Primary School Mathematics: What Are We Testing?

Dr John Potter*

At the primary school level students are expected to learn maths terminology and develop competence in strategies superior to finger counting and guessing. They are also expected to apply their new found maths skills in real life situations - hence ‘word problems’.

But it is commonly found that primary school students find ‘word problems’ difficult. Is this because they do not see the relevance of mathematical processes in the problem or do not know how to apply them? Or is it because the language used in the question is opaque to them? The latter dysfunction has been shown to be important for students studying in a non-home language (e.g. Duminy, 1973, etc.) but the problem also exists when the home language is the language of instruction. This paper reports a study that tests the hypothesis that students’ marks maybe correlated with their mis-understanding of ‘word problem’ language, both technical terms and language of common use.

A secondary objective of the study was to evaluate the utility of a simple survey procedure by which teachers could rigorously test the utility of particular teaching practices.

Background

A review of mathematics text books shows that there is general agreement about the broad areas of study to which children aged between 9 and 12 years should be introduced. Most strongly focused are basic mathematical processes in relation to whole numbers and common and decimal fractions. Table 1 lists the traditional strategies appropriate to teaching these areas of competence. Powers, percentages and geometric parameters like perimeter and area are also taught but the main emphasis is arithmetic. It is expected that children entering secondary school will carry with them a ‘kitbag’ of arithmetic tools for application in advanced disciplines like algebra and physical science, but the empirical evidence suggests that the majority of students are not so equipped[1].

Various explanations have been given to explain why primary students do not develop competence in basic arithmetic processes. Theorists like Carraher et al (1991) have noted the inability of street vendors to translate street-learned arithmetic competence to class room maths literacy practice. Others like Siegel (1991) have pointed to

Subject	Process	Teaching Strategy
Whole numbers	+ and -	Integer position/value, columns
	x and ÷	Primes, Factorisation,
Multiples		Highest/Lowest Common Factors
Multiplication and Division		Rounding off
Decimal Fractions	+ and -	As for whole numbers
	x and ÷	Manipulation of the decimal Multiplication and Division
Common Fractions	+ and -	Equivalent fractions, Simplification, Common Denominator
	x and ÷	Numerators and Denominators Reciprocal Rule, Simplification
Inter-Relationships		Mixed Numbers to Improper Fractions and vice versa Common Fractions to Decimals and vice versa

TABLE 1: Arithmetic Strategies: Late Primary Schooling

.

difficulties, disparities and individual variableness in the process of extrapolating from domain general to domain specific knowledge. To such notions must be added the ever present possibility of poor teaching and the objective realities of students lives (e.g. McLaughlin & Talbert, 1990). And beyond that there is the de-motivational impact of students finding that they do not perform well in school exams. The experience of many twelve year old students is that seven years in school has sadly eroded the openness and desire for learning they manifested on their first day at school (Potter 2010). Schools are now seen as competitive sites in which they have learned that they are not capable of making the grade, especially in mathematics (Pateman, 1980).

Contrary to the above, the author has found that even a slight improvement in maths test performance will result in a quick return to a positive classroom motivation. By this view, any bland acceptance that students will inevitably not perform well in ‘word problem’ tests seems unacceptable. If it can be shown that the lack of competence in ‘word problem’ solving is due to the students’ lack of comprehension of mathematics terminology and strategies, then a teaching deficiency will have been uncovered. If, on the other hand, the problem is shown to drive from non-comprehension of ‘word problem’ language then a hindrance to traditional maths testing will have been identified.

The Survey Group

The survey was conducted at a well regarded South African primary school. The study population consisted of 89 Year 7 students. They aged from 11 years to 13 years. The language of instruction was English, and while the students derived from various backgrounds, the majority (82%) came from either wholly English speaking or fully bilingual homes where one of the languages was English (Table 2).

The students had been tested via a series of traditional tests over the three months prior to the survey. Their results showed the usual spread of performance (Table 3). Table 3 also shows the range of IQ measurements of these students as measured by the South Africa GSAT Test, but note that the data in column 2 have no direct link with the data in column 3; the parameters are listed in the same table simply for convenience.

Origin	Number	Home Language		Bilingual
Afrikaans	3	Afrikaans	2	1
English	56	English	56	
Indo-Asian	3	English	3	
Sotho	4	Sotho	3	1
Xhosa	2	Xhosa	2	
Zulu	5	Zulu	5	
USA	1	English	1	
Poland	1	English	1	
Portuguese	6	Portuguese	2	4
Greek	1			1
German	2			2
Italian	2			2
Zambia (Bemba)	1			1
Taiwan	1	Mandarin	1	

Table 2: Survey Group: Origins and Language Competency

Decile Range	Term Marks (%)	IQ (GSAT)
First (Top)	71-81	122-131
Second	67-70	114-121
Third	60-66	111-113
Fourth	56-59	108-110
Fifth	53-55	104-107
Sixth	50-52	102-103
Seventh	44-49	98-101
Eighth	39-43	92-97
Ninth	35-38	85-91
Tenth (Bottom)	25-34	68-84

Table 3: Range of Term Marks and IQ by Decile

Survey Methodology

Students were asked to fill in two questionnaires in normal class time. They were told that participation was non-compulsory but that their answers would help teachers see where they needed further instruction. Questionnaire 1 listed twenty five (25) maths terms drawn from the syllabus (Table 4, column 1). Seventeen (17) of the terms were descriptive nouns like ‘perimeter’, ‘sum’, ‘prime number’, etc. The other eight (8) terms were commands like ‘convert’, ‘express’ and ‘estimate’.

The students were asked to give definitions of the listed terms. They had been introduced to them many times in prior years so it was taken for granted that they should be familiar with them. As it happened, precise definitions of six (6) technical nouns had been given to the students during the three months immediately prior to the survey; the students had been asked to write these definitions in their work books and memorise them. So the survey provided an opportunity to test the utility of this rote learning teaching practice.

Questionnaire 2 looked at twenty seven (27) common use words extracted from previous year-end examinations set at the school (Table 5). The students were asked: “Can you say what these words mean?”

Results

Eighty usable results were obtained from the eighty
nine students surveyed (90% return). The mean percentage of students *attempting* answers in Questionnaire 1 was 69% and the mean number of *correct* answers given by those who attempted answers was 43% (Table 4). That is to say, with regard to maths terminology on average 3 out of 10 students gave no indication of what they thought the term meant and *of the rest* only 4 out of ten gave a correct answer. With regard to technical terms, on average 76% of the students attempted an answer and of these 39% (range 6-84%) were correct.

The results for the six terms specifically defined and memorized in class were higher than the results for the terms which had been taken for granted; 90% of students responded in the case of the defined terms with 70% correct answers as compared with a 68% response and 22% (range 8-36%) correct answers for the remainder of the terms.

The results for the eight mathematics commands was: attempted 62%, correct 52% (range 14-83%).

The mean number of students attempting answers in Questionnaire 2 was 75% (range 50-95%) of which 50% (range 12-75%) were correct (Table 5).

Term	Attempted %	Correct Answers %
A Perimeter	89	36
A Rectangle	95	11
An Operation Key	40	12
A Product*	92	65
An Hectare	56	6
A Sum*	100	84
A Number	77	35
A Key Sequence	55	33
A Factor*	86	68
A Thousandth	70	33
A Difference	96	84
Index Notation	81	8
A Square	92	11
A Multiple*	81	53
The Eighth Power	52	33
A Prime Number*	87	63
A Fractional Number	37	24
Convert	85	76
Express	55	43
Estimate	91	83
Simplify	84	79
Re-Write	79	76
In Decreasing Order	39	24
Evaluate	29	14
Factorise	37	20
Overall Mean	69	43
Mean (Tech. Terms)	76	39
Mean Def. Given*	90	70
Mean Unrehearsed	68	22
Mean Commands	62	52

Table 4: Survey Results: Questionnaire 1

Word	Attempted %	Correct %
Actual	88	49
Carefully	95	31
Study	70	12
Represent	84	51
A Salary	65	43
A Hammer	89	74
In a Line	85	55
Distance	74	54
Average Speed	80	36
A Touring Holiday	81	62
A Revolution	90	74
Covered	50	34
Allow	89	48
A Shaded Portion	79	70
Smallest	93	65
The Following	80	56
Different	76	57
Nearest	90	75
Equivalent	53	43
Short Hand	74	61
A Rule	69	63
A Bell Toll	55	48
Respectively	64	34
Together	64	45
Carried	50	29
Altogether	55	39
Mean %	75	50

Table 5: Survey Results; Questionnaire 2

Discussion

It was disturbing that 24% of students did not attempt to define the technical terms and 36% could not say what the eight commands might mean. There was no evidence during the survey or from the papers that the students had taken the task lightly. It is most likely that the students simply had no idea what the terms/commands meant. It must also be regarded as unsatisfactory that the mean for correct answers for technical mathematics terms was only 40%.

The marked improvement for terms deliberately defined and memorized suggests that there needs to be a shift away from any assumption that clear and precise understanding of mathematics terms will develop *en passant* in classroom conversations.

Surprisingly, concrete terms like ‘rectangle’, ‘square’, ‘hectare’ and ‘operational key’ performed below abstract arithmetic terms. This suggests that students may know what a particular geometric figure looks like but fail to identify its crucial properties. In the case of the ‘rectangle’, most students could say that the opposite sides were equal but only 11% mentioned the all important right angles. Perhaps this result is not surprising when it is considered that definition of a geometric figure involves a ‘formal operation’ (Piaget, 1964) and these students could be predicted to have not entered that phase of their development. The 36% correct response for ‘perimeter’ was disappointing, but perhaps not surprising to those who battle daily to help students grasp what appear to them to be simple notions.

It was also disappointing that only half of the students had any clear idea of what the common usage words in Questionnaire 2 meant, seeing that 82% of the students were from English speaking homes. Looking at the results more closely, it seems that 70% of the students had a good grasp of concrete words like ‘hammer’ and ‘revolution’ but only 30% could give an adequate definition for adverbs like ‘carried’ and ‘altogether’. Sadly, most students failed to give correct meanings for instructional words like ‘study’ or ‘carefully’. Perhaps less surprising, few could define an ‘average speed’; after all, this is a notion of some complexity normally introduced in the early secondary school years. Why it had been included in a primary school exam paper was not apparent.

Diagnosis

The following analysis looks at two hypotheses: (1) the ability of students to recall and define word meanings is correlated with their innate intelligence as defined by an Intelligence Test – in this case the South Africa GSAT; and (2) the test/exam performance of students is dependent on their comprehension of the meanings of mathematical terms and words of common usage included in mathematics papers.

With regard to hypothesis 1, students' term marks and their performance in the above Questionnaires was compared with their GSAT results on file in the school office. With regard to the second hypothesis, student performance in the survey was compared with their term marks.

The data were first presumed to be parametric and subjected to regression analysis; regression coefficients were calculated, the data graphed and a line of best fit determined[2]. The R-squared values in every case were below accepted levels of significance (Table 6), i.e. neither hypothesis was supported by this method of analysis.

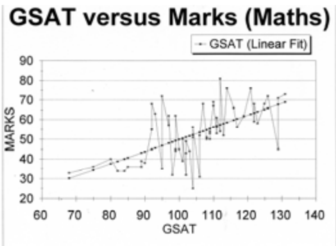

Figure 1: Line of Best Fit by Regression Analysis: GSAT v Mathematics Marks (Primary Students Aged 11-13)

Notwithstanding, the graphs appeared to show general trends[3] (e.g. see Figure 1) so the data were subjected to a non-parametric analysis. To facilitate this, in the case of the first hypothesis the data for GSAT were transformed to five classes and a bottom and top fifth of the population determined. Corresponding scores for Term Marks and Survey Results for each group were listed and the probability that the results from the top and bottom classes were drawn from the same population calculated using the Fisher Exact Probability Test (see Mizelditch, 1959). With regard to the second hypothesis the students' responses were ranked in accordance to their survey responses to Technical Terms, Commands and Words of Common Usage and their Term Marks listed. In each case the Term Marks for the top and bottom classes were compared using the same probability test. The results of these analyses are summarized in Table 7. They reveal significant levels of relationship in the case of the GSAT v Term Marks (1%), GSAT v Common Usage Words (5%), Technical Terms v Term Marks (1%), Commands v Term Marks (1%) and Common Usage Words v Term Marks (1%). That is to say, a non-parametric test supported both hypotheses.

Regression	**R-Squared Value**	**Significance[#]**
GSAT v Term Marks	0.39	N.S.
GSAT v Technical Terms	0.04	N.S.
GSAT v Maths Commands	0.08	N.S.
GSAT v Common-use Words	0.09	N.S.
Tech Terms v Term Marks	0.25	N.S.
Commands v Term Marks	0.15	N.S.
Common-use Words v Marks	0.23	N.S.

[#]The nearer the R-squared value is to 1, the better the correlation between the variable

Table 6: Showing the Degree of Correlation revealed by Regression Analysis between Intelligence (GSAT) and the Exam Performance and Ability to Define Mathematical Terms, Commands and Words of Common Usage demonstrated by Primary School Pupils aged 11-13 years

Comparison	**Probability**	**Significance[#]**
Top v Bottom Deciles	**Fisher Exact Test**	
GSAT v Term Marks	0.009	**
GSAT v Technical Terms	0.13	N.S.
GSAT v Maths Commands	0.15	N.S.
GSAT v Common-use Words	0.03	*
Tech Terms v Term Marks	0.004	**
Commands v Term Marks	0.004	**
Common-use Words v Marks	0.01	**

[#]Conventions of Significance:

- * Significant: 5% probability or less
- ** High Significance: 1% probability or less
- *** Very High Significance: 0.1% or less

Table 7: Relationships between Intelligence (as measure by the GSAT) and Term Marks and Comprehension of Mathematical Language; Primary School Students aged 11-13 years.

Conclusions

The significant correlation at the 1% level between GSAT and Term Marks and at the 5% level between GSAT and Word Meanings (Fisher Exact Test) will be encouraging for those who conduct such tests but the question remains (as with 'word problem' language) whether due recognition is given to the fact that the language of GSAT questions may disadvantage some participants.

The significant correlations between word meanings and Term Marks suggest that the failure of teachers to give specific attention to instruction in maths language is liable to produce a depreciated view of a student's abilities. That this may have a serious de-motivational effect is evident with consequences for subsequent teaching and learning.

The evidence that most students could readily reproduce dictionary meanings of words which had been deliberately taught and memorized is encouraging. This kind of pedagogic strategy is sometimes called into disrepute but the reality is that students who struggle to understand the language of a ‘word problem’ must be disadvantaged compared with those who have a good working knowledge of maths terms and commands. And this is not dissimilar to the case of students who have not had multiplication tables drilled into them are left guessing or counting fingers under the desk when faced with procedures like factorizing, multiplication and division. Simply put, these results reinforce the view that a little rote learning in relation to maths terminology should not be despised. On the contrary, it holds promise of being profoundly helpful.

With regard to the language of common usage, the results suggest that teachers need to recognise that they may be out of touch with their student’s range of life experience and this may lead them to include in ‘word problems’ situations with which the students are unfamiliar. By doing so they may defeat the purpose of the exercise. Avoiding ambiguity is another obvious necessity, although it has to be said that this is not always easy. In one case the author asked the class to consider what would happen when a 23mm long nail was driven through a piece of timber 17mm thick. To his surprise the whole class, even the ‘bright’ students, found the question difficult so he asked the students to draw a nail 17mm long. Nine girls drew a finger nail complete with polish! These are not matters to be taken lightly if we want to get genuine information from maths testing by ‘word problems’. More than that, in the long run, de-motivating students by setting them up to produce bad exam results has important implications for their economic future and their life-long perception of their innate abilities.

With regard to the second objective of the study, the development of a practical procedure for teacher self-evaluation, the study showed that a simple survey technique is capable of providing reliable data from a group of students. With regard to statistical analysis, regression analysis failed to deal adequately with the complexity of factors influencing student performance but data classification and the use of a non-parametric probability test found significant relationships between the parameters tested.

The results further suggest that teachers should be encouraged to conduct on-going rigorous testing of their teaching and assessment methods. Simple classroom surveys can provide data by which an enhanced *milieu* of teaching and learning may be created in the classroom.

Notes

[1] In South Africa an E symbol (40-50%) is sufficient for a student to advance to the next class, even for entry into Universities and Technical Training Colleges.

[2] This type of analysis is easily performed with modern computer soft ware.

[3] The disparate nature of the middle order data appeared to be critical in preventing the regression analysis demonstrating correlations; and this could be attributable to the complexity of factors which constitute student learning and performance.

References

Carraher, T. et al (1991): “Mathematics in the Streets and in Schools” in Light, R., Sheldon, S. and Woodhead, M. (eds) Learning to Think, Open University Press.

Duminy, P.A. (1973): African Students and Teaching Them, Van Schaik, Pretoria.

McLaughlan, M. & Talbert, J. (1990): “How the World of Teachers and Students Challenges Policy Coherence” in Fuhrman, S. (ed) Towards a Coherent Education Policy, San Francisco.

Mizelditch, M. Jnr (1959): A Basic Course in Sociological Statistics, Holt, Rinehart and Winston.

Pateman, T. (1980): “Can Schools Educate”, J.Phil.Educ., Vol.14, No.2.

Piaget, J. (1964): “Development and Learning” in Ripple, R.E. and Rockcastle, E. (eds) Piaget Rediscovered, Cornell University Press, Ithaca.

Potter, J.S. (2010): An Agency Crisis in Schools and Its Relationship with an Uncertain Ontology, Veritas (St Clements), Vol.1 No.3.

Siegler, R.S. (1991): How Domain-General and Domain Specific knowledge interact to produce strategy choices” in Light, R., Sheldon, S. and Woodhead, M. (eds) Learning to Think, Open University Press.

*Dr John Potter holds a professional teaching qualification in addition to degrees in Agricultural Science, Education Management, Education Philosophy and Theology. Throughout his 50+ years of professional life, Dr Potter has mainly been involved in adult education but in the 1990s he was involved in curriculum writing for early secondary schooling. The study reported in this article derives from his investigation into reasons why students commencing secondary education were not competent in mathematics. Dr Potter can be contacted on paracamp@senet.com.au.

.