

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Performance assessment tasks in the TIMSS study:
can we learn from them?**

A thesis presented in partial fulfilment of the requirements

for the degree of

Master of Educational Studies

in Mathematics at Massey University

Robyn Vivian Caygill

1998

ABSTRACT

Within the context of the reform of curricula in the education system, assessment methods and activities are also being reformed. There has been little research into the new methods and activities of assessment or of the impact these methods and activities will have on both the learning of students and the assessment of that learning. The International Association for the Evaluation of Educational Achievement (IEA) in its comparative study, the Third International Mathematics and Science Study (TIMSS), included some *hands-on* investigations, called performance assessment tasks, as some of the activities that assessed student learning. The student performances on two of the mathematics performance assessment tasks, *dice* and *packaging* were examined in this thesis, particularly in relation to student performances on some of the multiple-choice tasks also used in the study. In addition, the performances of some subgroups of the 207 standard three and 276 form three students who attempted each task were compared. The subgroupings were based on student responses to questions on gender, ethnicity, language of home, socio-economic status, and value of mathematics.

Many students were found to perform differently when their performances were compared in the multiple-choice and performance assessment questions that had similar content. Students were more likely to give no response to the performance assessment tasks than the multiple-choice tasks, particularly at the standard three level. For some, but not all, of the performance questions there was a smaller difference between the educationally disadvantaged subgroups of students and their peers, when compared with the differences between them on the multiple-choice tasks.

ACKNOWLEDGMENTS

Many people have contributed, in different ways, to the completion of this thesis. Firstly, I wish to thank the staff of the Comparative Education Research Unit within the Ministry of Education, particularly Steve May and Megan Chamberlain, for supplying the data files and answering my questions promptly. I also wish to thank the many friends who kept me sane, listened to me moaning, looked after Paul, and generally kept me going.

Thanks to Glenda Anthony who supervised and guided me through the challenges of the thesis work.

Finally, to Graham, my friend, husband, severest and kindest critic, thank you for all the accommodations and contributions you have made to a thesis which was *harder to give birth to than Paul*.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iii
1 INTRODUCTION	1
1.1 The Third International Mathematics and Science Study.....	2
1.2 What is assessment?.....	3
1.3 Assessment terms and activities.....	5
1.4 Aim of the study.....	6
1.5 Relevance of this research.....	7
2 ASSESSMENT IN MATHEMATICS EDUCATION	9
2.1 Historical review of assessment.....	9
2.2 The purposes of assessment.....	11
2.3 The impact of assessment on teaching and learning.....	13
2.4 Criticisms of assessments.....	16
2.5 Review of current philosophies of assessment in mathematics education.....	19
2.6 A good assessment.....	22
2.7 The trend towards the use of Performance Assessment tasks.....	27
2.8 Changes needed.....	30
2.9 Challenges faced in changing assessment.....	32
2.10 Technical issues arising with performance assessments.....	35
2.11 Performance in mathematics - assessment and the educationally disadvantaged.....	39
2.12 Summary.....	44
3. ASSESSMENT AND THE TIMSS STUDY	45
3.1. What is TIMSS?.....	45
3.2. Sampling.....	46
3.3. Administration.....	47
3.4. Tasks and Materials.....	51
3.5. Coding.....	53
3.6. Validity and Reliability.....	55
4. METHOD	57
4.1. Task Selection.....	57
4.2. Data Analysis.....	58

4.3. Selection of background variables.....	63
4.4. Limitations of this analysis.....	66
4.5. Summary.....	67
5 DICE.....	69
5.1 The dice task.....	69
5.2 The multiple-choice and short-answer tasks associated with the dice task.....	71
5.3 Face validity - how do these tasks compare with curriculum expectations?.....	74
5.4 Results of the dice task.....	77
5.5 Results of the multiple-choice tasks.....	85
5.6 Short-answer questions that cover similar mathematical content and their results.....	93
5.7 Comparison of the results of the dice task and the multiple-choice and short-answer tasks.	94
5.8 How do the sub-groups of students fare in the dice task and its associated tasks.	96
5.8.1 Gender differences.....	96
5.8.2 Ethnic differences.....	99
5.8.3 Comparisons between students from English-speaking homes and those from non-English-speaking homes.....	102
5.8.4 Differences between students from differing socio-economic status groups.	106
5.8.5 Personal value of mathematics.....	109
5.9 Summary.....	113
6 PACKAGING.....	115
6.1 The packaging task.....	115
6.2 The multiple-choice tasks associated with the packaging task.....	116
6.3 Face validity - how do the packaging task and the associated multiple-choice tasks compare with curriculum expectations?.....	118
6.4 Results of the packaging task.....	120
6.5 Results of the multiple-choice tasks.....	125
6.6 Comparison of the results of the packaging task and the multiple-choice tasks.....	127
6.7 How do sub-groups of students fare in the packaging task and its associated tasks.....	129
6.7.1 Gender differences.....	129
6.7.2 Ethnic differences.....	132
6.7.3 Differences found between students from English-speaking homes and those from non-English speaking homes.....	134
6.7.4 Differences found between students from differing socio-economic status groups.....	138
6.7.5 Personal Value of Mathematics.....	141
6.8 Summary.....	145

7 DISCUSSION	149
7.1 Appropriateness of tasks examined.....	149
7.2 Relationships between multiple-choice and performance questions.	149
7.3 Non-response to questions.....	150
7.4 Performances of disadvantaged sub-groups of students.....	151
7.4.1 Boys and Girls	152
7.4.2 Ethnic groups and language differences.....	154
7.4.3 Socio-economic status.....	155
7.4.4 Attitude to value of mathematics	155
7.5 Positives and negatives of the different assessment activities -performance assessment and pencil-and-paper.....	156
7.5.1 Test construction.....	156
7.5.2 Testing and Administration.....	157
7.5.3 Evaluation.....	159
7.6 Areas for future research.....	163
7.7 Concluding thoughts.....	164
 References	 165
 Appendix A	 A - 1
 Appendix B. Chi-test values for the dice task	 B - 1
 Appendix C. Chi-test values for the packaging task	 C - 1

LIST OF FIGURES AND TABLES

<i>Table 3.1 Assignment of Tasks to Stations</i>	48
<i>Table 3.2: Comparison of population one and two tasks for identical items.</i>	51
<i>Table 3.3: Summary of Expected Performances for the Mathematics Performance Assessment Tasks</i>	52
<i>Table 3.4: Typical code allocation for a question with maximum correctness score 2</i>	54
<i>Figure 4.1 Representation of the partial overlap between performance assessment tasks and the associated multiple-choice tasks</i>	59
<i>Table 4.1 Example of expected table where there is no difference between the scores on two tasks.</i>	60
<i>Table 4.2 An example of the expected table if there was no difference between groups</i>	61
<i>Table 4.3 Observed and expected proportions of girls and boys achieving each score on a hypothetical task</i>	62
<i>Table 4.4 Observed and expected proportions of girls and boys achieving a score of zero or more than zero on a hypothetical task</i>	63
<i>Table 5.1 Achievement objectives from the Mathematical Processes strand associated with the dice task</i>	75
<i>Table 5.2 Achievement objectives from the content strands associated with the dice task</i>	76
<i>Figure 5.1 Proportions of student responses to question one of the dice task.</i>	77
<i>Figure 5.2 Proportions of student responses to question two of the dice task</i>	79
<i>Figure 5.3 Proportions of student responses to question three of the dice task</i>	80
<i>Figure 5.4 Proportions of student responses to question four of the dice task</i>	81
<i>Figure 5.5 Proportions of student responses to question 5 part (a) of the dice task</i>	82
<i>Figure 5.6 Proportions of student responses to question 5 part (b) of the dice task</i>	83
<i>Figure 5.7 Total scores for the dice task</i>	85
<i>Figure 5.8 Proportions of student responses to question H8</i>	86
<i>Figure 5.9 Proportions of student responses to question J5</i>	87
<i>Figure 5.10 Proportions of student responses to question L13</i>	87
<i>Figure 5.11 Proportions of student responses to question K4</i>	88
<i>Figure 5.12 Proportions of student responses to question G1</i>	89
<i>Figure 5.13 Proportions of student responses to question L10</i>	90
<i>Figure 5.14 Proportions of student responses to question N18</i>	90
<i>Figure 5.15 Proportions of student responses to question M3</i>	91
<i>Figure 5.16 Proportions of student responses to U4</i>	93
<i>Figure 5.17 Proportions of standard three boys and girls who scored the maximum score</i>	97
<i>Figure 5.18 Proportions of form three boys and girls who scored the maximum score</i>	97
<i>Figure 5.19 Proportions of standard three girls and boys who did not respond to each question</i>	98
<i>Figure 5.20 Proportions of form three girls and boys who did not respond to each question</i>	98

Figure 5.21 Proportions of standard three students in the minority and majority ethnic groups who scored the maximum score.....	100
Figure 5.22 Proportions of form three students in the minority and majority groups who scored the maximum score.....	100
Figure 5.23 Proportions of standard three students in the minority and majority ethnic groups who did not respond to each question.....	101
Figure 5.24 Proportions of form three students in the minority and majority ethnic groups who did not respond to each question.....	101
Figure 5.25 Proportions of standard three students from English-speaking and non-English-speaking homes who scored the maximum score.	103
Figure 5.26 Proportions of form three students from English-speaking and non-English-speaking homes who scored the maximum score.	104
Figure 5.27 Proportions of standard three students from English-speaking and non-English-speaking homes who did not respond to each question.	105
Figure 5.28 Proportions of form three students from English-speaking and non-English-speaking homes who did not respond to each question.	105
Figure 5.29 Proportions of standard three students from each socio-economic group who scored the maximum score.....	107
Figure 5.30 Proportions of form three students from each socio-economic group who scored the maximum score.....	107
Figure 5.31 Proportions of standard three students from each socio-economic group who did not respond to each question.....	108
Figure 5.32 Proportions of form three students from each socio-economic group who did not respond to each question.	108
Figure 5.33 Proportions of students gaining the maximum score who agreed and disagreed with the statement I think it is important to do well in mathematics at school	110
Figure 5.34 Proportions of students gaining the maximum score who agreed and disagreed with the statement Mathematics is important to everyone's life.....	111
Figure 5.35 Proportions of students who agreed and disagreed with the statement I think it is important to do well in mathematics at school that did not respond to the packaging and multiple-choice questions.....	112
Figure 5.36 Proportions of students who agreed and disagreed with the statement Mathematics is important to everyone's life that did not respond to the packaging and multiple-choice questions.	112
Table 6.1 Achievement objectives from the Mathematical Processes strand associated with the packaging task.....	118
Table 6.2 Achievement objectives from the content strands associated with the packaging task.....	119
Figure 6.1. Proportions of student responses to question one of the packaging task.....	120
Figure 6.2 Proportions of student responses to question two of the packaging task.....	122

<i>Figure 6.3 Proportions of student responses to question three of the packaging task.</i>	123
<i>Figure 6.4 Graph of score totals for the packaging task.</i>	124
<i>Figure 6.5 Proportions of student responses to K3.</i>	125
<i>Figure 6.6 Proportions of responses to L5.</i>	126
<i>Figure 6.7 Proportions of responses to B11.</i>	127
<i>Figure 6.8 Proportions of girls and boys who scored the maximum score for each question.</i>	130
<i>Figure 6.9 Proportions of girls and boys who did not give any response to each question.</i>	131
<i>Figure 6.10 Proportions of students in the minority and majority ethnic groups who scored the maximum score for each question.</i>	133
<i>Figure 6.11 Proportions of the minority and majority ethnic groups who did not give any response to each question.</i>	134
<i>Figure 6.12 Proportions of students from English-speaking and non-English-speaking homes who scored the maximum score for each question.</i>	136
<i>Figure 6.13 Proportions of students from English-speaking and non-English-speaking homes who did not give any response to each question.</i>	137
<i>Figure 6.14 Proportions of students from each socio-economic group who scored the maximum score for each question.</i>	139
<i>Figure 6.15 Proportions of students from each socio-economic group who did not give any response to each question.</i>	140
<i>Figure 6.16 Proportions of students gaining the maximum score who agreed or disagreed with the statement I think it is important to do well in mathematics at school.</i>	142
<i>Figure 6.17 Proportions of students gaining the maximum score who agreed and disagreed with the statement Mathematics is important to everyone's life.</i>	143
<i>Figure 6.18 Proportions of students agreeing and disagreeing with the statement I think it is important to do well in mathematics at school who gave no response to each question.</i>	144
<i>Figure 6.19 Proportions of students agreeing and disagreeing with the statement Mathematics is important to everyone's life who gave no response to each question.</i>	144
<i>Table 7.1 Advantages and disadvantages of the use of performance assessment and multiple-choice questions.</i>	162