at the Universiti of Sains Malaysia. The reason we selected the Malay tests is because these tests were constructed using the Dagostino and Carifio model of reading comprehension (2004), which was developed for the English language. In the first study, we evaluated the comparability tests with their English translations on the basis of skills and levels according to a conceptual framework of reading comprehension developed by Dagostino and Carifio (2004; Dagostino, Carifio, Bauer, & Zhao, 2013). The results of that study showed strong correlations across the Malay and English versions of the test on the classification of reading skills and levels of reading comprehension.

The second study examined the English version of the test to see if there were any correlations of Bloom's classifications of cognitive dimensions of thought with the classification by reading skill and level of comprehension as determined in our previous work. To accomplish this task, three raters classified the test items using Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001). The results demonstrated high-levels of rater agreements among the classification for each test item, indicating that Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001) could appropriately be applied as a classification scheme for the tests. These findings were reported at the 2013 Annual Eastern Educational Research Association (EERA) Conference.

Present Work

After analyzing the results of the first two studies, we felt that more work needed to be done. The primary purpose of the present work was to draw upon the results of the previous studies, and compare and analyze the relationship between the Malay tests classification system and Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001). With this purpose in mind, we set out to explore the following research question:

What relationship, if any, exists between the original Malay Classification System and Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001)?

Organization of the Paper

With this research question in mind, we will begin with a description of the Malay tests, followed by an overview of their development and content. Next we will describe Bloom's Revised Taxonomy of the Cognitive Dimension (Anderson & Karthwohl, 2001) and explain its application to the present study. Next, we will outline this study, including the parameters and limitations, methodology, procedures, results and subsequent data analysis. Finally, we will discuss our findings, their implications, and possible future research.

The next two sections of the paper, describing the Malay Tests and Bloom's Revised Taxonomy: The

Cognitive Dimension, were originally published as part of the author's previous studies (see Dagostino et al., 2013).

The Malay Tests

The Description and Construction of the Malay Tests

The original two Malay tests, constructed by a team of researchers at the Universiti of Sains Malaysia, were developed for the purpose of evaluating reading comprehension abilities of students in the primary grades (Test I for grade 1-3, Test II for grades 4-6) in Malaysia (Hashim, 2006). The following section of this article describes the process for the development and the content of these tests.

Steps for Design and Content of the Malay Instruments

Using the Dagostino-Carifio model of reading comprehension (1994) as a theoretical basis, the development of the test focused on three components: a) defining and selecting the category of the comprehension level as well as of the comprehension skill, b) selection and development of the reading texts, and c) the development of the test questions and the answers. The two tests were designed by conducting a preliminary

Table 1

Specifications for Malay Reading Comprehension Tests with this general Template being the same for Test I and Test II

Reading Comprehension Category	Reading Skills	Reading Skills
Literal (L)	L1A, L1B, L1C	identifying meaning of word/ phrase/ sentence
	L2	identifying main idea
	L3	identifying important point
	L4	making comparison
	L5	identifying cause-effect
	L6	identifying sequence of ideas/events
Inferential (F)	F1	interpreting main idea
	F2	interpreting important point
	F3	interpreting comparison
	F4	interpreting cause-effect
	F5	making a conclusion
Critical Creative (K)	K1	evaluating
	K2	making a conclusion
	K3	internalizing
	K4	identifying the moral of the story/lesson

survey that included a discussion with Malay teachers, a review of teaching learning materials and observations of teachers teaching in a classroom. Once the survey was completed, a first draft was developed for Test I and for Test II. The writing of the first draft was accomplished through a series of workshops with Malay language teachers, experts from Curriculum Development Center, administrators from the District Education Office and State Education Department, lecturers of School of Educational Studies from the Universiti Sains Malaysia (Hashim, 2006). As a result of this work, the researchers established the following Table of Specifications (Table 1), which outlines the relationships between the reading comprehension levels and reading skills underlying the construct of both tests.

Defining the Reading Comprehension Levels and the Reading Comprehension Skills

The reading comprehension levels and the reading skills determine the difficulty and the nature of the reading texts and the test items. The Malaysian tests have three comprehension levels defined as follows (Hashim, 2006):

    a) *Literal (message extraction) Reading Comprehension, which refers to the memorization of facts in texts where information is explicitly stated at a basic level of thinking;*

    b) *Inferential (message interpretation) Reading Comprehension which refers to the ability of students to interpret meaning where they need to use overt information along with intuition, reasoning, and experience to attain the higher level of thinking assessed by the Malay tests; and,*

    c) *Critical/Creative (message evaluation) Reading Comprehension, which refers to the student's ability to do an overall critical evaluation of certain information or an idea that has been read in terms of the precision and/or suitability of the given information of a new idea, encountered. This critical evaluation may require some divergent thinking and depend to some degree upon the knowledge and personal experience of the student, but it focuses mostly on convergent critical thinking being done by the student.*

    Reading comprehension skills. There are ten reading comprehension skills that are assessed by the Malay tests (Hashim, 2006): (a) identifying meaning of word/phrase/sentence; (b) identifying the main idea; (c) identifying the important point; (d) identifying the cause-effect relationship; (e) identifying the sequence of ideas/events; (f) making a comparison; (g) drawing a conclusion; (h) evaluating; (i) internalizing; (k) identifying the moral of the story/lesson. These ten skills range from simple reading comprehension to what is called deep or deeper understanding, which is a first step towards what is called evaluative reading. These skills are

the ones that usually constitute the classification of items assessed in most reading tests.

Types and Contents of Reading Texts

There are several types of text that make the text broad in scope and representative of various types of reading of non-technical materials that are encountered in daily reading situations (Hashim, 2006). There are essays, fiction, reports, letters, poems, biographies, speeches, dialogues, and news reports. There are 12 texts for Test I and 12 Texts for Test II. There are various subjects (literature, history, etc.) The individual texts are 100 words or less for Test I and 100 words or more for Test II. The passages in the test for grades (1-3) are simpler in structure as well as expectations for level of reading comprehension than those used for grades 4-6. A research group, three expert teachers, teacher trainers, psychometric and experts from the university developed the texts, with ideas for the texts coming from books and magazines.

Development of the Test Items

The question and answer formats for the tests took various forms such as a) sentences from text that needed completion with a choice of answers, b) items that needed a choice of answers in multiple choice form, and c) instructions and blanks to be filled in with multiple choice form. An item specification table was developed to categorize the types of items in the test (Hashim, 2006). Each test consists of 50 multiple choice items designed to evaluate reading comprehension with consideration given to reading skill ability and reading comprehension level. Some specific things were considered in the item development. They are as follows: a) arrangement of each item was based upon reading comprehension skill (forms, style, pupils' existing knowledge), and b) implicit information and inferential definition. In the case of implicit information, the text considers information in the text and students' background. In the case of inferential definition the test considers an integrated synthesis of literal with existing knowledge, intuition and reader's imagination.

The following Table of Specifications include the classification by reading comprehension level and reading skill for each test item. Both Malay tests were built from the same general Table of Specifications, but classification by reading skill varied for each test (Table 2 and 3; see Appendix A).

Design and Choice of Answers and Distracters

A multiple-choice format was used because it was considered as most objective. Each answer had four options (A, B, C, D for each item with each option coded A=1, B=2, C=3 and D= 4). The correct answer was scored 1, and the wrong answer was scored 0. The design of the answers and distracters required a) the suitability of choice of answers relative to the cognitive task that was related to the content and the texts, and b) syntax and semantic forms needed to be different from the texts so

that students could be assessed on how well they understood the context of the meaning (NorHashim, 2006).

Reliability Measures of the Two Malay Tests

The Malay researchers examined three types of internal consistency reliability estimators for both tests with the results being almost identical for both tests. The first internal consistency (of test-taker overall performance) reliability estimator computed was the Cronbach alpha coefficient, which was $r = +.66$ ($N = 2763$) for Test I and $r = +.61$ ($N = 4101$) for Test II. As is well known, test length, sample size, and test content and item type *heterogeneity* affect and limit the size of the Cronbach alpha one will observe in any given context. As test content and the cognitive levels and operations assessed are so heterogeneous for both tests, the Cronbach alphas observed for each test are quite good to excellent given that test lengths (50 items each) and sample sizes ($N = 2763$ and $4101+$) and are in the range that one would expect given the qualitative characteristics of both tests.

The second internal consistency reliability estimator the Malay researchers computed was the Guttman reliability coefficient, which assess the degree to which students' performances on the test are hierarchical in character (i.e., students who do well on low level items are not doing well on high level items and vice versa), which performances should be for Test I and Test II given how they were constructed and their qualitative characteristics. The Guttman reliability coefficient for Test I was $r = +.77$ ($N = 2763$) and for Test II was $r = +.72$ ($N = 4101$), which are excellent to outstanding and indicate that this particular qualitative characteristic of both tests are as hypothesized and purported.

The third internal consistency reliability estimator the Malay researchers computed was the Kuder-Richardson odd-even items reliability coefficient, which assesses the degree to which items types and their characteristics are *evenly balanced* across the test, as well as students' performances on the items on the test. For example, the Kuder-Richardson reliability coefficient would be low if all of the odd items were easy (or recall) items and all of the even items were difficult (or skill) item, or if all of the poorly constructed and non-functioning items were easy items as opposed as opposed to this characteristic being evenly balanced across both the odd and even items. The Kuder-Richardson odd-even items reliability coefficient for Test I was $r = +.77$ ($N = 2763$) and for Test II was $r = +.73$ ($N = 4101$), which are good to excellent and indicate that the various types of items and their various characteristics were *evenly balanced* across each test as were student performances.

As one-administration internal estimates of various types of consistencies in student performances across each of these two tests and thus internal consistency reliabilities estimates, the results obtained by the Malay researchers of the three different indicators of internal reliabilities estimates were excellent. High one-administration internal consistency estimates of reliabilities, however, are no guarantee that test-retest reliabilities will be equally high as they could actually be lower or higher which is why the Malay researchers are currently collecting the data to generate the test-retest reliability coefficients as these coefficients are key in the assessment of change across time on these measures. But to date, the reliabilities estimates for each test that are available are excellent and particularly so given the internal complexity of each test, and each is also initially supportive empirically of specific aspects of the construct validity of each test, although not as direct or strong evidence as other analyses might indicate.

Bloom's Revised Taxonomy:
The Cognitive Dimension (Anderson & Karthwohl, 2001) and Its Application to the Present Study

Bloom's original taxonomy was designed to help teachers establish objectives for instruction, learning and assessment. This revised taxonomy has served to guide the design and the implementation of accountability programs and standards-based curriculum. The revision of the original taxonomy that was used in the present study has been refined to incorporate new knowledge into the original framework. This revised taxonomy gave us a good conceptual framework for determining the cognitive levels and ability reflected in test items on the reading comprehension test that we expect to use as an assessment instrument in subsequent studies. The test already has been examined for general levels of reading comprehension and reading skills. What we hoped to accomplish in the present study was to compare the classification system developed by the Malay test writers with a table of specifications developed using Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001). This taxonomy was chosen from other ways to evaluate cognitive abilities because it is most applicable, familiar and understandable to the classroom teacher, yet detailed enough to give valuable insight into cognitive processes that are considered necessary to learning and to the assessment of success in instruction and learning (Anderson & Karthwohl, 2001). Further work is planned to compare Bloom's Revised Taxonomy with other classification frameworks for measuring cognitive abilities as they may manifest themselves in tests of reading comprehension.

Using Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001) gave us a standard, well-recognized classification system for our immediate goals, and it also indirectly should be useful for guiding instruction and curriculum guidelines that may be generated by our present work. This consistency across these tasks should simplify the work of the classroom teacher and the researcher. In sum, Bloom's Revised Taxonomy gives us definitions for classifying the learning, teaching and assessing of the cognitive dimension of thought that is

central to instruction in most subject areas, and in relationship to our work in reading comprehension as an aspect of assessment of literacy in a way that differs from most current measures of reading comprehension directly.

It should be noted that although our classification process focuses on the categories of the cognitive process dimension of the taxonomy, it also takes into consideration two categories of the Knowledge Dimension, that of Factual and Conceptual Knowledge, but not Procedural and Metacognitive Knowledge. However, categorizing test items using the two categories of Knowledge (Factual and Conceptual Knowledge) was not part of the present study.

In sum, Bloom's Revised Taxonomy gives us objectives for classifying the learning, teaching and assessing of the cognitive dimension of thought that is central to instruction in most subject areas, and in relationship to our work in reading comprehension as an aspect of assessment of literacy in a way that differs from most current measures of reading comprehension. In Table 4 (Appendix B) is a Chart of the Bloom's Revised Taxonomy and descriptions of each categories as found in *A Taxonomy for Learning, Teaching and Assessment (*Anderson & Karthwohl, 2001).

Methodology

Procedures

The Malay version of the test was translated into English, and in a previous study, three expert raters completed the inter-rater judgments of the answers, levels and skills of each test item. In a subsequent study, the raters applied Bloom's Taxonomy as a method of categorization for each test item.

The raters had either a Ph.D. in language arts and literacy, or were completing work for that degree. One of the raters spoke both English and Chinese, and another works with young children from several cultures and language backgrounds. Previous ratings by these same raters had judged the items for skills and levels as described earlier in this paper with good results (Dagostino et al., 2013).

The raters first made a comparison between the Malay Table of Specifications and Bloom's Revised Taxonomy and then derived a 'cross-walk" from the first Malay classification system to the other (Bloom's RT). Using the definition of levels and skills as outlined in the Malay Table of Specifications, the raters assigned each individual test item to a level of Bloom's Revised Taxonomy using the definitions of the levels, which essentially estimated the overlap, intersection or "correlation" of the two classification systems as a results of the 50 items so classified for each Malay test. The resulting "cross-walked" joint Table of Specifications is presented in Tables 5 and 6 (see Appendix C).

Next, the raters evaluated each test item from both Malay tests and classified them using Bloom's Revised Taxonomy of Cognitive Abilities. The three expert raters completed their individual judgments by first reading each item of Test I and Test II independently, and then determining the Bloom's Revised Taxonomy level of cognitive ability they felt best applied to the dimension of reading comprehension being tested. The categories for classification are as follows: 1) Remember, 2) Understand, 3) Apply, 4) Analyze, 5) Evaluate, 6) Create. (As a reminder of each category, see Table 4.)

After independent readings and ratings of the test items using the Revised Taxonomy were completed, the raters compared their judgments for all of their ratings. There was not a need for a reconciliation process among the raters based upon this discussion because of the high level of agreement among the three raters. After the quantitative analysis of the ratings was completed, the raters met again to discuss the results to evaluate the meaning of the raters' agreements on the item ratings.

Results and Data Analysis

The analysis and results section of this paper presents the data and its interpretation for the following research question:

What relationship, if any, exists between the original Malay Classification System and Bloom's Revised Taxonomy: the Cognitive Dimension (Anderson & Karthwohl, 2001)?

To analyze the data, we compared the Cross-Walk Tables of Specifications of Test Items by Malay Levels/Skills and Bloom's Revised Taxonomy (Tables 5 and 6), which predicted the Bloom's Revised Taxonomy classification of each test item based on the Malay Table of Specifications, with the actual Bloom's Revised Taxonomy classifications (Anderson & Karthwohl, 2001) the raters assigned to each test item. Table 7 (see Appendix D) represents the comparison between these expected results and the actual results for the evaluation of the test items.

As can be seen from Table 7, there was a major discrepancy between the predicted results and the actual results. According to the original Malay Table of Specifications, there are three levels of reading comprehension for each test, literal, inferential, and critical/creative. These levels correspond to Bloom's Revised Taxonomy cognitive processes of remember, understand, and evaluate (Anderson & Karthwohl, 2001). However, when we sorted each individual test item for both test, they found that only 40% of the items for Test 1, and 38% of the items for Test 2 matched the predicted results found in the Cross-Walk Tables of Specifications of Test Items by Malay Levels/Skills and Bloom's Revised Taxonomy (Tables 5 and 6). Additionally, 22% of the items on Test 1, and 16% of the items on Test 2, were sorted into categories of Bloom's Revised Taxonomy that were not on the Cross-Walk Tables of Specifications.

Furthermore, though the Malay Table of Specifications indicates a mix of literal, inferential, and

critical/creative test items for each test, when we applied the Bloom's Revised Taxonomy framework (Anderson & Karthwohl, 2001) to the test items, we found that the majority of the test items fell into the category of understanding (62% for Test 1, 58% for Test 2 – see the full results in Figure 1).
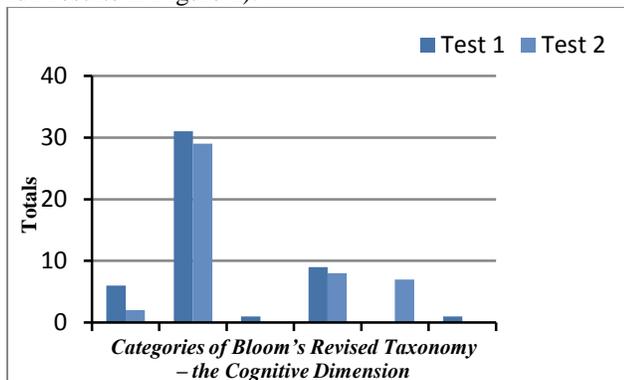

Figure 1. Reconciled answers for each Malay test. This figure illustrates the reconciled answers of the three raters when evaluating the individual answers of the two Malay tests using Bloom's Revised Taxonomy: The Cognitive Dimension as a framework.

Findings and Discussion

Based upon the results of the analyses, we have found three categories of concerns and questions that should help us focus our discussion. The first category focuses on the comparison of the Malay classifications and the Bloom's Revised Taxonomy classifications (Anderson & Karthwohl, 2001) by addressing two questions. The two questions are:

a) Why did the predicted Bloom's and the actual Bloom's classifications not match as we hoped that they would?

b) Does the Bloom's classification more accurately describe test items than the Malay classification, or does it simply refine or transform the Malay classification in some way?

In response to these questions, we wish to suggest that the mismatch may be attributed to the fact that the predicted classifications are drawn from the Malay classification system, whereas the actual classifications were based upon an evaluation of the individual test items using the Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001). In this case then, the Malay classification of the test items may reflect something quite different from the cognitive processes that we believe are at the heart of reading comprehension and understanding as they are identified Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001), or that the Taxonomy intends to do. This discrepancy in the two classifications suggests that each system may actually represent a different view of reading comprehension and point to strengths and weaknesses in these two views. We

propose that Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001) be considered as a construct for transforming the Malay system or levels, as the Bloom's Revised Taxonomy classifications challenge the Malay classification system as a way to measure reading comprehension. We also believe that the Bloom's classifications (Anderson & Karthwohl, 2001) reflects the Dagostino-Carifio model (1994) model of reading comprehension, as described in their original work, better than the Malay system does, raising questions about how the Dagostino-Carifio model (1994) was actually applied in the original Malay work. What we believe is that this discrepancy reflects the idea that what the Malay classifications are doing is reflecting readers' behaviors rather than the readers' understanding of the text, where understanding means getting meaning and showing cognitive level of thought. This notion of reading comprehension is part of the substance of the Dagostino-Carifio model (1994), and will direct our future work.

This thinking in addressing the first concern and questions leads us to the second category of concern, which is the question:

Is the skills model, as reflected in the Malay classifications, adequate or too limited for the needs of the reader in today's world?

That is, is looking at behaviors, as has been done for years in the assessment of reading comprehension, rather than at meaning and cognition, misleading us about the readers' optimal performance in comprehending a text? We believe that it is, and that changes need to be made in the model used for assessing reading comprehension so that it reflects the reader's understanding (ability to get meaning and cognitive level) rather than the skills reflected in the reader's behavior.

We believe that our data challenges the perception that assessing skills assesses reading comprehension adequately, and raises the question:

Do we need a different way to test reading comprehension? And what then is the goal in doing so?

If we consider the Dagostino-Carifio model of reading comprehension (1994), we find the answer that the ability to evaluate a text is a reasonable, attainable and necessary goal for the kind of assessment that we seek. Implicit in the Dagostino-Carifio model (1994) is the principle that evaluation runs constantly throughout the reading comprehension process, and that without evaluation, understanding of the text is incomplete and the reading process quite limited. And it is this dimension of reading comprehension that is missing from the present assessment of reading comprehension, particularly in the instruments like the Malay tests, for the primary and intermediate grades. We think that the omission is based upon a limited view of the cognitive abilities of young readers as well as the difficulty of assessing this dimension of reading comprehension with psychometric

certainty, particularly given the formats for measurement that are presently used. Our belief is that when the content of the passages of a reading comprehension assessment instrument is familiar and relevant to the young reader's background that they can do the higher-level thinking required to evaluate a text. We also believe that using a classification of test items like Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001) will give a better profile of how well a reader understands a text and gets meaning rather than simply reflect the skills applied to processing the text. We acknowledge that some assessment instruments may include items that claim to require some critical or evaluative reading, but we do not believe that, in general, most instruments do it adequately. In turn, instruments do not tell us a reader's optimal ability, or inform educators as to how well the reader will meet the demands of today's world.

We believe that the results of the present study point in the direction that we need to go to accomplish this goal; that is, to measure higher-levels of thinking as they apply to reading comprehension and find a good format to do so.

The third and final category of concern has to do with the possible limitation of our work and whether our analyses of the data gives adequate support to the above discussion and final conclusion about the difficulty and present inadequacy of assessing reading comprehension as we do today. However, if we are correct in our analyses and in concluding that reading comprehension can be better assessed by using Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001) for item classification and profiling student reading ability, we will be challenging many of the instruments as well as frameworks such as the Common Core that are presently used to guide instruction and assessment of reading comprehension.

In general, we believe the testing industry needs to shift from a simple behaviorist paradigm to one that reflects the cognitive view as it is reflected in the Dagostino-Carifio model of reading comprehension (1994) and subsequently in the Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001). Our thinking and conclusions do not necessarily mean that the items in this Malay test are bad, but that the test does not go far enough in its assessment of a reader's ability, and that a different classification system may give a better profile of the reader's optimal performance and ability.

Further Work

The next step in our work and our series of studies is to establish a table of specifications that reflects this added dimension of reading comprehension so that such an instrument, or section of an instrument, may be developed. Next comes the development and validation

of items that can be added to the Malay test with their classifications according to Bloom's Revised Taxonomy (Anderson & Karthwohl, 2001). This work then can be used to test children in the primary and intermediate grades with an instrument that aims to reflect the goal of assessing higher-level thinking, such as that found in evaluative reading, as it is reflected in reading comprehension assessment instruments. Doing this will not only extend educators' ability to assess optimal performance, but also test the applicability of the Dagostino-Carifio model of reading comprehension (1994) for primary and intermediate grade readers.

References

Anderson, L. E., & Karthwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching and assessment.* New York: Longman.

Cain, K., & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology, 76*(4), 697–708.

Dagostino, L., & Carifio, J. (1994). *Evaluative reading and literacy: A cognitive view*. Boston: Allyn and Bacon.

Dagostino, L., Carifio, J., Bauer, J., & Zhao, Q. (2013). Cross-cultural reading comprehension assessment in Malay and English as it relates to the Dagostino and Carifio model of reading comprehension. *Current Issues in Education*, *16*(1).

Dagostino, L., Carifio, J., Bauer, J. D., Zhao, Q., & Hashim, N. H. (2014). Assessment of a reading comprehension instrument as it relates to cognitive abilities as defined by Bloom's Revised Taxonomy. *Current Issues in Education*, *17*(1).

Harman, H. H. (1976). *Modern factor analysis.* Chicago: University of Chicago Press.

Hashim, H. N. (2006). *Reading comprehension in Malaysia primary grades 1-6.* Penang, Malaysia: Universiti Sains Malyasia.

McNamara, D. S., & Kendeou, P. (2011). Translating advances in reading comprehension research to educational practice. *International Electronic Journal of Elementary Education, 4*(1), 33-46.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934-947.

Winstead, L. (2004). Increasing academic motivation and cognition in reading, writing, and mathematics: Meaning-making strategies. *Educational Research Quarterly, 28*(2), 30-49.

Appendix A

Tables 2 and 3

Table 2
Malay Table of Specifications Including Test Items by Classification for Test I

Test Item	Malay Classification	
	Level	Skill
1	Literal (L)	Main Idea
2	Literal (L)	Important Point
3	Literal (L)	Cause and Effect
4	Literal (L)	Sequence of Ideas/Events
5	Literal (L)	Main Idea
6	Literal (L)	Important Point
7	Literal (L)	Cause and Effect
8	Literal (L)	Meaning of Word/ phrase/ sentences
9	Literal (L)	Main Idea
10	Literal (L)	Comparison
11	Literal (L)	Cause and Effect
12	Literal (L)	Sequence of Ideas/Events
13	Literal (L)	Meaning of Word/ phrase/ sentences
14	Literal (L)	Comparison
15	Literal (L)	Comparison
16	Literal (L)	Sequence of Ideas/Events
17	Inferential (F)	Important Point
18	Inferential (F)	Comparison
19	Inferential (F)	Cause and Effect
20	Inferential (F)	Conclusion
21	Inferential (F)	Main Idea
22	Inferential (F)	Comparison
23	Inferential (F)	Cause and Effect
24	Inferential (F)	Conclusion
25	Inferential (F)	Main Idea
26	Inferential (F)	Important Point

27	Inferential (F)	Cause and Effect
28	Inferential (F)	Conclusion
29	Inferential (F)	Main Idea
30	Inferential (F)	Comparison
31	Inferential (F)	Comparison
32	Inferential (F)	Cause and Effect
33	Critical Creative (K)	Evaluating
34	Critical Creative (K)	Conclusion
35	Critical Creative (K)	Internalizing
36	Critical Creative (K)	Moral of the Story
37	Critical Creative (K)	Conclusion
38	Critical Creative (K)	Conclusion
39	Critical Creative (K)	Internalizing
40	Critical Creative (K)	Evaluating
41	Literal (L)	Meaning of Word/ phrase/ sentences
42	Literal (L)	Cause and Effect
43	Inferential (F)	Main Idea
44	Inferential (F)	Conclusion
45	Critical Creative (K)	Evaluating
46	Literal (L)	Meaning of Word/ phrase/ sentences
47	Literal (L)	Main Idea
48	Inferential (F)	Important Point
49	Inferential (F)	Comparison
50	Critical Creative (K)	Moral of the Story

Table 3

Malay Table of Specifications Including Test Items by Classification for Test II

Test Item	Malay Classification	
	Level	Skill
1	Literal (L)	Meaning of Word/ phrase/ sentences
2	Literal (L)	Important Point
3	Literal (L)	Cause and Effect
4	Literal (L)	Sequence of Ideas/Events
5	Literal (L)	Meaning of Word/ phrase/ sentences
6	Literal (L)	Comparison
7	Literal (L)	Cause and Effect
8	Literal (L)	Sequence of Ideas/Events
9	Literal (L)	Main Idea
10	Literal (L)	Important Point
11	Literal (L)	Cause and Effect
12	Literal (L)	Sequence of Ideas/Events
13	Literal (L)	Main Idea
14	Literal (L)	Comparison
15	Literal (L)	Comparison
16	Literal (L)	Cause and Effect
17	Inferential (F)	Main Idea
18	Inferential (F)	Important Point
19	Inferential (F)	Cause and Effect
20	Inferential (F)	Conclusion
21	Inferential (F)	Main Idea
22	Inferential (F)	Important Point
23	Inferential (F)	Cause and Effect
24	Inferential (F)	Conclusion
25	Inferential (F)	Main Idea
26	Inferential (F)	Important Point
27	Inferential (F)	Comparison

28	Inferential (F)	Conclusion
29	Inferential (F)	Main Idea
30	Inferential (F)	Important Point
31	Inferential (F)	Comparison
32	Inferential (F)	Cause and Effect
33	Critical Creative (K)	Evaluating
34	Critical Creative (K)	Conclusion
35	Critical Creative (K)	Internalizing
36	Critical Creative (K)	Moral of the Story
37	Critical Creative (K)	Conclusion
38	Critical Creative (K)	Evaluating
39	Critical Creative (K)	Conclusion
40	Critical Creative (K)	Conclusion
41	Literal (L)	Meaning of Word/ phrase/ sentences
42	Literal (L)	Main Idea
43	Inferential (F)	Comparison
44	Inferential (F)	Cause and Effect
45	Critical Creative (K)	Moral of the Story
46	Literal (L)	Meaning of Word/ phrase/ sentences
47	Literal (L)	Important Point
48	Inferential (F)	Comparison
49	Inferential (F)	Conclusion
50	Critical Creative (K)	Evaluating

Appendix B

Table 4

Table 4
Definitions of the Categories of Bloom's Revised Taxonomy – the Cognitive Dimension (Remembering, Understand, Apply, Analyze, Evaluate and Create)

Remembering	*Recognizing* involves retrieving relevant information from long-term memory in order to compare it with presented information. Also identifying
	Recalling involves retrieving relevant information from long-term memory when a prompt is given. The prompt often is a question. Also retrieving.
Understand	*Interpreting* occurs when a student is able to convert information from one representation to another representation. Also translating or paraphrasing.
	Exemplifying occurs when a student gives a specific example or instance of a general concept or principle. Also illustrate.
	Classifying occurs when a student recognizes that something belongs to a certain category. It is a complementary process to exemplifying.
	Summarizing occurs when a student suggest a single statement that represents presented information or abstracts a general theme. Also generalize or abstract.
	Inferring involves finding a pattern within a series of examples or instances. The student abstracts a concept or a principle that accounts for a set of instances. Also extrapolating or concluding.
	Comparing involves detecting similarities and differences between two or more objects, events, ideas or situations. Also contrasting, matching.
	Explaining occurs when a student is able to construct and use a cause-effect model of a system. The model may be derived from a formal theory or may be grounded in research and experience. Also constructing a model.
Apply	*Executing* occurs when a student routinely carries out a procedure when confronted with a familiar task. Also carrying out.
	Implementing occurs when a student selects and uses a procedure to perform an unfamiliar task. It is carried out in conjunction with understand. Also using.
Analyze	*Differentiating* occurs when there is a determination of the relevant or important pieces of a message in relation to the whole structure.
	Organizing occurs relative to the way the pieces of a message are organized into a coherent structure.
	Attributing occurs when the underlying purpose or point of view of the message is related to the entire communication.
Evaluate	*Checking* involves testing for internal consistencies or fallacies in an operation, product, or communication to see whether data support or disconfirm hypothesis or conclusions as well as the accuracy of facts.
	Critiquing involves judging a product, operation or communication against externally imposed criteria and standard.
Create	*Generating* occurs when a problem is represented and alternatives and hypothesis that meet certain criteria are produced.

| | *Planning* occurs when a solution method is devised that meets a problem's criteria for developing a plan for solving the problem. |
| | *Producing* occurs when a plan is carried out for solving a given problem that meets certain specifications. |

Appendix C

Tables 5 and 6

Table 5
Cross-Walk Table of Specifications of Test Items by Malay Levels/Skills and Bloom's Revised Taxonomy – Test 1

Test Item	Malay Classification		Bloom's Revised Taxonomy Classification
	Level	Skill	
1	Literal (L)	Main Idea	1. Remember
2	Literal (L)	Important Point	1. Remember
3	Literal (L)	Cause and Effect	1. Remember
4	Literal (L)	Sequence of Ideas/Events	1. Remember
5	Literal (L)	Main Idea	1. Remember
6	Literal (L)	Important Point	1. Remember
7	Literal (L)	Cause and Effect	1. Remember
8	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
9	Literal (L)	Main Idea	1. Remember
10	Literal (L)	Comparison	2. Understand
11	Literal (L)	Cause and Effect	1. Remember
12	Literal (L)	Sequence of Ideas/Events	1. Remember
13	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
14	Literal (L)	Comparison	2. Understand
15	Literal (L)	Comparison	2. Understand
16	Literal (L)	Sequence of Ideas/Events	1. Remember
17	Inferential (F)	Important Point	2. Understand
18	Inferential (F)	Comparison	2. Understand
19	Inferential (F)	Cause and Effect	2. Understand
20	Inferential (F)	Conclusion	2. Understand
21	Inferential (F)	Main Idea	2. Understand
22	Inferential (F)	Comparison	2. Understand
23	Inferential (F)	Cause and Effect	2. Understand
24	Inferential (F)	Conclusion	2. Understand
25	Inferential (F)	Main Idea	2. Understand
26	Inferential (F)	Important Point	2. Understand

27	Inferential (F)	Cause and Effect	2. Understand
28	Inferential (F)	Conclusion	2. Understand
29	Inferential (F)	Main Idea	2. Understand
30	Inferential (F)	Comparison	2. Understand
31	Inferential (F)	Comparison	2. Understand
32	Inferential (F)	Cause and Effect	2. Understand
33	Critical Creative (K)	Evaluating	5. Evaluate
34	Critical Creative (K)	Conclusion	5. Evaluate
35	Critical Creative (K)	Internalizing	5. Evaluate
36	Critical Creative (K)	Moral of the Story	5. Evaluate
37	Critical Creative (K)	Conclusion	5. Evaluate
38	Critical Creative (K)	Conclusion	5. Evaluate
39	Critical Creative (K)	Internalizing	5. Evaluate
40	Critical Creative (K)	Evaluating	5. Evaluate
41	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
42	Literal (L)	Cause and Effect	1. Remember
43	Inferential (F)	Main Idea	2. Understand
44	Inferential (F)	Conclusion	2. Understand
45	Critical Creative (K)	Evaluating	5. Evaluate
46	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
47	Literal (L)	Main Idea	1. Remember
48	Inferential (F)	Important Point	2. Understand
49	Inferential (F)	Comparison	2. Understand
50	Critical Creative (K)	Moral of the Story	5. Evaluate

Table 6

Cross-Walk Table of Specifications of Test Items by Malay Levels/Skills and Bloom's Revised Taxonomy – Test 2

Test Item	Malay Classification		Bloom's Revised Taxonomy Classification
	Level	Skill	
1	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
2	Literal (L)	Important Point	1. Remember
3	Literal (L)	Cause and Effect	1. Remember
4	Literal (L)	Sequence of Ideas/Events	1. Remember
5	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
6	Literal (L)	Comparison	1. Remember
7	Literal (L)	Cause and Effect	1. Remember
8	Literal (L)	Sequence of Ideas/Events	1. Remember
9	Literal (L)	Main Idea	1. Remember
10	Literal (L)	Important Point	2. Understand
11	Literal (L)	Cause and Effect	1. Remember
12	Literal (L)	Sequence of Ideas/Events	1. Remember
13	Literal (L)	Main Idea	1. Remember
14	Literal (L)	Comparison	2. Understand
15	Literal (L)	Comparison	2. Understand
16	Literal (L)	Cause and Effect	1. Remember
17	Inferential (F)	Main Idea	2. Understand
18	Inferential (F)	Important Point	2. Understand
19	Inferential (F)	Cause and Effect	2. Understand
20	Inferential (F)	Conclusion	2. Understand
21	Inferential (F)	Main Idea	2. Understand
22	Inferential (F)	Important Point	2. Understand
23	Inferential (F)	Cause and Effect	2. Understand
24	Inferential (F)	Conclusion	2. Understand
25	Inferential (F)	Main Idea	2. Understand
26	Inferential (F)	Important Point	2. Understand
27	Inferential (F)	Comparison	2. Understand
28	Inferential (F)	Conclusion	2. Understand

29	Inferential (F)	Main Idea	2. Understand
30	Inferential (F)	Important Point	2. Understand
31	Inferential (F)	Comparison	2. Understand
32	Inferential (F)	Cause and Effect	2. Understand
33	Critical Creative (K)	Evaluating	5. Evaluate
34	Critical Creative (K)	Conclusion	5. Evaluate
35	Critical Creative (K)	Internalizing	5. Evaluate
36	Critical Creative (K)	Moral of the Story	5. Evaluate
37	Critical Creative (K)	Conclusion	5. Evaluate
38	Critical Creative (K)	Evaluating	5. Evaluate
39	Critical Creative (K)	Conclusion	5. Evaluate
40	Critical Creative (K)	Conclusion	5. Evaluate
41	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
42	Literal (L)	Main Idea	1. Remember
43	Inferential (F)	Comparison	2. Understand
44	Inferential (F)	Cause and Effect	2. Understand
45	Critical Creative (K)	Moral of the Story	5. Evaluate
46	Literal (L)	Meaning of Word/ phrase/ sentences	1. Remember
47	Literal (L)	Important Point	1. Remember
48	Inferential (F)	Comparison	2. Understand
49	Inferential (F)	Conclusion	2. Understand
50	Critical Creative (K)	Evaluating	5. Evaluate

Appendix D

Table 7

Table 7
Comparison of the Predicted Bloom's Revised Taxonomy Classification for Each Test Item with the actual Bloom's Revised Taxonomy Classification for each test item.

Item Number	Predicted Classification	Actual Classification	Item Number	Predicted Classification	Actual Classification
1	1. Remember	2. Understand	1	1. Remember	2. Understand
2	1. Remember	1. Remember	2	1. Remember	4. Analyze
3	1. Remember	4. Analyze	3	1. Remember	
4	1. Remember	4. Analyze	4	1. Remember	4. Analyze
5	1. Remember	2. Understand	5	1. Remember	4. Analyze
6	1. Remember	1. Remember	6	1. Remember	2. Understand
7	1. Remember	1. Remember	7	1. Remember	2. Understand
8	1. Remember	2. Understand	8	1. Remember	4. Analyze
9	1. Remember	2. Understand	9	1. Remember	2. Understand
10	2. Understand	2. Understand	10	2. Understand	1. Remember
11	1. Remember		11	1. Remember	2. Understand
12	1. Remember	4. Analyze	12	1. Remember	4. Analyze
13	1. Remember	2. Understand	13	1. Remember	2. Understand
14	2. Understand	1. Remember	14	2. Understand	2. Understand
15	2. Understand	2. Understand	15	2. Understand	2. Understand
16	1. Remember	2. Understand	16	1. Remember	4. Analyze
17	2. Understand	2. Understand	17	2. Understand	2. Understand
18	2. Understand	2. Understand	18	2. Understand	4. Analyze
19	2. Understand	2. Understand	19	2. Understand	2. Understand
20	2. Understand	2. Understand	20	2. Understand	2. Understand
21	2. Understand	2. Understand	21	2. Understand	2. Understand
22	2. Understand		22	2. Understand	1. Remember
23	2. Understand	4. Analyze	23	2. Understand	4. Analyze
24	2. Understand	3. Apply	24	2. Understand	2. Understand
25	2. Understand	2. Understand	25	2. Understand	2. Understand
26	2. Understand	4. Analyze	26	2. Understand	2. Understand

27	2. Understand	2. Understand
28	2. Understand	2. Understand
29	2. Understand	2. Understand
30	2. Understand	1. Remember
31	2. Understand	4. Analyze
32	2. Understand	2. Understand
33	5. Evaluate	4. Analyze
34	5. Evaluate	4. Analyze
35	5. Evaluate	2. Understand
36	5. Evaluate	2. Understand
37	5. Evaluate	2. Understand
38	5. Evaluate	6. Create
39	5. Evaluate	4. Analyze
40	5. Evaluate	
41	1. Remember	2. Understand
42	1. Remember	2. Understand
43	2. Understand	2. Understand
44	2. Understand	2. Understand
45	5. Evaluate	2. Understand
46	1. Remember	2. Understand
47	1. Remember	1. Remember
48	2. Understand	2. Understand
49	2. Understand	2. Understand
50	5. Evaluate	2. Understand

27	2. Understand	2. Understand
28	2. Understand	5. Evaluate
29	2. Understand	2. Understand
30	2. Understand	2. Understand
31	2. Understand	
32	2. Understand	2. Understand
33	5. Evaluate	5. Evaluate
34	5. Evaluate	
35	5. Evaluate	5. Evaluate
36	5. Evaluate	2. Understand
37	5. Evaluate	2. Understand
38	5. Evaluate	5. Evaluate
39	5. Evaluate	5. Evaluate
40	5. Evaluate	2. Understand
41	1. Remember	2. Understand
42	1. Remember	2. Understand
43	2. Understand	2. Understand
44	2. Understand	5. Evaluate
45	5. Evaluate	2. Understand
46	1. Remember	2. Understand
47	1. Remember	2. Understand
48	2. Understand	2. Understand
49	2. Understand	5. Evaluate
50	5. Evaluate	

Article Citation

Dagostino, L., Carifio, J., Bauer, J. D., Zhao, Q., & Hashim, N. H. (2014). Using Bloom's Revised Taxonomy to analyze a reading comprehension instrument. *Current Issues in Education*, *18*(2). Retrieved from http://cie.asu.edu/ojs/index.php/cieatasu/article/view/1379

Author Notes

Lorraine Dagostino Mail
University of Massachusetts Lowell
O'Leary Library 519, 61 Wilder St., Lowell, MA 01854-3047
Lorraine_Dagostino@uml.edu

Lorraine Dagostino is a professor at Graduate School of Education in the University of Massachusetts Lowell. Her research interests include literacy, critical thinking and evaluative reading, theoretical issues in reading and literacy, and literature issues.

James Carifio
University of Massachusetts at Lowell
O'Leary Library 527, 61 Wilder St., Lowell, MA 01854-3047
James_Carifio@uml.edu

James Carifio is a professor in Psychology and Research methods at the University of Massachusetts Lowell. He has done extensive research and publishing in the respective fields

Jennifer D. C. Bauer
University of Massachusetts Lowell
jennifer_bauer@student.uml.edu

Jennifer D. C. Bauer is an assistant professor and Communications Department Chair at Middlesex Community College. Additionally, she is a doctoral student at the University of Massachusetts at Lowell, pursuing her Ph.D. in Literacy Studies. Her research interests include Visual Arts and Literacy, Digital/New Literacies, Language Variation, Urban Education, and Culturally Responsive Pedagogy.

Qing Zhao
University of Massachusetts Lowell
O'Leary Library, 61 Wilder St., Lowell, MA 01854-3047
Qing_Zhao@uml.edu

Qing Zhao recently earned her Ed.D. in Language Arts & Literacy from the Graduate School of Education, University of Massachusetts Lowell in 2013. Her research interests currently include the study of teaching and learning English as a foreign/second language (EFL/ESL) and development of EFL/ESL learners' language proficiency in both North American and Asian learning context.

Nor Hashimah Hashim
Universiti Sains Malaysia
shimah@usm.my

Nor Hashimah Hashim is a professor of education at the Universiti Sains Malaysia Her research interests include preschool education, primary school education, and curriculum studies

Correspondence concerning this article should be addressed to Jennifer Bauer at Jennifer_bauer@student.uml.edu.

Current Issues in Education

Mary Lou Fulton Teachers College · Arizona State University
PO Box 37100, Phoenix, AZ 85069, USA