International Journal of Assessment and Evaluation in Education
Vol 1/Dec 2011 (38-51) ISSN 2232-1926

Assessing A Hierarchy of Pre-Service Secondary Mathematics Teachers' Algebraic Thinking

Lim Hooi Lian
Wun Thiam Yew
Universiti Sains Malaysia

Noraini Idris
Universiti Pendidikan Sultan Idris

Abstract

The purpose of this study was to assess a hierarchy of Malaysian pre-service secondary mathematics teachers' algebraic thinking in using equation. The content domains incorporated in this framework were linear pattern (pictorial), direct variation, concept of function, arithmetic sequence and inverse variation. The SOLO model was used for coding the pre-service secondary mathematics teachers' responses and the Rasch model (rating scale analysis) was used to clarify the construction of the hierarchy. The participants of this study consisted of 120 pre-service secondary mathematics teachers who were first year and second year students in a local university. They were given a pencil-and-paper test. The test comprised ten superitems. Results of the study revealed that seven different levels of algebraic thinking were identified, namely prestructural, unistructural, multistructural, lower relational, relational, upper relational and extended abstract. Results of the study also indicated that 57.5% of the pre-service secondary mathematics teachers performed at least at the lower relational level (algebraic thinking level), 42.5% of them performed at the multistructural level and below (pre-algebraic thinking level). The results provided evidence of the significance of the SOLO model in assessing algebraic thinking amongst the pre-service secondary mathematics teacher.

Keywords *algebraic thinking, assessing, equation, Rasch Model, SOLO Model, pre-service secondary mathematics teacher*

INTRODUCTION

The dawn of the new reform in mathematics education recognizes algebra as a tool for problem solving, a method of expressing relationship, describing, analyzing and representing patterns, and exploring mathematical properties in a variety of problem situations (Fernandez & Anhalt, 2001).Thus, 'algebraic thinking for the new era' has become a catch-all phrase for the recent research. Many recommendations have been made to transform algebra from a skill-drilling sequence of practice into a meaningful topic that involves the generation of powerful algebraic thinking in the classroom. A reform of the school curriculum in mathematics education had also been done in some countries. For example, United States, Australia and Malaysia are characterized by Principles and Standards for School Mathematics (NCTM, 2000), National Statement on Mathematics for Australian School *(AEC,* 1991) and Integrated Curriculum for Secondary School or *Kurikulum Bersepadu Sekolah Menengah* (KBSM*)* (Ministry of Education, 2000) which have moved to an increase in emphasizing on algebraic thinking processes, understanding and applying the algebraic concepts in making generalization and representation in authentic problem situation. They moved away from the traditional view of algebra which is tight focusing on procedural skills and symbol manipulative facility.

Since the new curriculum (Integrated Curriculum for Secondary School or *Kurikulum Bersepadu Sekolah Menengah)* has been implemented for fourteen years (1988-2012) in Malaysia, it is time to discuss the achievement of the aims of this curriculum. Are students able to think mathematically, creatively, and critically in learning algebra? This question can be answered by reflecting on the assessment system, especially the test items. According to Cheah and Malone (1996) and Teng (2001), the current written tests still dominate the traditional format where answers are unique for each item. For instance, the format of the open-ended item in upper secondary school national level examination (*Sijil Pelajaran Malaysia*) seems to discourage students to apply the algebraic equation method to solve it. The need for using the equation method is not obvious (Lim, 2007). The students working through these items may not see why an algebraic representation is useful. It is easier to solve it without any algebraic knowledge, either by trial and error, or by logical reasoning and arithmetic. In short, the ability to operate with unknowns, developing of algebraic thinking is being sidelined.

In order to address the desired outcomes such as algebraic problem solving skills, algebraic reasoning and representation, the item format needs to offer the problem that challenges students' abilities as well as typically assessing ability to perform mathematical procedures. In Malaysia, while research has been done to investigate the variety of misconceptions in algebra (Cheah & Malone, 1996; Teng, 2001), very little consideration has been given to the assessment of algebraic thinking. Thus, the question of how to assess thinking ability in algebra may still be new and challenging for many educators. Some of them do not realize the assessment of algebraic thinking is a very useful foundation for the understanding of algebraic concepts. Moreover, this situation places the secondary teacher in a crucial and critical position; in fact the degree to which they are capable of developing the students' algebraic thinking may determine the success of the algebra reform. However, most teachers have little experience with the rich and connected aspects of algebraic thinking that need to become the norm in schools (Blanton & Kaput, 2003). Zazkis and Liljedahl (2002) indicated that although most of the pre-service elementary school teachers were able to express generality verbally, they were unable to use algebraic notation comfortably and confidently. Ahuja (1998) found that a majority of the pre-service primary teachers in his study cannot see algebra as a generalized arithmetic. They even faced serious difficuilties in understanding the basic concept of algebra, namely variable and algebraic expressions. Thus, this study focused on the pre-service secondary mathematics teachers' level of algebraic thinking as particularly important factors impacting algebra reform in the secondary school.

In this study, the SOLO model, known as the Structure of the Observed Learning Outcome, developed by Biggs and Collis (1982), was used to construct a superitem test which reflects the four levels of the SOLO model.

THEORETICAL FRAMEWORK

The SOLO model was designed mainly as a means to measure student's cognitive ability in academic learning contexts (Biggs & Collis, 1982, 1989). It has been used to analyze the structure of student's mathematical thinking, understanding of mathematical concepts and problem solving ability over a wide educational span from primary to tertiary levels (Callingham & Pegg, 2009; Chick, 1988; Collis, Romberg, & Jurdak, 1986; Lam & Foong, 1998; Panizzon, Callingham, Wright & Pegg, 2007; Reading, 1999; Vallecillos & Moreno, 2002; Watson, Chick, & Collis, 1988; Wilson & Iventosch, 1988). It also has been applied in the area of science (Lake, 1999; Levins, 1997), counseling (Burnett, 1999) and practice subject (Chan, Tsui & Chan 2002).

The SOLO model suggests that when students answer the tasks given, their responses display a similar sequence across the task. This leads to the identification of the stages at which a student is currently operating (Biggs & Collis, 1982). In these consistent sequences, the following stages occur:

1. Prestructural - response consists only of irrelevant information to the task. In other words, the task is not attacked appropriately; the students have not really understood the task given.
2. Unistructural - response includes one relevant piece of information and it is treated independently. Thus, it may display a premature response because all available information has not been utilized.
3. Multistructural - response includes several pieces of relevant information without relating them to each other. They are seen as discrete and unrelated elements.
4. Relational - response integrates all relevant pieces of information. It includes several conclusions from the available information given. However, the explanation is still relatively based on the context.
5. Extended abstract - response not only includes all relevant pieces of information given but extends the response to integrate relevant pieces of information which are not given. It also includes hypothetical situation from a generalization.

As originally developed, the SOLO model used an open response format in which student responses were examined for structural organization by an assessor. A later development enabled the technique to be used with a closed format, which is also called "superitem". Collis, Romberg and Jurdak (1986), Lam and Foong (1998), Wilson and Iventosh (1988), and Aoyama (2007) developed the use of superitem based on the SOLO taxonomy as an alternative assessment tool for monitoring the growth of students' cognitive ability in solving mathematics problems. A superitem consists of a problem situation and four different complexity levels of items related to it. The problem situation is often represented by text, diagram or graphic. The SOLO model assumes a latent hierarchical and cumulative cognitive dimension. While the items represent four main levels of reasoning defined by the SOLO model, it found that when students answered the tasks given, their responses to the tasks could be summarized in terms of the four levels (Biggs & Collis, 1982; Biggs & Collis, 1989; Chick, 1988; Chick, 1998; Wilson & Iventosch, 1988; Wongyai & Kamol, 2004; Wilson & Chavarria, 1993), ranging from unistructural to extended abstract. Table 1is an example of a superitem that was used to assess pre-service secondary mathematics teachers' algebraic thinking in using equation to solve the problem situation.

Table 1 Superitem 9 (Concert Hall)

The first row of Mega Mall concert hall has 10 seats. Each row thereafter has 2 more seats than the row in front of it.		
Level	Item	Descriptor
Unistructural	a. How many seats are in row 2 in Mega Mall concert hall?	The item requires that the student identifies the next term in the sequence by referring directly to the given information in the stem (the first term and common difference).
Multistructural	b. How many seats are in row 13 and 18 in the concert hall?	The item requires the given information that is handled serially. The student identifies the recursive relationship between the terms to solve the specific cases.
Relational		The item requires that the student integrates all the given information to make generalization by forming an algebraic expression and equation. If the student provides this response, it would demonstrate his/her thinking ability in identifying the linear relationship between variables and applying algebraic symbols to make the representation. Besides, the student would show his/her ability to apply the rule to solve the problem given.
a. Lower Relational	How many seats are in row ***k*** in the concert hall?	
b. Intermediate Relational	Write an equation to find the number of seats for any number of rows. Let ***s*** represents the number of seats and ***r*** represents the number of rows.	
c. Upper Relational	If the last row has 406 seats, try to use the equation to find the number of rows in the hall.	
d. Extended Abstract	c. The manager planned to prepare 1000-1500 seats in row 100 for a musical concert. Will he make it? If yes, explain your answer. If no, try to suggest a new equation in order to help the manager.	The item requires that the student extends the understanding of linear relationship by evaluating the relevance and applicability of equation in the related situation and forms an appropriate alternative solution for it.

The four responses (unistructural, multistructural, relational and extended abstract) listed in Table 1 represented both an increase in the use of the information available and an increase in the complexity of structural response. For example, the unistructural and multistructural responses may only involve one or relevant aspects of given information in the stem and thus there is little relationship between the question and the given information. Further, in these levels, a student needs to only encode the given information and use it directly to give a response. Whereas, at the relational or extended abstract level, the student needs to make a generalization within the given information or abstract principle which was not given directly in the stem. Besides, at the relational or extended abstract level, the student needs to understand the task in a way that is personally meaningful and links up with the existing knowledge. Thus, within any superitem, a correct response to an item would indicate the cognitive ability to respond to the information in the stem at the certain level reflected in the SOLO structure.

In the area of algebra, the SOLO model has been used to describe student's elementary equation solving (Biggs & Collis, 1982) and made the comparison with various learning theories in describing development of algebraic ideas (Pegg, 2001) but there is no coherent description of the student's algebraic thinking sufficient to inform instructional decisions. Thus, in this study, we claim that the proposed framework enables pre-service secondary school mathematics teachers' algebraic thinking to be described across four main levels of the SOLO model.

PURPOSE OF THE STUDY

As the emphasis on algebraic thinking in mathematics increases, the need for assessing pre-service secondary mathematics teachers' algebraic thinking ability becomes more pressing. This study aims to assess a hierarchy of pre-service secondary mathematics teachers' algebraic thinking in using an equation to solve a problem situation. In order to capture the manifold nature of algebraic thinking, the framework of this study incorporates five content domains of equation based on the Malaysian school mathematics syllabus. These include: linear pattern (pictorial), direct variation, concept of function, arithmetic sequence and inverse variation.

SIGNIFICANCE OF THE STUDY

This study may provide a guideline to educators who want to identify the levels and the processes of algebraic thinking among the pre-service teachers in using equation across the five content domains. Subsequently, the instrument of the study might also be used as a diagnostic assessment reference to evaluate the weaknesses of the student's conceptual understanding of equation. The findings of this study might give indications concerning students from different levels of algebraic thinking and may provide many possible answers under algebraic thinking situations. The favorable results from this study imply that there are possible alternatives for the assessment of algebraic thinking. Consequently, the findings might provide useful information to the assessment developer particularly in the subject of mathematics.

THE STUDY

This study used a quantitative approach to assess the pre-service secondary mathematics teachers' level of algebraic thinking based on the SOLO model. The dataset was submitted to rating scale analysis.

Participants

In Malaysia, the first topic of algebra is taught during lower secondary school (Form One or Grade Seven). The students begin to face the more complex topics such as straight line; gradient and area under graph; index and logarithm; matrix; variation; graph of function and quadratic equation which are related to algebra in Form Four, Form Five and Form Six. Thus, the construction of an assessment guide to assess algebraic thinking amongst pre-service secondary mathematics teachers is important as they had followed the Mathematics syllabuses at the school level.

The participants of this study consisted of 120 pre-service secondary mathematics teachers who were first year and second year students in a local university. They took mathematics in Form Six (pre-university level) or Matriculation level and now they are majoring or minoring in mathematics. These participants were randomly chosen.

Instrumentation

In this study, the instrument for data collection consisted of a pencil-and-paper test of ten superitems. Two superitems were constructed for each content domain to be assessed. The concert hall problem discussed above is an example of a superitem designed for this study.

Data Analysis

The data analysis had been done based on the findings from a pencil-and-paper test. The test paper results were analyzed by using the Rasch model. The Rasch model (Wright & Masters, 1982) is based upon the difficulty of item, assuming that item difficulty is the main characteristic influencing the responses. The Rasch model used in Winsteps for the polytomous item (an item which has more than two possible responses) analysis is the "Rating Scale" analysis with the equation:

$$\log[P\mathrm{nij}/Pni(j\text{-}1)] = Bn - Di - Fj$$

Where Pnij is the probability that person encountering item i is observed in category j, Bn is the "ability" or rater-severity measure of person n, Di is the difficulty-to-endorse measure of item i, and Fj is the "calibration" measure of category j relative to category (j - 1) (Bradley, Cunningham & Gilman, 2006).

The Rasch model uses the sum of the item ratings simply as a basic point for estimating probabilities of those responses. Because it is based upon the ability to endorse a set of items and the difficulty of a set of items, it is assumed item difficulty is the main characteristic influencing responses (Bradley & Sampson et al, 2005). Here, two facets are involved, the instrument's items and the respondents. From a Rasch perspective, a respondent's willingness to endorse interacts with an item's difficulty to assign a certain score to produce an observed outcome (Bradley & Sampson, 2005). In general, people are more likely to endorse easy items than those that are difficult and people with higher willingness to endorse scores are more agreeable than those with low scores. In the context of testing, students with higher ability are more able to solve more difficult items than those who are lower ability. Rasch analysis reports person willingness to endorse and item difficulty to endorse estimates along a logit (log odds unit) scale, "a unit interval scale in which the unit intervals between the locations on the person-item map have a consistent value or meaning" (Bond & Fox, 2001, pp. 29). Bond and Fox explain that employing Rasch techniques allows for the ordering of respondents along this continuum of willingness to endorse items and orders items along a continuum according to their difficulty to endorse. "Based on this logic of order, the Rasch analysis software programs perform a logarithmic transformation on the item and person data to convert the ordinal data to yield interval data…actual item and person performance probabilities determine the interval sizes" (Bond & Fox, 2001, pp. 29). For example, the ordinal values 0, 1, and 2 might be applied to an item which has three ordered performance category levels as following: 0 = disagree, 1 = neutral correct and 2 = agree. In this study, the ordered values 1, 2, 3, 4, 5, 6 and 7 might be applied to the superitem as following: 1 = totally wrong or no response (prestructural level), 2 = unistructural level, 3 = multistructural level, 4 = lower relational level, 5 = intermediate relational level, 6 = higher relational level and 7 = extended abstract level. Code 1, 2, 3, 4, 5, 6 and 7 covered all the response possibilities in the test. Code 1 is the code for the lowest possible response level and code 7 is the code for the highest possible response level in each superitem.

In this study, the Rasch model was also used to determine the validity of the underlying theorised construct. In the Rasch model, construct validity can be examined by identifying the fit to the model of both items and persons. Infit statistic is the most commonly used measure of fit. The mean infit statistic may be considered acceptable if its values lie between 0.77 and 1.3 (Adams & Khoo, 1996). If all the superitems can be shown to be systematically related to each other along the variable, this is taken as confirmation that a single construct is being measured in this study.

RESULTS

Table 2 shows the reliability of test items and the reliability of participants of this study. Item reliability index scores indicates the replicability of item placements along the pathway if these same items were given to other samples with comparable ability levels. For example, if other participants were given these same items, the items would estimate stability. In this analysis, the item reliability of estimate was 0.95, meaning that researchers can quite readily rely on this order of item estimate to be replicated when it is given to other participants for whom it is suitable. Person reliability index indicates the replicability of participants ordering that could be expected if the participants were given another set of items measuring the same construct (Bond & Fox, 2001). In this study the value of person reliability index 0.97 was high, indicating a consistency of the participants' ability estimates.

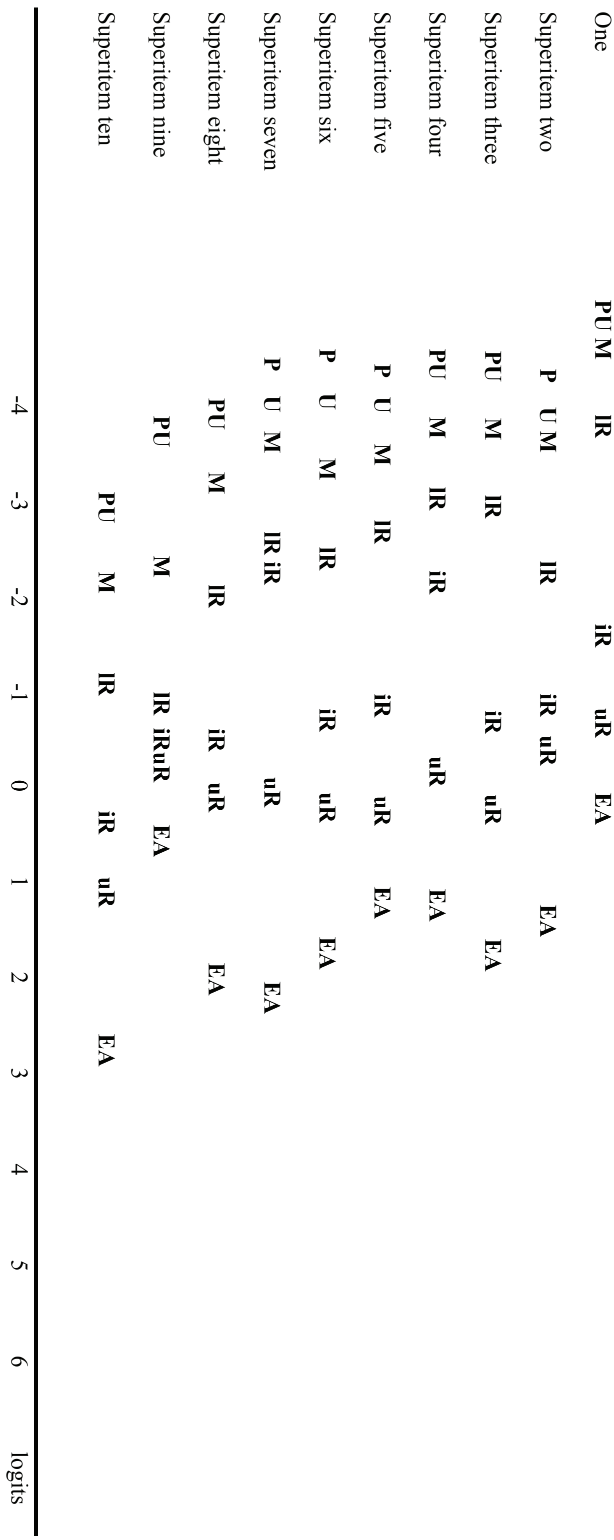
"Infit Mean Square" relates the differences between expected scores and actual scores. If the fit statistics (infit and outfit) of an item or a participant is acceptable, the expected value of the mean square (variation in the observed data) is shown between 0.7 and 1.3. In the analyses, the item and person infit mean square fell within the acceptable range, 0.95 and 0.94 respectively, which indicates that the tasks form a hierarchical one-dimensional scale. Cronbach Alpha is the test reliability coefficient that estimates the quotient between the variances of expected and observed value (Aoyama, 2007). The value of 0.98 implies that the data of this study was reliable. In conclusion, all indicators were high enough to suggest the existence of a hierarchical construct within these test items.

Table 2 Reliability and Fit Indices

1.	Item reliability	0.95
2.	Item infit mean square	0.95
3.	Person reliability	0.97
4.	Person infit mean square	0.94
5.	Cronbach Alpha	0.98

Figure 1 shows the difficulty of each level item within each superitem. All the superitems were ordered by their difficulty levels within the test. In other words, the items increased in difficulty as they moved up the SOLO model level, as one might have predicted. The level of difficulty of each superitem is consistent and easily interpreted. It provides support to the sequence of SOLO levels of response and helps to confirm the construct validity of the test. The findings presented above are in line with those of Wilson and Iventosch (1988), who found that all their seven open-ended questions' difficulties increased as they moved up the levels of taxonomy. Within each superitem, the order of the items is unistructural, multistructural, lower relational, relational, upper relational and extended abstract. This expected order was found to hold for all the superitems.

Figure 1 also shows the presence of differences in difficulties between the levels of the items. The difference increases as the thinking level increases especially from lower relational level to intermediate relational level and upper relational to extended abstract. The unistructural level items and multistructural level items tested the arithmetic concept of the pattern and tended to be easy compared to their respective lower relational level items which required the pre-service secondary mathematics teachers to make generalization in the form of algebraic expression. On the other hand, upper relational level items and extended abstract level items tested much higher level thinking, namely pre-service secondary mathematics teachers' have to show their abilities to apply the concept of algebraic equation to make a generalization of the pattern given, new pattern or new situation.



P = totally wrong or no response (prestructural)
U = unistructural level
M = multistructural level
IR = lower relational level
iR = intermediate relational level
uP = upper relational level
EA = extended abstract level

Figure 1 The Rasch and superitem logit scale

		8.0	
X			
		7.0	
			Extended abstract
		6.0	
X	7	5.0	
XX			
XXX		4.0	
XXX	18		**Upper relational**
XXXXXXXXXXXXXXX		3.0	
XXX	13		
XXXX		2.0	**Intermediate relational**
XXXXXX			
XXX	31	1.0	
XXXXX			
XXXX		0	**Lower relational**
XXXXXXXXXXXXXXXXXXX			
		- 1.0	
XXXX	24		**Multistructural**
XXXXXXXXXXXXXXXXXXXX		- 2.0	
XXXXXXXX	21		
XXXXXXXXXXXXX		- 3.0	**Unistructural**
XXXXXX	6	- 4.0	**Prestructural**

Figure 2 Distribution of Participants and Levels

Figure 2 shows the distribution of pre-service secondary mathematics teachers in each of the SOLO levels. From the findings, 57.5% of the pre-service secondary mathematics teachers performed at least at lower relational level. Seven of them were able to extract the abstract general principle (algebraic equation concept or other related concept) from the information given to form an alternative solution for the new pattern (extended abstract level), 18 of them showed their abilities to apply the rule in the form of algebraic equation to solve the related problem (upper relational level), 13 of the pre-service secondary mathematics teachers could make a generalization of the pattern in the form of algebraic equation and 31 of them were able to form an algebraic expression as the rule of the pattern. In other words, 69 of the pre-service secondary mathematics teachers in this study were able to apply algebraic concepts to solve the task given.

The findings also indicated that 42.5% of the pre-service secondary mathematics teachers performed at multistructural level and below. Twenty-four performed at multistructural level, 21 performed at unistructural level and 6 of them performed at prestructural level. All the pre-service secondary mathematics teachers at these levels could be classified into lower level of algebraic thinking. Generally, most of them were only able to numerically solve a variety of problems involving specific cases. They encountered difficulties in generalizing the arithmetic through the use of algebraic symbols.

DISCUSSION AND CONCLUSION

The SOLO model is claimed to be applicable in assessing cognitive learning outcomes among different levels of pre-service secondary mathematics teachers. This study had found evidence to support such claim (Chick, 1988; Collis, Romberg & Jurdak, 1986; Lam & Foong, 1998; Reading, 1999; Vallecillos & Moreno, 2002; Watson, Chick &Collis, 1988; Wilson & Iventosch, 1988). Seven levels of sophistication in algebraic thinking can be found in pre-service secondary mathematics teachers' responses to the tasks: prestructural, unistructural, multistructural, lower relational, intermediate relational, higher relational and extended abstract. The data analyses demonstrated that 42% of participants achieved algebraic thinking at prestructural level, unistructural level and multistructural level. They were able to numerically solve a variety of problem involving specific cases but they encountered difficulties in making generalization through the use of an equation. These findings are consistent with previous research finding (Zazkis & Liljedahl, 2002) that majority of the pre-service secondary mathematics teachers were able to solve the problems involving specific cases, explain the sequence of pattern only in terms of difference between successive terms and very few of them were able to generalize the problem into algebraic form.

In order to develop and stimulate the new era of algebraic thinking amongst the students, it is essential that the pre-service teachers have a minimum relational level of algebraic thinking. However, this study shows that almost half of the prospective teachers surveyed were at the prestructural level, unistructural level and multistructural level. The findings revealed that many of these future teachers cannot see algebra as a powerful tool to generalize patterns. In fact, lack of understanding of the concept of an unknown, a variable and an equation is an important factor for these prospective teachers. The study agrees with Ahuja (1998) who found that many pre-service teachers learned manipulation rules without reference to the meanings of the expressions being manipulated.

As a consequence, teacher educators should help their pre-service teachers to develop their knowledge base related to algebra. The pre-service teachers should be engaged in exploring algebraic concepts and procedures from new perspectives, for instance, exposing the pre-service teachers to some activity based courses in algebraic thinking in order to provide them with a deeper understanding of algebraic concepts and development of algebraic reasoning.

As noted above, the SOLO model can be used in the writing of an item with the format of a superitem, it also can be used to score the item and allow for crediting partial knowledge. Thus, it provides educators with an indication of different levels of algebraic thinking that they can expect to encounter in their classroom. Thus this model has the potential to contribute to both instruction and assessment. In the instructional perspective, it would seem prudent for educators to use thinking ability level descriptors as broad guidelines for organizing instruction and building problem task. From an assessment perspective, it appears to be valuable in providing educators with useful background information on pre-service secondary mathematics teachers' initial solving ability, and in enabling them to monitor general growth in algebraic thinking.

One of the purposes for developing the paper and pencil test used in this research was to provide an instrument that would be practical for use in a range of secondary school classrooms and university. In this study, the coding was carried out by researchers. It would be of interest for others to use the test in the classroom. It is the researchers' view that the test may offer rich potential for documenting students' progress in the area of algebra over the secondary school years.

In this study, the high value of reliability and validity strongly supported the existence of uni-dimensional scale for algebraic thinking. The findings also support the idea of progressive levels of algebraic thinking for pre-service teachers within the five content domains, described by uni-dimensional sequence (from prestructural level to extended abstract level). However, this should not rule out the possibilities of two dimensional or even more complex constructions (Aoyama, 2007).

ACKNOWLEDGEMENTS

This project is made possible with funding from the Incentive Grant of University Sains Malaysia, Penang, Malaysia.

REFERENCES

Adams, R. J., & Khoo, S. T. (1996). *Quest: Interactive item analysis system. Version 2.1* [computer software]. Melbourne, Australia: Australian Council for Educational Research.

Ahuja, O. P. (1998). Importance of algebraic thinking for pre-service primary teachers. *The Mathematics Educators*, *3*(1), 72 – 92.

Australian Education Council (1991). *A national statement on mathematics for Australian school*. Cariton, VIC: Curriculum.

Aoyama, K. (2007). Investigating a hierarchy of pre-service secondary mathematics teachers' interpretation of graphs. *International Electronic Journal of Mathematics Education, 2*(3), 298 – 318.

Blanton & Kaput (2003). *Developing Elementary teachers' algebra 'Eyes and Ears' understanding characteristics of professional development that promote generative and self-sustaining change in teacher practice.* Retrieved May 11 2009, from http://www.scps.k12.fl.us/scctm/TextFiles/pdf.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome).* New York: Academic Press.

Biggs, J. B., & Collis, K.F. (1989). Towards a model of school-based curriculum development and assessment using the SOLO taxonomy. *Australian Journal of Education*, *33*(2), 151 – 163.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement* i*n the human sciences*. New Jersey: Lawrence Erlbaum Associates.

Bradley, K. D., Cunningham, J. D. & Gilman, R. (2006). *Using the Rasch Rating Scale Model to test the fit and function of MSLSS scale across groups*. Retrieved May 9 2009, from http://www.uky.edu/~kdbrad2/MSLSS-AeRA2006.pdf.

Bradley, K. D. & Sampson, S. O. (2005). *Improving data collection through Raschmeasurement: A continuing study of teacher supply and demand*. Retrieved May 9 2009, from www.uky.edu/~kbrad2/AERASupplyDemandHandout.pdf.

Burnett, P. C. (1999). Assessing the structure of learning outcome from counseling using the SOLO Taxonomy: An exploratory study. *British Journal of Guidance & Counseling*, *7*(1), 567–580.

Callingham, R. & Pegg, J. (2009). Using developmental frameworks to support curriculum outcomes. *Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia.* Fremantle: MERGA.

Chan,C. C., Tsui, M. S., & Chan, M. Y. C. (2002). Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on student's learning outcomes: An empirical study. *Assessment and Evaluation in Higher Education*, *27*(6), 511 – 528.

Cheah, C. T., & Malone, J. A. (1996). Diagnosing misconceptions in elementary algebra. *Journal of Science and Mathematics Education in South East Asia, 14*(1), 61 – 68.

Chick, H. L. (1988). Student responses to polynomial problem in the light of the SOLO taxonomy. *Australian Senior Mathematics Journal, 2*(2), 91 – 110.

Chick, H. L. (1998). Cognition in the formal modes: Research mathematics and the SOLO Taxonomy. *Mathematics Education Research Journal*, *10*(2), 4 – 26.

Collis, K. F., Romberg, T. A., & Jurdak, M. E. (1986). A technique for assessing mathematical problem-solving ability. *Journal for Research in Mathematics Education*, *17*(3), 206 – 221.

Fernandez, M. L., & Anhalt, C. O. (2001). Transition toward algebra. *Mathematics Teaching in the Middle School, 7*(4), 236 – 242.

Kuchemann, D. (1981). Algebra. In K. Hart (Ed.), *Children's understanding of mathematic: 11 – 16 (pp 102 –119).* London: Murray.

Lake, D. (1999). Helping students to go SOLO: Teaching critical numeracy in the biological science, *Journal of Biological Education*, *33*(4), 191–199.

Lam, P., & Foong, Y. Y. (1998). Assessment of mathematics structure of learning outcome proficiency attainment level using hierarchical items in testlets. *Educational Research Quarterly, 27*(2), 3 – 15.

Levins, L. (1997). Assessment of students' outcome using a theoretical framework. *Australian Science Teachers' Journal*, *43*(1), 56-60.

Lim, H. L. (2007). *Penggunaan Model SOLO dalam penilaian kebolehan penyelesaian persamaan linear pelajar Tingkatan Empat*. Unpublished PhD Thesis. University of Malaya .

Ministry of Education. (2000). *Kurikulum Bersepadu Sekolah Menengah: Huraian Sukatan Matematik Tingkatan Empat.* Kuala Lumpur: Pusat Perkembangan Kurikulum.

National Council of Teachers of Mathematics. (2000*). Principles and Standards for School Mathematics*. Reston VA: Author.

Orton, A., & Orton, J. (1994). Pre-service secondary mathematics teachers' perception and use of pattern and generalization. *Proceedings of the 18th Conference of the International Group for the Psychology of Mathematics Education, Lisbon: University of Lisbon* (pp. 404–414).

Panizzon, D., Callingham, R., Wright, T. & Pegg, J. (2007). Shifting sands: Using SOLO to promote assessment for learning with secondary mathematics and science teachers. In P. Jeffery (Ed.), *Proceedings of the AARE annual conference, Fremantle*. Retrieved from http://www.aare.edu.au/07pap/pan07110.pdf

Pegg, J. (2001). *Fundamental cycles in learning algebra.* Center for Cognition Research inLearning and Teaching, University of New England Research Retrieved September 7, 2003, from http://www.edfac.unimelb.edu.au.

Reading, C. (1999). *Understanding data tabulation and representation.* Centre for Cognition Research in Learning and Teaching, University of New England. Retrieved October 1, 2003, from http://fehps.une.edu.au./f/s/curric/cReading/

Stephens, A. C. (2006). Equivalence and relational thinking: Preservice elementary teachers' awareness of opportunities and misconceptions. *Journal of Mathematics Education,* 9, 249–278.

Swafford, J. O., & Langrall, C. W. (2000). Grade 6 pre-service secondary mathematics teachers' preinstructional use of equation to describe the represent problem, *Journal for Research in Mathematics Education, 31*(1), 89–112.

Teng, S. L. (2002). *Konsepsi alternatif dalam persamaan linear di kalangan pelajar Tingkatan Empat*. (Unpublished master's dissertation). Universiti Sains Malaysia.

Vallecillos, A., & Moreno, A. (2002). Framework for instruction and assessment on elementary inferential statistics thinking. *Proceedings of the Second International Conference on the Teaching of Mathematics, Greece* (pp. 1–9).

Watson, J., Chick, H., & Collis, K. (1988). Applying the SOLO taxonomy to error on area problems. *Proceedings of the 12th Biennial Conference of the Australian Association of Mathematics Teachers, Newcastle* (pp. 260-281).

Wilson, L.D., & Chavarria, S. (1993). Superitem test as a classroom assessment tool. In N.L. Webb & F. A. Arthur (Eds.), *Assessment in the mathematics classroom* (pp. 135 – 142). Virginia: NCTM.

Wilson, M., & Iventosch, L. (1988). Using the partial credit model to investigate responses to structured subtest. *Applied Measurement in Education, 1*(4), 319 – 334.

Wongyai, P., & Kamol, N. (2004). *A framework in characterizing lower secondary school pre-service secondary mathematics teachers' algebraic thinking.* Retrieved December 23, 2004, from http://www.icme- organisers.dk/tsg09/PiyavadeeWongyai.pdf

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.

Zazkis, R., & Liljedahl, P. (2002). Generalization of patterns: The tension between algebraic thinking and algebraic notation. *Educational Studies in Mathematics, 49,* 379 – 402.