Newman Error Analysis: A Comparative Study Involving Year 7 Students in Malaysia and Australia

Nerida F. Ellerton
Edith Cowan University

and

M. A. ("Ken") Clements
The University of Newcastle

After 206 Year 7 students—145 in 4 schools in Malaysian, and 61 in 2 schools in Australia—had answered 24 mathematics questions, they were interviewed, in accordance with the Newman interview technique. Types of errors made by the students in the two countries were analysed and compared. Data revealed that (a) around 70% of all errors were in one of the Comprehension, or Transformation, or Careless categories; and (b) strikingly different error patterns occurred for different questions.

In the mid-1970s an Australian language educator, M. Anne Newman (1977a,b) developed a systematic procedure for analysing errors made by students who had given answers to questions on written mathematics tests. According to Newman, any person confronted with a written mathematics task needs to go through a fixed sequence: Reading (or Decoding), Comprehension, Transformation (or Mathematising), Process Skills, and Encoding. Errors can also be the result of unknown factors, and Newman (1983) assigned these to a composite category, termed "Careless."

Over the past two decades the "Newman method" has been widely used throughout the Asia-Pacific region—in Australia (eg, Casey, 1978; Clarkson, 1980; Clements, 1980; Clements & Ellerton, 1992; Faulkner, 1992; Tuck, 1983; Watson, 1980); in Brunei (Mohidin, 1991); in India (Kaushil, Sajjin Singh & Clements, 1985); Indonesia (Ora, 1992); in Malaysia (Ellerton & Clements, 1992; Teoh Sooi Kim, 1991; Kownan, 1992; Marinas & Clements, 1990; Sulaiman & Remorin, 1993); in Papua New Guinea (Clarkson, 1983, 1991; Clements, 1982); Singapore (Kaur, 1995); in the Philippines (Jiminez, 1992); and in Thailand (Singhatat, 1991; Sobhachit, 1991; Thongtawat, 1992).

Analyses of data based on the Newman procedure have drawn special attention to the influence of language factors on mathematics learning, and by so doing have challenged mathematics educators and teachers to redefine what is "basic" in school mathematics. Newman research has generated a large amount of evidence pointing to the conclusion that far more children experience difficulty with the semantic structures, the vocabulary, and the symbolism of mathematics than with standard algorithms. Also, findings of Newman research have consistently pointed to the inappropriateness of many "remedial" mathematics programs in schools in which there has been an over-emphasis on the revision of standard algorithms, with hardly any attention being given to difficulties associated with Comprehension and Transformation.

In many Newman studies the proportion of errors first occurring at the "Transformation" stage has been great (in some studies, more than 50% of breakdowns first occurred at that stage).

Newman Research and Tests in Different Languages

Clarkson (1983) and Newman (1977a) reported Newman data for which 12 per cent of the initial errors were in the Reading category. With all other Newman error analysis studies, however, not more than 2% of initial errors have been in the Reading category (Clements & Ellerton, 1992). In case it is concluded that reading has less influence on students' responses to mathematics word problems than might be expected, it should be pointed out that by "Reading" Newman merely meant "Decoding," that is to say the ability to sound the words or symbols correctly. The second Newman Category, Comprehension, has proved to be of considerable importance in Newman research, and it is interesting that most language experts would regard comprehension as being an integral component of the act of reading.

A number of recent Newman error analysis studies carried out in schools in Southeast Asian nations have suggested that a high proportion of initial errors occurred at the Comprehension stage or at the Transformation stage. Thus, for example, Marinas and Clements (1990), working with Year 7 students in Malaysia, found that 71% of errors made on typical Year 7 mathematics questions were in the Comprehension or Transformation categories. Singhatat (1991), working with Thai students, found that 68% of errors were in either of these two categories. Neither Marinas and Clements nor Singhatat found that students in their samples made any Reading (i.e. Decoding) errors.

However, as far as we know there have been no Newman error studies carried out in which students with different language backgrounds took the "same" test in their own first language. In one sense this would impossible to achieve because different languages have different syntactic and semantic structures. However, it would be of interest to try to create two tests covering the same mathematics content but in different languages, with every attempt being made to preserve the semantic structures of the questions—that is to say, to match corresponding questions so far as semantic structure was concerned.

This paper describes an attempt to construct two forms of a mathematics test: one written in *Bahasa Melayu* and the other in the English language. Once the two forms of the test were constructed, Newman error patterns generated by Year 7 students in Malaysia and Australia responding to the form which represented their first language (*Bahasa Melayu* in Malaysia, English in Australia), were analysed and compared.

Comparing Error Patterns of Malaysian and Australian Year 7 Students on Comparable Forms of the Same Test

Methodology

It was decided to investigate the extent to which Newman error patterns for Australian children were similar to those of children in Malaysia, of the same age and grade level. A sample of Australian Year 7 students took an English-language form of a test comprising 24 pencil-and-paper questions, and a sample of Malaysian Year 7 students took the *Bahasa Melayu* form of the same test.

The instrument(s). In order to carry out this investigation it was necessary to establish two comparable tests, one in *Bahasa Melayu* (the language of instruction in Malaysia) and one in the English language (the language of instruction in Australia). After discussion with linguists and with experienced mathematics educators and teachers in Malaysia and Australian, a 24-question test was developed containing only items which were regarded as representing essential knowledge and skills for Year 7 students in both nations.

The test was first written in English. It was then translated into *Bahasa Melayu* by Malaysian mathematics educators. Back translation methods (Brislin, 1970) were used to check that each question in a pair of corresponding questions was as much like its pair as possible—with particular attention being paid to content, vocabulary, and semantic structure (Riley, Greeno, & Heller, 1983). Every teacher of the pupils involved in the study agreed that the questions were well worded, and covered material which they had taught to their students.

The questions were carefully chosen to include basic skills questions, standard word problems, and unfamiliar word problems for which an appropriate method was not obvious. The three types of questions are illustrated by Questions 2, 13, and 17, which were all concerned with subtraction:

Question 13:

(English): Find the value of 940 – 586.

(Bahasa Melayu): Selesaikan 940 – 586.

Question 17:

(English): There are 950 students in a school, but one day only 587 of them were present. How many students were absent that day?

(Bahasa Melayu): Terdapat 950 orang pelajar di sebuah sekolah, tetapi hanya 587 orang pelajar hadir pada satu hari tertentu. Berapa ramaikah pelajar yang tidak hadir padar hari itu?

Question 2:

(English): There are 20 boys and girls altogether in a room. If there are 4 more girls than boys, how many boys are there in the room?

(Bahasa Melayu): 20 orang budak lelaki dan perempuan berada di dalam sebuah bilik. Jika terdapat 4 orang budak perempuan lebih daripada budak lelaki, berapa orangkah budak lelaki di dalam bilik itu?

Some of the questions tested understanding of measurement concepts (time, area, volume, money, mass, distance/speed relationships), some were concerned with spatial ideas, and many involved important aspects of number (the four operations on whole numbers, ratio, fractions, percentages).

The sample. The location of samples was determined by the availability of trained interviewers. In Malaysia, 145 Year 7 students in four Year 7 classes (in three different schools) participated in the study; in Australia, 61 Year 7 students in two Year 7 classes (in different schools) were involved. One of the Malaysian classes was a top-stream group in a school located in a middle-class suburb in Kuala Lumpur; another of the classes was a non-streamed group in a highly regarded school in Penang; the other two classes were from a school located in a rural part of Malaysia. Both of the Australian classes were in schools in Perth: one was a top-stream group in a school in a middle-class suburb (referred to as m/c), the other was a heterogeneous group in a school in a working-class suburb (referred to as w/c)

The interviewers. All interviewers were senior and experienced educators who had been trained by the authors to carry out Newman interviews. The interviews were conducted, using the normal Newman interview method, in the students' language of instruction (which, in all cases, was the first language of the interviewers)—*Bahasa Melayu* in Malaysia, and English in Australia. All Newman error classifications made by the interviewers during the interviews were subsequently checked by the authors.

Results

Table 1 summarises the results, and it can be seen that the overall error profiles for the Malaysian and Australian samples were fairly similar. However the following within- and between-sample differences should be noted:

1. In each of the Malaysian subsamples, the sum of the percentages for Reading, Comprehension, and Transformation was at least 70%, but this was not the case for the two Australian groups.
2. Higher proportions of the errors made by Australian students than of errors made by the Malaysian students were classified as being of the "Process Skills" variety (14% for Australian children, but only 3% for Malaysian children).
3. In all 6 schools, only a small proportion of errors were classified as Reading errors.
4. A relatively high proportion of errors made by high achieving students (especially in the Perth m/c school and in the Kuala Lumpur group) were "Careless." This is in line with results obtained in an early study carried out in Papua New Guinea by Clements (1982).

Table 1
Results of Newman Interviews of Year 7 Students in Malaysia and Australia

Nation	*Location*	*N*	*No. of Errors Analysed*	*% of Total Errors in Category*					
				R	*C*	*T*	*PS*	*E*	*X*
Malaysia	Kuala Lumpur	31	210	0%	12%	62%	2%	3%	21%
	Penang	44	473	2%	26%	42%	5%	5%	19%
	Rural	70	780	4%	28%	42%	3%	5%	18%
TOTAL	(Malaysia)	145	1463	3%	25%	46%	3%	5%	18%
Australia	Perth (m/c)	28	143	1%	13%	28%	15%	10%	33%
	Perth (w/c)	33	350	1%	12%	51%	14%	1%	20%
TOTAL	(Australia)	61	493	1%	12%	44%	14%	4%	24%

Code: R = Reading; C = Comprehension; T = Transformation; PS = Process Skills; E = Encoding; X = Careless

Different Error Profiles for Different Students

Previous Newman error analysis researchers have often noted that different students can generate quite different error profiles, even though they respond to the same set of questions. This phenomenon was again evident in the study being reported in this paper. Thus, for example, one student in the Malaysian rural subsample made 9 Careless errors out of 17 errors, whereas three other students in the same subsample did not make a Careless error, even though they made 23, 24, and 24 errors. It would seem to be the case that, over the past two decades, Newman error analysis research has established that differences in error profiles are to be expected among students in all classrooms. No further consideration of the phenomenon will be given in this paper.

Different Error Patterns for Different Questions

Curiously, given the recognition that different students are likely to generate different error profiles, there has been little recognition of the possibility that different questions might generate different error patterns. Consider, for example, the different error patterns associated with Questions 13, 17, and 2 (see Table 2).

From Table 2, it can be concluded that the Australian students did not subtract as well as the Malaysian students. A similar result was found on all questions requiring straightforward applications of numerical skills. Yet the Australian Year 7 students had a higher mean score on the 24-question test than the Malaysian students. (The mean number of questions correct (out of a possible 24) for the 61 Australian students was 15.9, and the mean for the 145 Malaysian students was 13.8.)

Data for Question 2 (in Table 2) indicate that Australian students did better than their Malaysian counterparts on word problems which demanded careful analysis of the semantic structure. This pattern was generally evident in the overall analysis. These results are similar to those reported by Kaushil et al. (1985), who found that students in Delhi (India) did much better than Australian students of the same age on questions requiring straightforward applications of numerical skills, but much worse on word problems with non-trivial semantic structures.

Table 2 indicates that although two questions on the same test can generate similar error patterns, they can also generate quite different error patterns. For example, both Questions 13 and 17 generated a large proportion of Careless errors, in both their Bahasa Melayu and English forms. In Australia. but not so much in Malaysia, there were also many Process Skills errors.

By contrast, for Question 2, most of the errors made by both Malaysian and Australian students were in the Transformation category. Neither Australian nor Malaysian Year 7 students made Process Skills errors on Question 2.

Table 2
Errors on Three Questions by Year 7 Students in Malaysia and Australia

*Question**	*Nation*	*% Responses Which Were Errors*	*% of Total Errors in Category*					
			R	*C*	*T*	*PS*	*E*	*X*
13. Find the value of 940 – 586.	Malaysia (145 students)	13%	11%	16%	5%	16%	0%	53%
	Australia (61 students)	33%	15%	10%	0%	35%	0%	40%
17. 950 students, 587 present. How many absent?	Malaysia	12%	6%	6%	6%	6%	6%	71%
	Australia	33%	0%	0%	0%	45%	0%	55%
2. 20 boys & girls. Four more girls. How many boys?	Malaysia	90%	0%	21%	78%	0%	0%	1%
	Australia	59%	0%	6%	86%	0%	0%	9%

* Note: The wording given under "Question" in Table 2 provides a summary only of each question.

If one analyses the wording and structure of Questions 13, 17, and 2, the above results are hardly surprising. Most experienced teachers of mathematics would predict that with Year 7 students, Questions 13 and 17 would generate mainly Process Skills and/or Careless errors, and Question 2, mainly Transformation and/or Comprehension errors.

One might agree to label Questions 13 and 17 as "Careless" questions in the Malaysian education context, and "Careless & Process Skills" questions in the Australian education context. From a similar perspective, Question 2 might be labelled a "Transformation" question in both the Malaysian and Australian contexts. A convention might be developed that if more than 25% (say) of errors for a question are in the same category, then the name of that category would be included in the label for the question. If two (or even three) categories are such that each of the categories contained more than 25% of the errors for a question, then the category associated with the highest proportion of errors will be named first in the label for a question.

However, it is not always easy to predict which categories of errors will be associated with which questions. The reader is invited to consider the following two questions (which were Questions 20 and 23 among the 24 questions used in the study reported in this paper), and to predict which type(s) of errors would be most common if the questions were answered by (a) the Australian, and (b) the Malaysian Year 7 students in the study. Predictions can be checked against pertinent data from the Malaysian and Australian samples which are given in an Appendix to this paper (immediately before the Reference list). The English versions only of the questions are given below:

> *Question 20:* There are 12 apples on a table. If you picked up one-third of the apples, how many apples would remain on the table?
>
> *Question 23:* Bella takes exactly 3 hours to walk 15 km. How long would it take her (if she walks at the same speed) to walk 20 km?

Some Spearman Rank-order Correlations

In order to check the extent of agreement between relative difficulties of corresponding questions on the two forms of the test, the 24 questions were ranked in difficulty—one rank being based on the Malaysian sample, and the other on the Australian sample. A between-country Spearman rank-order correlation co-efficient was then calculated (Lumsden, 1969). The correlation obtained was a moderate 0.30.

However, when in-country rank-order correlations between the ranks for Kuala Lumpur-Penang, Kuala Lumpur-rural (Malaysia), Penang-rural (Malaysia), and Perth (m/c)-Perth (w/c) were calculated, these all ranged between 0.70 and 0.80. It would appear to be the case that there are strong in-country influences on the difficulty of mathematics questions. The extent to which these arise from linguistic differences or from approaches to teaching and learning mathematics needs to be further investigated.

Concluding Comments

Although Newman research has invariably appealed to teachers whose students have been involved in Newman studies because it generates important, authentic, and interesting data, the implications of this research for teachers and policy makers have not been clear. Can results of Newman studies based on data generated from interactions between English-speaking students and tests written in the English, be legitimately generalised to predict likely error patterns when non-English-speaking students take the "same" tests—with the tests *not* in English but in the students' own first languages?

From reports of previous Newman studies, and from the data presented in this present paper, it would appear to be the case that the answer to the question in the last paragraph is a qualified "Yes." Newman research data from a range of Asia-Pacific nations suggest that the most basic components of a mathematics education should be "learning to read, write, comprehend, and mathematise." In fact, the four operations on natural numbers seemed to have been well handled by most pupils in upper primary and secondary classes in all nations in which Newman research has been conducted thus far.

Findings from Newman research indicate that many lower secondary students in many Asia-Pacific nations have little understanding of, or ability to apply, percentages and decimal and vulgar fraction concepts. It also seems to be the case that many Australian students do not have their numerical skills as well developed as their age- and/or grade-counterparts in other Asia-Pacific nations. However, as data in this paper reveals, that does not imply that the Australian students are better (or worse) off so far as school mathematics is concerned. Many Australian students do better than corresponding students in other Asia-Pacific nations at comprehending and "mathematising."

Analyses presented in this paper indicate that not only are the interactions of particular questions with particular groups of students likely to generate predictable Newman error patterns, but also the patterns for one question are likely to be quite different from those for other questions. This would appear to be relevant to those involved in test development, or who are concerned with best-test design. In particular, it is likely that psychometric considerations (such as reliability, validity, and item discrimination) need to be complemented by findings of Newman research.

Findings from Newman research raise the difficult issue of what educators can do to improve learners' comprehension of mathematical text and their ability to transform (that is to say, to identify an appropriate sequence of operations which will solve a given word problem). At present, little progress has been made on this issue, and it should be an important focus of mathematics education researchers in Asia-Pacific nations over the next decade. At issue is the extent to which this is related to linguistic features, as opposed to teaching and learning patterns found in mathematics classrooms.

Appendix

Table 3 shows data for Questions 20 and 23. Entries in Table 3 suggest that Question 20 is a "Comprehension & Transformation" question for both Malaysian and Australian Year 7 students. Question 23 is a "Transformation" question for Australian Year 7 students (and a "Transformation & Comprehension" question for Malaysian Year 7 students).

Table 3
Errors on Two Questions by Year 7 Students in Malaysia and Australia

Question	*Nation*	*% of Responses Which Were Errors*	*% of Total Errors in Category*					
			R	*C*	*T*	*PS*	*E*	*X*
	Malaysia (145 students)	56%	1%	40%	44%	0%	0%	15%
20. 12 apples on a table; 1/3 are removed. How many remain?.								
	Australia (61 students)	26%	0%	56%	25%	0%	0%	19%
	Malaysia	46%	1%	37%	46%	5%	0%	10%
23. 3 hours to walk 15 km. How many hours for 20 km?								
	Australia	44%	0%	4%	89%	0%	0%	7%

References

Brislin, R. (1970). Back-translation for cross-cultural research. *Journal of Cross-cultural Psychology, 1*(3), 185–216.

Casey, D. P. (1978). Failing students: A strategy of error analysis. In P. Costello (Ed.), *Aspects of motivation* (pp. 295–306). Melbourne: Mathematical Association of Victoria.

Clarkson, P. C. (1980). The Newman error analysis—Some extensions. In B. A. Foster (Ed.), *Research in mathematics education in Australia 1980* (Vol. 1, pp. 11–22). Hobart: Mathematics Education Research Group of Australia.

Clarkson, P. C. (1983). Types of errors made by Papua New Guinean students. *Report No. 26*. Lae: Papua New Guinea University of Technology Mathematics Education Centre.

Clarkson, P. C. (1991). Language comprehension errors: A further investigation. *Mathematics Education Research Journal, 3* (2), 24–33.

Clements, M. A. (1980). Analysing children's errors on written mathematical tasks. *Educational Studies in Mathematics, 11*(1), 1-21.

Clements, M. A. (1982). Careless errors made by sixth-grade children on written mathematical tasks. *Journal for Research in Mathematics Education, 13*(2), 136–144.

Clements, M. A., & Ellerton, N. F. (1992). Overemphasising process skills in school mathematics: Newman analysis data from five countries. In W. Geeslin & K. Graham (Eds.), *Proceedings of the Sixteenth International Conference on the Psychology of Mathematics Education* (Vol. 1, pp. 145-152). Durham, New Hampshire: International Group for the Psychology of Mathematics Education.

Ellerton, N. F., & Clements, M. A. (1992). Implications of Newman research for the issue of "What is basic in school mathematics?" In B. Southwell, R. Perry, & K. Owens (Eds.), *Proceedings of the Fifteenth Annual Conference of the Mathematics*

Education Research Group of Australasia (pp. 276–284). Sydney: Mathematics Education Research Group of Australasia.

Faulkner, R. (1992). *Research on the number and type of calculation errors made by registered nurses in a major Melbourne teaching hospital.* Unpublished M.Ed. research paper.

Jiminez, E. C. (1992). *A cross-lingual study of Grade 3 and Grade 5 Filipino children's processing of mathematical word problems.* Penang: SEAMEO-RECSAM.

Kaur, B. (1995). A window to the problem solvers' difficulties. In A. Richards (Ed.), *Forging links and integrating resources* (pp. 228–234). Darwin: Australian Association of Mathematics Teachers.

Kaushil, L. D., Sajjin Singh, & Clements, M. A. (1985). *Language factors influencing the learning of mathematics in an English-medium school in Delhi.* Delhi: State Institute of Education (Roop Nagar).

Kownan, M. B. (1992). *An investigation of Malaysian Form 2 students' misconceptions of force and energy.* Unpublished manuscript, SEAMEO-RECSAM, Penang.

Lumsden, J. (1969). *Elementary statistical method.* Perth: University of Western Australia Press.

Marinas, B., & Clements, M. A. (1990). Understanding the problem: A prerequisite to problem solving in mathematics. *Journal for Research in Science and Mathematics Education in Southeast Asia, 13*(1), 14-20.

Mohidin, R. (1991). *An investigation into the difficulties faced by the students of Form 4 SMJA secondary school in transforming short mathematics problems into algebraic form.* Penang: SEAMEO-RECSAM.

Newman, M. A. (1977a). An analysis of sixth-grade pupils' errors on written mathematical tasks. In M. A. Clements & J. Foyster (Eds.), *Research in mathematics education in Australia, 1977* (Vol. 2, pp. 269–287). Melbourne: Swinburne College Press.

Newman, M. A. (1977b). An analysis of sixth-grade pupils' errors on written mathematical tasks. *Victorian Institute for Educational Research Bulletin, 39,* 31–43.

Newman, M. A. (1983). *Strategies for diagnosis and remediation.* Sydney: Harcourt, Brace Jovanovich.

Ora, M. (1992). *An investigation into whether senior secondary physical science students in Indonesia relate their practical work to their theoretical studies. Unpublished manuscript.* Penang: SEAMEO-RECSAM.

Riley, M. S., Greeno, J., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153-196). New York: Academic Press.

Singhatat, N. (1991). *Analysis of mathematics errors of lower secondary pupils in solving word problems.* Penang: SEAMEO-RECSAM.

Sobhachit, S. (1991). *An investigation into students' understanding of the electrochemical cell and the electrolytic cell.* Penang: SEAMEO-RECSAM.

Sulaiman, S., & Remorin, P. R. (Eds.). (1993). *Science-and mathematics-related project work abstracts of SEAMEO RECSAM participants* (April-June 1993). Penang: RECSAM.

Teoh Sooi Kim (1991). *An investigation into three aspects of numeracy among pupils studying in Year three and Year six in two primary schools in Malaysia.* Penang: SEAMEO-RECSAM.Tuck, S. (1983). *An investigation of the effectiveness of the Newman Language Mathematics Kit.* Unpublished M.Ed thesis, Monash University.

Thongtawat, N. (1992). *Comparing the effectiveness of multiple-choice and short-answer paper-and-pencil tests.* Penang (Malaysia): SEAMEO/RECSAM.

Watson, I. (1980). Investigating errors of beginning mathematicians. *Educational Studies in Mathematics, 11*(3), 319–329.