

1-1-1999

The measurement of classroom music learning using an outcomes framework in Western Australian schools

Beverley J. Pascoe
Edith Cowan University

Follow this and additional works at: <https://ro.ecu.edu.au/theses>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Pascoe, B. J. (1999). *The measurement of classroom music learning using an outcomes framework in Western Australian schools*. Edith Cowan University. Retrieved from <https://ro.ecu.edu.au/theses/1233>

This Thesis is posted at Research Online.
<https://ro.ecu.edu.au/theses/1233>

Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).
- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

**THE MEASUREMENT OF CLASSROOM MUSIC
LEARNING USING AN OUTCOMES FRAMEWORK
IN WESTERN AUSTRALIAN SCHOOLS**

by

**Beverley J Pascoe
Dip Teach, B Ed (Hons), M Ed**

**A Thesis submitted in partial fulfilment
of the requirements for the award of
Doctor of Philosophy
in the Faculty of Education, Edith Cowan University**

Date of Submission _____

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

DECLARATION

I certify that this thesis does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any institution of higher education; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

ACKNOWLEDGEMENTS

I wish to gratefully acknowledge the support, assistance and contributions of many people during the conceptualisation, research and writing of the thesis for this study.

Particular thanks and my sincere appreciation are extended to my supervisor, Dr Russell Waugh for his continued support, encouragement and the sharing of his skills and knowledge. My thanks are also extended to Associate Professor John Williamson and Dr Dawn Butterworth for their assistance.

The significant amount of work done on the data analysis by Educational Testing Centre is gratefully acknowledged. Particular thanks are extended to Annette Mercer and Terry Church for the many long and agonising hours we worked.

Thanks are extended to the Education Department of Western Australia for granting permission to access materials. I also wish to thank the many teachers and schools from metropolitan and country schools who piloted, trialed and marked the tests.

The support, patience and understanding extended to me by my dear husband Frank, as well as the encouragement from my wonderful mother and family during my years of study has been my inspiration and I sincerely thank them all.

ABSTRACT

This study addresses the problem of a lack of reliable and systematic methodology for evaluating progress in classroom music, using an outcomes framework, in primary and secondary schools in Western Australia. An innovative range of assessment tasks was developed for use at system, school or classroom level in order to develop a Music Achievement Scale, so that, in the future, meaningful reporting of student outcomes in music, in relation to *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), can occur. Music tasks were tested with a representative sample of 2191 government primary school students from Years 3 (946) and 7(921), and 324 secondary students from Year 10. The tests are designed to assess student progress in relation to outcome levels rather than for specific Year levels, thus making them useful for the collection of information on student achievement between Years 3, 7 and 10.

The Music Achievement Scale is composed of two parts: an Analysis test and a Process test. The Analysis test is designed to address the two 'appreciating' strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) and the Process test is designed to address the two 'expressing' strands of the statements. The Analysis test is a pencil and paper test in which individual students respond to excerpts of taped music. The Process test is a developmental process in which students respond to a stimulus by creating a short musical composition, which is developed, rehearsed and performed in a group situation, and which includes students' critical appraisal of their performance.

The tests consist of a combination of multiple choice and extended answer questions types and where possible, tasks are open-ended in order to provide the opportunity for students to perform to the maximum of their

abilities. Through the use of common items and common stimulus materials, tasks allow for the linking of items through Years 3, 7 and 10, thus providing valuable information on student progression through the outcome levels. The open-ended tasks are polychotomous, allowing for the partial credit of student responses, rather than being either 'right' or 'wrong.'

The tests were administered, in school classrooms that reflected students' usual learning environments, by their usual teachers of music. In primary schools this was sometimes a specialist music teacher and sometimes their usual classroom teacher. In secondary schools, the specialist music teacher administered the tests. All teachers used explicit administration instructions, which included time allocations to be apportioned for specific sections of the tests.

All tests were centrally marked by experienced specialist music teachers who underwent training in marking procedures that included processes of moderation. Markers used marking keys that addressed the complexities of open-ended and polychotomous items to allocate a raw score to each student on both the Analysis and the Process test. An extended logistic model of Rasch (Andrich, 1988a) through the use of the RUMM (Andrich, Sheridan & Luo, 1996) item analysis computer program, was employed to analyse the data. The Scale has good content validity and the tasks fit the measurement model, providing further evidence of validity. Reliability of the scale is high: the Person Separation Index is 0.900 and the Item Separation Index is 0.928. The Test-of-Fit Power is 'excellent,' indicating that a valid and reliable Scale of Music Achievement has been created.

Results indicate that the mean level for each year group shows a clear

pattern of student development in music appreciation and music expression. Around 80 per cent of year 3 students demonstrated skills associated with level 2 outcomes in classroom music, in excess of 55 per cent of year 7 students demonstrated skills associated with level 3 outcomes, and over 80 per cent of year 10 students demonstrated skills associated with level 4 outcomes.

To increase the awareness of teachers and Principals in the differential performances in music of sub-groups, a collection of data was undertaken on the performances of boys and girls, Aboriginal Torres Strait Islander (ATSI) students and non-Aboriginal students, and English speaking background and non-English speaking background (NESB) students. There were significant differences in the performances of some of these sub-groups that raise issues to be addressed in the future.

The study is of importance to Western Australian teachers and schools because, for the first time, specialist and generalist teachers will have access to reliable, authentic assessment materials that reflect exemplary classroom practice, as well as an instrument that allows for the mapping of student progress on a continuum of achievement related to the outcomes framework. Reporting to parents using the method of assessment developed in this study will provide more information on students' skills and abilities than in the past.

Issues related to the differential performances of sub-groups as well as issues of access and inclusivity, will be important at the system level for future developers of curriculum, as well as future developers of music assessment materials. Now that baseline data has been gathered and new methods pioneered, the way has been paved for future, improved methods of assessment in *the Arts*, and music in particular.

TABLE OF CONTENTS

	Page	
TITLE PAGE	I	
USE OF THESIS	II	
DECLARATION	III	
ACKNOWLEDGEMENTS	IV	
ABSTRACT	V	
TABLE OF CONTENTS	VIII	
LIST OF TABLES	XIV	
LIST OF FIGURES	XVI	
LIST OF APPENDICES	XVII	
CHAPTER 1	INTRODUCTION	1
	Background	1
	Problems with Arts measurement	3
	Significance	7
	Limitations	9
	Aims of the study	12
	Structure of the thesis	15
CHAPTER 2	LITERATURE REVIEW	18
	Accountability	18
	Assessment and evaluation in education	23
	Assessment in The Arts	25
	Music assessment in Western Australia	29
	Performance assessment	36

	Portfolio assessment	39
	Student self-assessment	45
	Computer-based assessment in The Arts	47
	Norm-based vs criterion-based assessment	51
	Rasch measurement in The Arts	54
	Ethnographic strategies	57
	Outcomes based education	59
	Summary	61
	Research questions	64
CHAPTER 3	THE MODEL AND THEORETICAL FRAMEWORK	65
	The model	69
	Expressing as part of the model	69
	Appreciating as part of the model	69
	Test design	71
	Development of the test items	74
	Structure of the Analysis test	75
	Structure of the Process test	81
CHAPTER 4	MEASUREMENT AND INSTRUMENTS	85
	Assessment	85
	Evaluation	86
	Item response theory	87
	Rasch models	89
	The Extended Logistic Model of Rasch	91
	The differences between raw scores and person ability measures	92
	The RUMM program	93
	Item fit	94

	Development of the Marking Keys	96
	Analysis Marking Key	96
	Process Marking Key	96
	Training the markers	98
	Marking the Music analysis tests	99
	Marking the Music Process tests	100
CHAPTER 5	SAMPLE, DATA COLLECTION AND PRELIMINARY ANALYSIS	
	ANALYSIS	102
	The samples for Years 3 and 7	102
	The sample for Year 10	104
	Data collection	108
	Process for student ability level estimates	110
	Student performance	113
	Patterns of development	118
	Conclusions	128
CHAPTER 6	DATA ANALYSIS (PART A): PSYCHOMETRIC CHARACTERISTICS AND MEANING OF THE SCALE	
	SCALE	131
	Introduction to the scale of music achievement	131
	General psychometric characteristics of the scale	134
	Fit of students to the model	134
	Fit of items to the model	136
	Validity of the measure of music achievement	145
	Threshold values	146
	Reversed thresholds	149
	Scale values	161
	Power of tests-of-fit	162

Increasing the probability	164
Establishment of levels	165
Transforming the logit values	166
Student performance levels	167
Implications	171
Conclusions	172

CHAPTER 7

DATA ANALYSIS (PART B):

SUB-GROUPS OF THE SCALE	174
Introduction	174
Sub-group analyses	178
Differential item performance	178
Year 3 sub-group analysis – males/females	179
Differential item performances Year 3 males and females	180
Year 3 sub-group analysis – Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students	180
Differential item performances of Year 3 Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students	181
Year 3 sub-group analysis – English speaking background students and non-English speaking background students	183
Differential performances of Year 3 English speaking background students and non-English speaking background students	184
Year 7 sub-group analysis of males and females	186
Differential item performances of Year 7 males and	

Females	186
Year 7 sub-group analysis – Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students	187
Differential item performances of Year 7 Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students	189
Year 7 sub-group analysis – English speaking background students and non-English speaking background students	190
Differential performances of Year 7 English speaking background students and non-English speaking background students	191
Year 10 sub-group analysis of males and females	193
Differential item performances of Year 10 males and Females	194
Year 10 sub-group analysis – Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students	195
Differential item performances of Year 10 Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students	196
Year 10 sub-group analysis – English speaking background students and non-English speaking background students	197
Differential performances of Year 10 English speaking background students and non-English speaking background students	198
Summary	200

CHAPTER 8	SUMMARY, CONCLUSIONS AND IMPLICATIONS	202
	Summary of the study	202
	Conclusions from the study	205
	Implications for teachers	210
	Implications for school administrators	212
	Implications for centrally administered educational systems	213
	Implications for theory	217
LIST OF REFERENCES		221

LIST OF TABLES

Table	Page
3.1 Analysis test item links and levels	79
3.2 Process test item links and levels	82
5.1 Arts enrolments for Year 10, semester one, 1996	105
5.2 Numbers of schools offering the range of arts disciplines	106
5.3 Arts enrolments by gender, semester one, 1996	107
5.4 Year 3 Frequency tables for Item MUA12	119
5.5 Year 7 Frequency tables for Item MUA12	119
5.6 Year 10 Frequency tables for Item MUA12	119
5.7 Year 3 Frequency tables for Item MUA08	120
5.8 Year 7 Frequency tables for Item MUA08	120
5.9 Year 10 Frequency tables for Item MUA08	120
5.10 Year 3 Frequency tables for Item MUA11	121
5.11 Year 7 Frequency tables for Item MUA11	122
5.12 Year 10 Frequency tables for Item MUA11	122
5.13 Year 3 Frequency tables for Item MUP01	123
5.14 Year 7 Frequency tables for Item MUP18	123
5.15 Year 10 Frequency tables for Item MUP18	124
5.16 Year 3 Frequency tables for Item MUP02	126
5.17 Year 7 Frequency tables for Item MUP19	126
5.18 Year 10 Frequency tables for Item MUP19	126
5.19 Year 3 Frequency tables for Item MUP11	127
5.20 Year 7 Frequency tables for Item MUP14	128
5.21 Year 10 Frequency tables for Item MUP16	129
6.1 Item parameters for Location and Scale	162
6.2 Test-of-fit summary statistics	163
6.3 Conversion of logit values to scale values	167
6.4 Summary of student performance in music	169

7.1	Number of schools offering arts disciplines at Year 10 in Semester one, 1996	176
7.2	Arts enrolments for Year 10, Semester one, 1996	176
7.3	Year 10 arts enrolments by gender, 1996	177
7.4	Percentages of Year 3 girls and boys attaining SOS levels	179
7.5	Percentages of Year 3 ATSI and non-ATSI students attaining SOS levels	181
7.6	Percentages of Year 3 ESB and NESB students attaining SOS levels	184
7.7	Summary of sub-group performances in music in Year 3	185
7.8	Percentages of Year 7 girls and boys attaining SOS levels	186
7.9	Percentages of Year 7 ATSI and non-ATSI students attaining SOS levels	188
7.10	Percentages of Year 7 ESB and NESB students attaining SOS levels	191
7.11	Summary of sub-group performances in music in Year 7	193
7.12	Percentages of Year 10 girls and boys attaining SOS levels	194
7.13	Percentages of Year 10 ATSI and non-ATSI students attaining SOS levels	196
7.14	Percentages of Year 10 ESB and NESB students attaining SOS levels	198
7.15	Summary of sub-group performances in music in Year 10	199

LIST OF FIGURES

Figure	Page
3.1 Model of music test design using an outcomes framework	73
5.1 Process for student ability level estimates	111
6.1 Distribution of student music achievement and item location estimates	132
6.2 Item characteristic curve for Item MA21	139
6.3 Item characteristic curve for Item MA18	140
6.4 Item characteristic curve for Item MA08	142
6.5 Item characteristic curve for Item MP25	143
6.6 Item characteristic curve for Item MP08	144
6.7 Item characteristic curve for Item MP19	145
6.8 Category probability curve for Item MA19	149
6.9 Category probability curve for Item MA06	151
6.10 Category probability curve for Item MA08	152
6.11 Category probability curve for Item MA18	153
6.12 Category probability curve for Item MA23	155
6.13 Category probability curve for Item MP02	157
6.14 Category probability curve for Item MP04	158
6.15 Category probability curve for Item MP05	159
6.16 Category probability curve for Item MP08	161
6.17 Student performance in music by level and mean score	170

LIST OF APPENDICES

	Page	
i	Student Outcome Statements Framework	235
ii	Year 3 Music Analysis test	239
iii	Year 3 Music Analysis Marking Key	242
iv	Year 7 Music Analysis Test	246
v	Year 7 Music Analysis Marking Key	252
vi	Year 10 Music Analysis Test	258
vii	Year 10 Music Analysis Marking Key	265
viii	Year 3 Process Test Administration Procedures	272
ix	Year 7 Process Test Administration Procedures	282
x	Year 10 Process Test Administration Procedures	291
xi	Music Process Marking Key	298
xii	Administration Procedures Year 10 Music Analysis	304
xiii	Administration Procedures Year 7 Music Analysis	308
xiv	Administration Procedures Year 3 Music Analysis	315
xv	Individual Item Fit statistics	321
xvi	Item Parameters combined Analysis and Process Tests	322
xvii	Histogram of item clusters	323
xviii	Frequencies of score ranges Year 3 girls/boys	324
xix	Frequencies of score ranges Year 3 ATSI/non-ATSI	325
xx	Frequencies of score ranges Year 3 ESB/NESB	326
xxi	Frequencies of score ranges Year 7 girls/boys	327
xxii	Frequencies of score ranges Year 7 ATSI/non-ATSI	328
xxiii	Frequencies of score ranges Year 7 ESB/NESB	329

xxiv	Frequencies of score ranges Year 10 girls/boys	330
xxv	Frequencies of score ranges Year 10 ATSI/non-ATSI	331
xxvi	Frequencies of score ranges Year 10 ESB/NESB	332
xxvii	The Expressing strands of <i>The Arts Student</i>	
	<i>Outcome Statements</i>	333
Xxviii	The Appreciating strands of <i>The Arts Student</i>	
	<i>Outcome Statements</i>	339

CHAPTER ONE

INTRODUCTION

Background

The need to gather information about the effectiveness of education in *The Arts* has been emphasised by the current push for accountability in education and recognition of *The Arts* as one of the eight compulsory learning areas in the Western Australian K-10 curriculum. The generic title, *The Arts*, subsumes the disciplines of dance, drama, media, music and the visual arts. In Western Australia, it is intended that, during the primary school years, students have the opportunity to experience several art forms and develop broadly-based achievements in each discipline with a view to specialisation in particular art forms at secondary school (Education Department of Western Australia 1994c, p.2). The present study, within a climate of educational accountability and a wider offering of *The Arts* in Western Australian schools, focuses on the measurement of achievement in one aspect of *The Arts* namely, music education.

The recognition of *The Arts* as one of the important learning areas in education systems, as evidenced in initiatives such as the British National Curriculum (1993), the American National Standards (Consortium of National Arts Education Associations, 1994), *The Arts - a curriculum profile for Australian Schools* (Curriculum Corporation, 1994b) and the Western Australian Student Outcome Statements (Education Department of Western Australia, 1996), is resulting in more positive attitudes to *The Arts* and wider recognition of its importance in the development of the whole person. Arts educators involved in the writing of the Australian and Western Australian documents received overwhelming support for the central role of *The Arts* in school curricula in response to the draft versions of the documents (Emery

1994, p.6) and this support and recognition of the importance of *The Arts*, together with an emphasis on accountability in schools, has led to an increased awareness of the necessity to evaluate student achievement in this area.

In this age of 'accountability', where the general perception is "you test what is important" and with the recognition of *The Arts* as one of the important learning areas, *The Arts* cannot be ignored when gathering information about what students know and what they are able to do. Feedback on student progress is important in any learning area, including *The Arts*, for several reasons, not the least being that data on the effectiveness of instruction will assist the teaching process and assist in motivating student learning, self criticism and evaluation. It will also inform parents, community, teachers and school decision-makers of progress towards meeting objectives and standards. At a system level, feedback on student progress is essential in informing governments and policymakers and assisting in driving curriculum initiatives.

Although teachers regularly use methods of observation, checklists and anecdotal records within the classroom, the most common form of formal assessment used to establish levels, or compare students with the rest of the population, is test data. This is a formal gathering of information involving a structured situation in which performance is assessed under standard conditions. This form of assessment is usually a requirement of entry into special educational courses or tertiary institutions and successful achievement in formal assessment is often a requirement of employers (Griffin 1991, p.13). In learning areas that have been regarded as the 'core' subjects such as mathematics and English, schools regularly use this type of formal testing to establish student grades or levels and, indeed in the area of music, formal

testing of performance is commonplace. This testing in Western Australian schools, however, has been confined to the playing of a number of pieces and identifying students' knowledge of the musical elements, such as rhythm, melody, harmony, texture, and notation, and there has been no obvious attempt to gather information on students' creativity skills or their knowledge in the areas of aesthetics, criticism, or past and present contexts. The absence of an attempt to assess these skills is probably due to the difficulties involved in designing assessment instruments in these areas which, by their very nature must be subjective, and to the difficulties in reaching consensus as to how levels can be identified. *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) have now provided a framework for levels of development in *The Arts* and, consequently, the opportunity to develop some methods of assessment in music learning is taken up in this study.

For the purposes of this study, the term 'assessment' refers to the overall process of making analytical judgements, the term 'evaluation' refers to the process of determining the extent to which individuals or groups possess certain skills, knowledge or abilities, the term 'measurement' refers to the collecting of quantitative information and the term 'test' refers to the use of a series of questions or activities to measure the skills, knowledge or abilities of individuals or groups (Lehman, 1996, p.1).

Problems with Arts measurement

The Arts have often been regarded by educators as a frill or extracurricular area and not as a 'real' subject (Gordon, 1992, p.24; Jorgensen, 1994, p.26; Carlton, 1987, p.45; Kemp & Freeman, 1988, p.21; Lehman, 1996, p.6) and primary school reports to parents traditionally have

placed undue emphasis on non-curricular factors, rather than on the skills and abilities of students. Typical teachers' comments related to the child's achievement in music are "participates enthusiastically" or "enjoys music" or "attends practice regularly." Whilst these are probably important traits, they do not indicate the child's expertise or development in music. The notion that music is a frill subject has obviously been accepted by parents. One wonders what their reactions would be if "participates enthusiastically" was the only information they were given regarding their child's progress in mathematics! The practice of placing undue emphasis on non-curricular factors when reporting on music may be seen by educators in other disciplines as evidence that music lacks curricular substance (Lehman, 1996, p.6), thus perpetuating the notion that it is not a 'real' subject. With the inclusion of *The Arts* as one of eight compulsory learning areas in Western Australian schools, this attitude must change and assessment in Arts achievement will be important in the gathering of data for schools' management information systems, as well as for reporting to parents.

Eight compulsory learning areas were identified in *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) which was the first step in directing education in Western Australia to focus on student outcomes rather than teacher input. The framework format of the Curriculum Corporation (1994a), in which the school curriculum was divided into eight compulsory learning areas during the compulsory years of schooling from year 1 to year 10, was adapted. The statements provide a framework for curriculum development, describing the sequence for developing skills and knowledge and the intention is to enable the teacher to identify the achievements of students, focus on needs and provide further opportunity for student learning, growth and development. Further explanation of *The Arts*

Student Outcome Statements (Education Department of Western Australia, 1996) is given in chapter three and the framework is given in appendix i.

The original working edition of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1994c), describes the outcomes as being a "focus for school development planning" and as providing a "basis for teachers and schools to monitor and account for their performance" (1994c, p.5). The statements reflect artistic experience as a unified whole for the student. There are four strands containing identical statements for the five Arts disciplines of dance, drama, media, music and visual arts. The progression of student achievement is described from level one to level eight, and the indicators to describe what students actually do at each level are specific descriptions of achievement reflecting each discipline.

Identifying achievements of students is comparatively simple in subject areas where, traditionally, testing has been an ongoing part of the everyday classroom and where an answer can be right or wrong. However, this is not the case in *The Arts*, and assessment in such things as critical thinking skills and aesthetics is difficult because of the apparent personal and subjective nature of *The Arts*. It involves the assessment of tasks that measure imagination and creativity and it involves the development of measuring keys to score such things as musical compositions, dance improvisations, pieces of art or dramatic improvisations. Achieving reliability in assessment of those tasks will be a challenge. Clear frameworks need to be developed along with marking keys to define and describe precisely what evidence is sought to demonstrate achievement of standards. The introduction of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) has now provided the opportunity for the development of a framework on which to base the achievement of standards.

The present study addresses the problem of a lack of reliable and systematic methodology for evaluating progress in the discipline of music in schools. It attempts to do this by developing an innovative range of authentic assessment tasks appropriate for use at system, school or classroom level so that meaningful reporting of student outcomes in music can occur. For the purposes of this study, the term 'authentic' describes assessment tasks that reflect exemplary classroom practice. The assessment tasks reflect good teaching and assessment practice in classroom music and the skills and understandings identified in the authentic tasks are placed onto a continuum of students' skills which is matched to a standards framework based on *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996).

Achievement tasks were developed for students in Year 3 (aged 8), Year 7 (aged 12) and Year 10 (aged 15). The reason for selecting these three levels is that they represent three significant stages of students' compulsory schooling - the conclusion of junior primary school, the conclusion of primary school and the conclusion of the compulsory years of education. Themes and stimulus material were linked across year levels to provide a continuum of achievement so that it was possible, subsequently, to use the assessment materials for students between these levels.

The knowledge, skills and abilities of students in the discipline of music were measured using the Extended Logistic Model of Rasch (Andrich 1988a, 1988b), which was derived from the Rasch measurement model of analysis (Rasch, 1960/80; Andrich, 1988a, 1988b; Wright, 1995). Student raw scores were transformed into ability estimates and these, together with item difficulties, were calibrated onto a common scale of music achievement. The Rumm (Andrich, Luo & Sheridan, 1996) computer program was used to

undertake the analysis. The music achievement scale was matched to outcome levels from *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) and outcome levels from the data were established.

Significance

The study adds to knowledge in four ways. First, it tests a theoretical model of standards based on *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), as it is applied to music learning. The model has been trialled in Western Australia over the past two years and is due to be operational in Western Australian government schools from 1998 onwards. The model has not been analysed using a Rasch measurement model before and this study will provide the first test of the model of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996).

Second, the study adds to knowledge of measurement of standards in music learning. It will be of importance to teachers in Western Australia, as the assessment methods and instruments developed will mean that specialist and generalist teachers in Western Australia will have access to reliable, authentic assessment material reflecting exemplary classroom practice. It will not only provide teachers with a useful set of instruments with which to measure student progress in music but it will also provide them with authentic models on which to base future assessable classroom activities. It will significantly contribute to teacher knowledge in music education and to the use by teachers of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) to measure progress because there

are no current standardised benchmarks of student achievement in music at government schools in Western Australia.

Teachers engaging in classroom music programs will be able to use the material in four ways. First they will be able to map activities to the outcome statement levels to provide clear examples of requirements at that level, and while all music teachers will find this useful, examples are needed, particularly by generalist teachers. Second, teachers, both specialist and generalist will be able to identify activities which can be matched to specific strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). Third, they could access examples of activities demonstrating the aesthetically-oriented strands of *Responding, reflecting and evaluating* and *Understanding the role of the Arts in society*, which are currently unavailable. Fourth, they could link items across different levels of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). For instance, the marking keys will provide opportunities to measure open-ended responses at different levels on the continuum, providing links from one level to the next. At present there are no syllabus documents in Western Australia that provide any of this information to either specialist or generalist teachers in music education.

The third way in which this study will add to knowledge is by helping educational administrators in gathering whole-school information in *The Arts*. School administrators are obliged to develop a Management Information System in their school which provides whole-school data in each of the eight learning areas for reporting to the District Superintendent and for planning priorities and future teaching programs. This study will provide data enabling them to gather reliable material in music achievement that can be interpreted and linked to achievement in other aspects of *The Arts*.

The fourth way in which this study will add to knowledge is within the broader field of arts assessment worldwide. It makes a major contribution to the literature in this field and the innovative materials and methods developed will be of use to others who are developing models to design standards-based assessments in other content areas as well as *The Arts*.

Limitations

There are six limitations to this study. These are associated with the sample and its generalisability, restrictions of data to the subject of music, the financial and logistical constraints of large-scale testing and consistency of marker judgements of open-ended achievement tasks.

The first limitation refers to the population of Year 3, Year 7 and Year 10 students to whom the tasks were administered. The students were drawn from government schools only. No students from private schools or independent schools were tested. Hence, strictly, the results of this study are only representative of Year 3, Year 7 and Year 10 students in government schools in Western Australia. Because of the nature of group activities, whole classes were tested and only one Year 3 or Year 7 class was tested in each school. Schools and classes were drawn randomly from all Western Australian metropolitan and country primary schools with a minimum Year 3 or Year 7 population of six. In secondary schools, the whole class samples consisted of students who were currently studying music and were drawn randomly from all secondary schools in Western Australia which offered music at Year 10 and which had a minimum population of six in the music class.

The second limitation refers to the tests designed to test students' abilities and performances in music only. Therefore, although *The Arts*

Student Outcome Statements (Education Department of Western Australia, 1996) outline generic levels across the five disciplines of dance, drama, media, music and visual arts, it is not possible to generalise about levels in disciplines other than music. That is, if students are reported as having achieved a level three in music, it is not possible to assume that they have achieved a level three in drama or dance, for example.

The third limitation refers to the period of time available within the testing situation. Only the knowledge, skills and understandings that could be tested within the specified testing time periods could be included. Although all students completed the tests within the same time allocation, it is possible they may have achieved a higher standard if given more time. To be validly assessed, some knowledge, skills and understandings of music would require students to engage in sustained activity over an extended period of time.

The fourth limitation refers to the fact that only the knowledge, skills and understandings that were amenable to testing in a classroom setting with a generalist teacher could be tested. The assessments were not designed to be individually administered by specialist music teachers because every school does not have a music specialist or the resources to test students individually. This limited the range of music skills that could be assessed.

The fifth limitation to the study refers to the fact that, due to financial and logistical constraints, the assessment instrument made use of only a limited number of stimulus materials. Although it would have been possible to

conduct the assessment using the same set of test items but with other stimulus materials, some choices had to be made. It is possible that, if students had been given different stimuli, the difficulty level of the task may have been altered.

The sixth limitation refers to the consistency of marker judgements in contexts different from the standardised procedures used in this study. The student responses to the open-ended tasks had to be marked consistently. Markers were given one day's training and moderation so that they were able to establish and maintain consistent standards. Markers then took the student responded tasks away for marking. Spot checks were made on the marks and where discrepancies were found, these were re-marked, although double marking of papers was not viable. The standards, scales and profiles created in this study are only valid where teachers use the same marking standards. Further explanation of the marker training procedures is given in chapter 4. Further descriptions of the marking standards are given in chapters 6 and 7.

While it is appreciated that the two proposed assessment forms are limited in their content, it should be recognised that, like most testing tasks they are illustrative and not exhaustive. They will form appropriate benchmarks on which teachers will be able to base future assessments, as well as demonstrating appropriate methodologies to generalist classroom teachers who may have limited or no knowledge of methods for assessment in classroom arts subjects.

The year levels selected for testing in this study are limited to Years 3, 7 and 10 and benchmarks will be established in these year levels. However, it should provide opportunity for linking across year levels and the possibility,

subsequently, of developing assessment materials for students between these levels using adaptations of the proposed materials.

It is appreciated that there may be limitations in the abilities of the cohort to be tested. However, by testing in forty classes (around 1000 students) in each of the Years 3 and 7 and in twenty classes (approximately 400 students) in Year 10 in both urban and rural Western Australia, a representative sample will be obtained.

Aims of the study

The aims of this study are to:

1. Develop a Music Achievement Scale comprising both the appreciation and expression of music appropriate for each of Year 3 (8 year olds), Year 7 (12 year olds) and Year 10 (15 year olds) in Western Australia;
2. Show patterns of development from Year 3 through Year 7 to Year 10 by including common or 'link' items in the tests;
3. Trial the music assessment instruments and generate marking keys based on data gathered at Western Australian schools;
4. Mark the tests and analyse the data using the Extended Logistic Model of Rasch to create interval level measurements for the instruments;
5. Match the music achievement scale to outcome level statements and determine level cut-off points;
6. Analyse the data to provide state means for Year 3, Year 7 and Year 10 to provide teachers with comparisons of student performance;
7. Analyse the data to provide comparative information on the performance of sub-groups; and
6. Develop student profiles to provide teachers with descriptions of performance.

The first assessment instrument consists of a set of stimulus material to which students respond, primarily in relation to the 'appreciating' strands of *Responding, Reflecting and Evaluating* and *Understanding the role of the Arts in Society*. Students produce responses in relation to aesthetics, critical analysis, interpretation of meaning and music concepts, such as beat, rhythm, melody, dynamics, shape, mood and tension. Developmental processes involved comparisons and contrasts and the exploration of critical and contextual understanding focusing on particular periods of music history. Where possible, tasks were open-ended in order to provide students with the opportunity to demonstrate their maximum levels of ability. The analysis task was designed to cover a time duration of approximately one lesson period at the appropriate level; that is, approximately 45 minutes at Year 3, 50 minutes at Year 7 and 60 minutes at Year 10.

A multi-media CD Rom version of the Year 3 Analysis task was designed by the candidate and developed in consultation with a teacher colleague, who is not only an experienced Year 3 teacher but who is also a producer of educational computer software. The CD Rom was developed in an attempt to determine whether the limited literacy skills of Year 3 students, as well as the limitations involved in whole classroom access to stimulus materials, have an effect on students' results. The CD Rom includes visual material in high quality colour, sound digitised for music and a capacity for moving images.

The CD Rom interface was designed so that students could complete tests at the screen, on an individual basis, thus allowing them the opportunity to listen to and view stimulus materials, as often as necessary, as well as having the questions read aloud, as often as necessary. Student responses were entered to a computer and, at the end of testing the whole class, the data were saved on a disk by the teacher, thus eliminating the need for large

quantities of paperwork. There is already a high penetration of CD-Rom in schools through school libraries and a proposal such as this may assist in increasing efficiency in the collection of data for future assessment. A small scale study involving approximately 120 Year 3 students will be conducted at the conclusion of this study using the CD Rom version vs the hard copy version of the task.

The second assessment instrument offers a broad view of student abilities through their documentation of the steps in the learning process which lead to the performance of their final products. The process addresses *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) 'expressing' strands of *Creating, exploring and developing ideas* and *Using skills, techniques, technologies*, and provides evidence of students' planning processes towards a simple composition and performance. The activities in which students engage provide opportunity for inquiry and the use of Arts language which are fundamental elements in the creative process leading to the development of worthwhile art. These activities will provide direct evidence of the students' skills and learning, as well as concrete evidence for evaluation using marking keys that will be developed during trials. An important feature of this instrument is the opportunity for students' reflection and self-appraisal of their work. The process assessment task is designed to cover a time duration of approximately two lesson periods at the appropriate level and is based on a clearly structured framework beginning with an appropriate stimulus and culminating in the performance of the composition.

The framework of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) provides a series of descriptions of standards against which performance can be gauged. Test items are a set of

developmental indicators of achievement that are mapped against the skills and abilities described at each level of the outcome statements. For purposes of reporting, descriptions of typical understandings which can be expected at each level are calibrated onto a measurement scale - the higher the calibration, the more difficult the item. Student levels of achievement are simultaneously calibrated on to the same scale and mapped as an arbitrary numerical scale which is organised at equal levels along the continuum, thus facilitating reporting of student performance data.

Literacy competence is not a criterion and hence, spelling, grammar and sentence-construction are not assessed. The criteria for evaluation is emphasised during item writing and the design of marking keys, and students are made aware of these criteria during testing. Literacy levels are kept at an understanding appropriate to the year level. Ongoing consultation with classroom practitioners was undertaken to refine items and language for the relevant year levels.

Structure of the Thesis

The main purpose of this study is to develop two music assessment instruments to assess aesthetics, composition and performance for Year 3, Year 7 and Year 10 students in Western Australian schools. The items of the instruments include open-ended tasks and teacher judgements of performance. Item difficulties and student performance measures are calibrated on the same scale using the Extended Logistic Model of Rasch (Andrich, 1988a, 1988b, 1978). Another purpose of the study is to establish state means that can be used as benchmarks at the Year levels 3, 7 and 10 so that non-specialist classroom teachers and music specialist teachers can measure students' music abilities against the achievements of other Western

Australian students. The thesis describing how this was done is set out in eight chapters. The first chapter sets the scene for the study and provides an introduction.

Chapter two describes a review of the relevant literature, including issues of accountability in education and the need for teachers and schools to be accountable to governments, at both state and federal levels in Australia, as well as to parents and the community. Different methods of assessment used in education and the changes in attitudes related to the different methodologies are discussed. Methods of assessment used specifically in Arts education, both in Australia and overseas, are reviewed. This chapter also provides a description of outcomes based education and the shift in curriculum policy and practice from an emphasis on what teachers have taught to an emphasis on what students have learned, as applied in Western Australia and overseas in recent times.

Chapter three describes the design of this research. It explains the model of arts measurement used, including the skills and knowledge students are expected to display at each level. It also describes the process used in the derivation of the tests and the generation of the scores.

Chapter four provides a description of the research methodology, including a description of measurement and the purposes of assessment of student achievement, together with a review of norm-referenced vs criterion-referenced testing and current attitudes about their benefits to the teaching and learning process. Item Response theory and both the Simple Logistic Model (Rasch, 1980) and the Extended Logistic Model of Rasch (Andrich, 1978, 1988a, 1988b), as well as item fit, are described. Explanations of the development of the test items, the marking keys and training of the markers

are included to emphasise the importance placed on the validity of the tasks developed and the reliability of the marker judgements.

Methods used to select the sample are detailed in chapter five, together with details of the characteristics of the sample. This includes constraints involved in the selection of the sample at Year 10 where music is an option and where there is an imbalance between numbers of girls and boys. Procedures used in administering the tests and collection of the data as well as a preliminary qualitative analysis are also included.

A psychometric analysis of the data which includes a discussion of the validity and reliability of the measures, analysis of scores using a Rasch model, ability estimates and fit, and student level cut-off estimates is provided in chapter six.

Chapter seven contains the analyses of the data for the measurement of the music learning. The difficulty of the items, the measurement of student ability and the interpretation of the measures in relation to *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) levels is presented together with a summary of the data analysis. The means for student standards in music, their derivation, and the way they can be used by teachers as benchmarks are explained.

Chapter eight comprises a summary and conclusions drawn from the whole study as well as implications for teachers, implications for the theory of music learning in the classroom and implications for administrators.

CHAPTER TWO

REVIEW OF LITERATURE

Much has been written about school assessment and the various methodologies used both at a school level and at a system level (Wright, 1994, LeMahieu, 1995, Worthen, 1993, Linn, Baker & Dunbar, 1991, Marzano, 1994, Gardner, 1996, Leonhard, 1990, Roberts, 1994, McGuire, 1983). This review outlines some of these methodologies and their relevance to *The Arts*. Although some Arts educators expound the virtues of comprehensive arts programs in schools and the necessity for the establishment of 'standards' and reliable assessment methodologies in *The Arts* (Ross, Radnor, Mitchell & Bierton, 1993, p.9; Straus, 1992, p.x; Cancel, 1992, p.xv; Armstrong, 1994, p.9; Mitchell, 1992, p.3; Gurin, 1994, p.85; Down, 1994, p.3), it seems that little has been done to develop those methodologies.

The issue of accountability and the impact of these methodologies on accountability in schools is discussed and a brief examination of standards, together with a brief history of teaching and assessment practices in music in Western Australian schools and the current status of student outcome statements in this state is made. A brief report of journal articles relating to methodologies of assessment and evaluation in education generally and *The Arts* in particular, in Britain, Canada, the United States of America and Australia is also given.

Accountability

Since the mid-1980s in Australia, both state and federal governments have demanded increased accountability in education both at school and

system levels (Bates, 1992a, p.15; Nierman, 1985, p.20; Carlin, 1994, p.29; Beatty, 1992, p.35; Down & Mahlmann, 1994, p.xiii, Duke, 1994, p.15; Boston, 1994, p.2). Results of system level assessment programs reported to parents, government and community during this time have resulted in the dissatisfaction of parents, politicians and lay people with students' literacy and numeracy skills (Ministry of Education and Training, 1993; Ministry of Education, 1994a; Education Department of Western Australia, 1994; West Australian Newspapers, 1996) and, as the costs of education escalate, it is likely that it will be subjected to even greater scrutiny by the community (Nierman 1985, p.20; Griffin 1991, p.1).

The tendency towards more autonomy and responsibility at the school level and the fact that principals, staff and the community in Western Australia are gaining more control over the running of schools, is resulting in greater public interest in the quality and effectiveness of schools. This, in turn, provides more focus on the methods used to evaluate and monitor school performances and, with an increasing level of cooperation by the States at a national level, it is acknowledged that accountability is firmly on the national education agenda (Masters, 1992, p.i; Ebel & Frisbie, 1986, p.1; Reid, 1992, p.55).

The Education Department of Western Australia has taken steps towards addressing the demand for accountability with its Monitoring Standards in Education program which randomly samples students at Years 3, 7 and 10 in two of the eight learning areas each year, in order to provide a snapshot of standards of performance in government schools. This helps to keep the general public informed about standards and schools also benefit from the provision of standards to compare their students' performances and to identify specific areas of need for future teaching and learning programs.

The need to keep the public informed brings about increased responsibility for principals, teachers and local school communities in the management of schools and curriculum and the shift in emphasis from an accountability model, based on school inputs, to a model based on an ongoing monitoring of outcomes in the form of "enhanced student learning" (Masters, 1992, p.56). In Western Australia this means teachers are committed to monitoring student outcomes. In discussing achievement standards in schools, Masters (1991, p.3) comments on the progress being made in other countries such as the United States of America, England and Wales, as well as Australia, in developing improved measures of student performance to provide for better monitoring of outcomes at all levels, including systems, schools, classrooms and individual students. The proposals for the California Assessment Program closely parallel those in England and Wales, as well as in other states of America (Masters, 1991, p.3), in that they feature better methods of assessing and reporting on achievements of individual students, the development of standards for system-wide assessment and reporting, a broad range of assessment methods and a significantly enhanced role for teacher observation and judgement in relation to the tasks.

It is important to recognise that Australia is already at the forefront in relation to the introduction of new approaches to the measurement of student achievement. These include the Western Australian Monitoring Standards in Education Program, the Victorian Achievement Studies, the New South Wales Basic Skills Testing Program and the Queensland Assessment of Student Performance. (Masters, 1991, p.3). Australia, England and Wales closely parallel each other in that they have all developed systems of student performance measures linked to the attainment targets of their National Curriculums, based on student outcomes rather than teacher input. Their

objectives for developing improved measures of student performance include the provision of information to policy-makers, as a basis for more informed decision making at all levels of schooling, from system manager, to principal, to teachers, and parents (Masters, 1991, p.3; Lehman, 1996, p.4).

The Education Department of Western Australia recognises that, if it has a set of standards for music achievement in the form of *The Arts Student Outcome Statements* (1996), then logically, ways must be found to determine whether these standards are being met or whether the outcomes are being achieved (Music Educators National Conference Committee on Performance Standards, 1994, p.1). The methods used to reflect exemplary classroom practice need to provide models to classroom practitioners and the Education Department of Western Australia is achieving this by including authentic methods of assessment such as group activities, hands-on activities, speaking and listening tasks and student interviews in system-level testing. They also provide assessment materials, including marking keys and student profiles for use in schools so that comparisons with system-level results can be made.

Many researchers and educators suggest that we should shy away from the tendency to mass-test students using standardised tests that are machine scoreable and require a minimum of expenditure as this type of testing has created a wariness among some educators and community members about the value of any kind of assessment for accountability in our schools. Cohen (1989), Darling-Hammond (1994), Dwyer (1989), and Shrubb (1989) are among those who warn us about the dangers of mass testing and any form of assessment that tends to narrow the curriculum to what can be measured. Cohen's (1989) opinion is strongly reflected in his statement:

"the pseudoscientific aroma of the statistical and psychometric superstructures of schools - marks, grades, aggregates and the like - have seduced so many people (including politicians who advocate statewide or national testing schemes) into accepting that quantitative approaches have some magical predictive qualities." (Cohen, 1989 p.14)

In addressing the issue of accountability, Darling-Hammond (1994, p.5) warns that testing students will not provide accountability in education, if assessment methods are not equitable. She insists that the goal of schooling must be to educate all children well, rather than selecting a "talented tenth" to be prepared for knowledge work. For this to occur, according to Darling-Hammond, assessment must support student learning and must include teachers within the process. She believes assessment must be aimed primarily at supporting more informed and student-centred teaching and must be an integral part of "ongoing teacher dialogue and school development" (Darling-Hammond, 1994, p.5).

While it is appreciated that the current trend in evaluation is to ensure that assessment tasks are 'authentic' and reflect what students can actually do (Cohen, 1989; Darling-Hammond, 1994 ; Dwyer, 1989 ; Shrubbs, 1989), there must still be a place for measurement where a number is assigned to the achievement of a student. These numbers or measurements are still useful for describing the amount of certain abilities that individual students have and they represent useful information in the evaluation process (Ebel & Frisbie, 1986, p.14). It is possible to satisfy a community's need for statistical data on student progress and, at the same time, provide teachers with authentic, meaningful tasks which reflect the *Arts Student Outcome Statements* (1996) framework and which can be used to gain reliable measures of student levels of achievement using a Rasch model of item response theory.

Assessment and evaluation in education

Assessment, as used here, refers to the overall process of making judgements about student progress and evaluation refers to the systematic process of collecting, analysing and interpreting information to determine the extent of the skills, knowledge or abilities possessed by students. Evaluation is the step in the assessment process at which judgement is made (Music Educators National Conference Committee on Performance Standards, 1994, p.1; Gronlund & Linn, 1990, p.3).

Ideally, assessment information should be gathered systematically from various sources within the classroom, using a variety of methods to evaluate what actual learning has taken place and, based upon the interpretation of evidence gathered, judgements should be made about the most appropriate educational program that points the way to future learning (Knight, 1992, p.25; Ministry of Education, British Columbia 1994, p.21). It must be emphasised, however, that there is more to evaluation than a collection of techniques. It is important that there is a systematic process which includes the identification of the intended outcome and an end result which identifies the extent to which this outcome has been achieved (Music Educators National Conference Committee on Performance Standards, 1996, p.5; Gronlund & Linn, 1990, p.vii).

Music educators and those in general education alike are of the opinion that all learning fundamental to achieving the goals of education should be evaluated (Bates, 1992, p.5; Gronlund & Linn, 1990; Griffin, 1991, p.2; Ebel & Frisbie, 1986, p.1; Marzano, Pickering & McTighe, 1993, p.v). This evaluation should include all dimensions of learning including performance, attitudes, values, aesthetic responses and critical judgement and should

consist of a variety of planned qualitative and quantitative procedures using formative and summative techniques which enable teachers to assess progress, improve instruction, monitor program effectiveness, and provide a basis for reporting (Bates, 1992, p.5). It should be learner-centred and contribute to student growth as well as reflecting the objectives of the program and should also provide for student self-evaluation to be an integral part of the process.

Performance assessment, which is sometimes referred to as alternative or authentic assessment (Marzano, Pickering & McTighe, 1993, p.13), refers to a variety of tasks in which students have the opportunity to demonstrate their understanding and to apply their knowledge and skills in a variety of contexts and, while performance assessment has been used within *The Arts* for generations, educators in other disciplines are beginning to realise its worth as an authentic assessment approach (Lehman, 1994, p.50; Ogilvie, 1992, p.205; Ministry of Education, British Columbia, 1994, p.2; Forster & Masters, 1996, p.1; Marzano, 1994, p.44). In the past, however, performance in music has mainly referred to the playing or singing of learned or set work and the identification of traditional musical elements. In accordance with more recent educational opinion, that performance should now include the ability of students to produce creative solutions to a given stimulus or criteria and to produce their own art works (Consortium of National Arts Education Association, 1994, p.10; Ministry of Education, Ontario, 1993, p.58; Education Department of Western Australia, 1994, p.2).

In considering all these points, the possibility of the inequitable effects on different populations of students through the use of educational testing, should not be ignored. These inequitable effects include the nature of the assessment tools and how they avoid bias, how they resolve issues about

subjectivity versus objectivity, how they are used to determine student placements and whether they will bring about changes for traditionally underprivileged students (Darling-Hammond, 1994, p.7). In discussing the situation in the United States, Darling-Hammond (1994, p.5) warns that, even with the development of assessment methods which are alternatives to the standardised test type, educators must pay careful attention to the ways that assessments are used. She argues strongly for the use of teacher involvement in the assessment process, and states that policies should ensure that teachers have access to practical information on student learning and that schools engage in the process of 'self-reflection, self-critique, self-correction and self-renewal.'

Assessment in the Arts

The increasing amount of research and publications that have been undertaken in recent years demonstrate the increasing realisation of the importance and worth of *The Arts* to society (Consortium of National Arts Education Associations, 1994, p.5; Emery, 1994, p.5; Rendell, 1994, p. 16; Worby, 1994, p. 13; Romer, 1994, p.vii; Gurin, 1994, p.6; Polisi, 1994, p.8) and the notion that, as in other important learning areas, we should be measuring students' knowledge, skills and progress in arts subjects (Lehman, 1996, p.3). However, while administrators and education systems are currently relying on standardised and system-level testing to assess learning in most subject areas, assessment of achievement in *The Arts* has been neglected. This is probably due to the difficulties entailed in the development of reliable and valid assessment instruments within *The Arts*. Many educators suggest that, apart from a certain amount of objective testing, producing concrete written evidence from students which displays their knowledge of Arts elements, students' success should be evaluated by their responses to

their experiences (Bartel, 1994, p.1; Ogilvie, 1992, p.205; McGuire, 1983, p.12; Lehman, 1994, p.5; Wiggins, 1994, p.202; Bannister, 1992 p.133; Porter, 1992, p.38; Armstrong, 1994, p.6). These responses to experiences in *The Arts* can be described as the students' awareness, understanding, perception and interpretation of artistic qualities and the value they place on those artistic experiences. Most assessment issues in *The Arts* relate to difficulties encountered in evaluating these qualities which, in *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) are referred to as aesthetics, criticism and appreciation.

There is a premise that individuals differ in their response to music and that affective response can be assessed in the linguistic mode; therefore the nature and content of the response to music should be measurable (Bartel, 1994, p.3). However, there are problems, when attempting quantification related to the 'like-dislike' reaction to music and this was demonstrated in a test developed in Queensland (Queensland Department of Education, 1985) in which an attempt was made at assessing aesthetics. However, because an effort was made to avoid subjectivity, the validity of the test is open to question. For instance, after listening to a segment of music on tape, students were asked how much they liked the music. If they said "very much" or "somewhat" they received a score; if they said "not much" or "not at all" they did not score anything. Subjectivity is necessarily a part of *The Arts* and cannot be avoided. Students should be entitled to express their opinions. However, although this information could be useful to a teacher within a classroom, it is difficult to see how it could be a measurable trait in a testing situation.

A more appropriate way of using student opinion is to ask the question, but rather than evaluating the response, use it as a prompt and ask them to

justify their opinion. The justification of their response can then be measured in terms of their knowledge of music. In another section of the Queensland Department of Education (1995) music test, students were asked to write as many words as they could think of to describe the music. However, answers were scored on the number of words written; not what the words were! This would seem to be more a measurement of the students' literacy and range of word skills rather than their knowledge of music. Also, the notion of 'more is better' is bad measurement practice. It assumes that a student who writes ten poor descriptors has more knowledge than the student who writes three relevant, meaningful descriptors. If no instruction is given as to the number and type of words required, this student could, justifiably, believe that he/she has adequately completed the task.

Similar problems were evident in the second national art assessment program (National Assessment of Educational Progress, 1981) conducted between 1974 and 1979 in the United States of America. Students were asked to value items such as furniture and sculptures by agreeing or disagreeing that it was 'all right for items to look like this'. It was reported, in the case of the furniture, that items with the highest responses were essentially representational or functional and, in the case of the sculptures, that items receiving the lowest positive response were those that displayed extreme simplicity or exaggeration in form, or those that employed unconventional techniques. Whilst reporting of results on these tasks attempts to avoid subjectivity, it is difficult to comprehend exactly what information about students learning in *The Arts* was gained. These two cases are cited purely in an attempt to emphasise the difficulties involved in the assessment of notions such as appreciation, aesthetics and criticism.

Lehman (1994, p.47) discusses the problems concerned with assessing the Arts, saying, "Assessment in the arts is a complex task, fraught with problems and pitfalls," although he agrees with music educators who believe there is a need for a systematic methodology of evaluating progress in *The Arts* (Bates, 1992, p.15; Willingham, 1992, p.41; Ross, Radnor, Mitchell & Bierton, 1993, p.xi). In the past, the main criteria for reporting student success in music have often been attendance and a positive attitude. Lehman (1994, p.47) believes that despite the problems involved in developing reliable, measurable test instruments, assessment of student learning will unquestionably become a major force in arts education in the near future. There are significant difficulties in assessing problem-solving ability and higher-order thinking skills (Ross, Radnor, Mitchell & Bierton, 1993, p.10), as well as in writing guidelines for scoring students' dance improvisations, dramatic scripts, or musical compositions. Although Lehman (1994, p.47) agrees it is difficult to define a task that measures imagination or creativity, he feels it can be done. He emphasises the necessity for both clear statements of the purposes of the assessment before techniques are developed, and for a specification of the context in which learning is assessed.

We are frequently warned against object-based evaluation measuring the product at the end of a project instead of evaluating students by their responses to authentic performance-based experiences on an ongoing basis which measures every stage in the teaching-learning process (McGuire 1983, p.12; Leyman 1994, p.47; Ross, Radnor, Mitchell & Bierton, 1993, p.x). While this notion is important in all areas of the school curriculum, it would seem even more important in *The Arts* where presentation and performance are a vital part of any arts discipline, including music.

The NAEP 1997 Arts Report card assessment of music, theatre, visual arts and dance (U.S. Department of Education, 1997) has addressed this issue and, in fact, there are a number of parallels between that large-scale assessment and strategies described in this study.

Music Assessment in Western Australia

The only official standardised music test used in music in Western Australian primary schools is the dated *Aural Foundations of Music Reading* test (Bentley, 1966). This test was designed as an aptitude test rather than a test of achievement and has traditionally been administered to students at the commencement of Year 3, 4, 5 or 6 (depending on the musical instrument of study). In order of aptitude for music, it ranks students who will be considered for selection for instrumental music tuition by an instrumental music teacher from the School of Instrumental Music. There are no creative skills, aesthetic considerations or performance involved in the *Aural Foundations of Music Reading* test (Bentley, 1966). It consists of a 20 minute tape recording of electronic sounds testing students' listening skills in the musical elements of pitch, rhythm and texture to which students respond in a multiple choice format. The electronic sounds used in the test do not reflect the music sounds to which students are accustomed and the multiple choice questioning technique does not give students the opportunity for any creative response. For these reasons, it would not be possible to gauge their knowledge and skills or their creative abilities in music from their results. For instance, items include the playing of two sounds, pitched a tone or semi-tone apart, and students are asked whether the music goes up or down. As its name suggests, the test is purely designed to test students' abilities to aurally distinguish sounds. Results of the tests are retained by the School of

Instrumental music and are not made available to primary schools for gathering school data or for the purposes of reporting student progress to parents.

Special music secondary schools in Western Australia currently use the *Advanced Measures of Music Audiation* test (Edwin & Gordon, 1989) to identify students with an aptitude for music in July of Year 7 (12 year olds). This test is used, together with auditions and interviews, to determine which students will be accepted into secondary school music programs. Although this test also uses tape recordings of electronic sounds, it is more difficult than the *Aural Foundations of Music Reading* test (Bentley, 1966). The *Advanced Measures of Music Audiation* test (Edwin & Gordon 1989) requires students to identify very slight, and difficult to identify, changes in rhythm and pitch and uses musical terminology and long examples. If students score reasonably well on this test, they are considered to be good candidates for the music program although there is no performance requirement or aesthetic or contextual component. Two distinct disadvantages of the test are the length of time taken for its administration and the difficulty involved in marking it. The marking process involves the use of four separate marking keys.

Music students enrolled in *Unit Curriculum* (Education Department of Western Australia, 1987a) Class Music programs in Western Australia undergo assessment procedures using standards-referencing based on grade-related descriptors where a grade of A, B, C, D or F is assigned before progressing from one stage to the next. For example, for the unit 'Perception & Basic Music Knowledge', a descriptor for an A grade is "memorize, recall and note *all* rhythms and melodies". For a B grade, the word *all* is substituted for *most*, for a C it is substituted for *some*, for a D the word is *few* and for an F it is *very few*. Teachers do not undergo any marker training to reach a

consensus on the meanings of words such as 'most', 'some' and 'few' and 'very few', and it is difficult to believe that consistency on these gradings is achieved. What is the difference between 'few' and 'very few'? Grade descriptors for each of the units use the same wording format to describe the levels. Schools are still using this system of unit progression although it was developed in the mid-eighties and philosophies on assessment have changed since that time. Although there are six stages in the unit which progress in level of difficulty from 1 to 6, schools do not always offer every stage and so students do not necessarily cover each stage sequentially. Assessment consists of 50% ongoing evaluation of activities within the classroom and 50% formal testing. For students undertaking the Instrumental and Ensemble Music programs, 70% of this testing consists of individual instrumental work which includes sight reading, technical knowledge and performance of prepared pieces, and 30% consists of ensemble playing. Again, grade related descriptors include words such as 'most', 'some' and 'few'. *Unit Curriculum* (Education Department of Western Australia, 1987) programs are undertaken mostly in secondary schools but there are a limited number of primary students enrolled in the Instrumental and Ensemble Music programs. These students are identified as being talented through administration of one or both of the tests mentioned above.

The most formal standardised testing in the area of music in Western Australian schools is the testing which is undertaken for Tertiary Entrance Examination and which is included as part of *The Curriculum Framework for K-12 Education in Western Australia* (Interim Curriculum Council, 1996). Written and performance tests are administered at the completion of Year 12 (17 year olds) and students' scores are aggregated to determine in which university courses they will be eligible to enrol. A significant part of this testing (40%) involves either performance, where students perform two or

three pieces which demonstrate their technical proficiency, or a project which students present and discuss during a 20 minute *Viva Voce* with the examiner. The performance pieces must represent a minimum of three different periods or genres of music which are chosen from a selection of five pieces submitted to the examiners by the student prior to examination. Students select one piece and the examiner selects another, which is usually a contrasting piece. The examiner may decide to select part of two other pieces. The project encompasses the history and literature of music and students must be prepared to answer a series of questions related to their topic. Testing also involves aural work which includes identification of such things as pitch, chords and rhythmic dictation, as well as a three hour written segment of the test in which students are required to write about specific composers or specific schools of music (Interim Curriculum Council, 1996). This three hour examination, together with compulsory class work including composition, make up the other 60% of the aggregate for the Tertiary Entrance Examination in music.

A criticism of the Tertiary Entrance Examination in music is that it does not fit the Curriculum Framework (Interim Curriculum Council, 1996): that is, rather than allow students to explore a range of cultural forms, it is based on Western Art tradition. The requirements are so strenuous that teachers in classes from as early as Year 8 (13 year olds) tend to 'train' students for the Tertiary Entrance Examination instead of providing them with a wide cultural range of material. Another criticism is that it does not take today's technology into account: that is, tools that can enhance music skills are not being regarded. For instance, there are computer programs for aural training that translate sound into musical script. This is becoming an acceptable way of writing and transposing in the music industry. A further problem is created by specialist music schools, who compete for the most talented students in their

courses, continually pushing standards higher and higher. The students undergoing these courses are usually students who have learnt music from a private teacher for many years and their high levels of expertise make it extremely difficult for a student who has only participated in school music to compete. With the level of competition to obtain a high aggregate for tertiary admissions, even the most competent music students sometimes elect not to include music in their selection of tertiary entrance examination subjects, as it involves many hours of practising performance and studying the history of music which could be spent studying subjects which are considered easier and which could be expected to attract a higher aggregate due to the larger population of students taking the tests.

It would seem that the time has come to question the demands being made on music students undergoing Tertiary Entrance Examinations as numbers of students opting to take these examinations are comparatively low. The latest figures available from the Secondary Education Authority indicate that 363 students from both government and non-government schools sat for the Tertiary Entrance Examination in music in 1996. This compares with a total of 11,572 students who sat for examinations in English, 10,910 in mathematics, 2676 in history, 4609 in geography, 2906 in physics and 1149 in art. A total of 12,072 students sat for examinations in at least one subject. (Secondary Education Authority, 1996, p.95)

The Western Australian music curriculum program, *Music in Schools* (Ministry of Education, 1989), is commonly in use in primary schools, where there is a music specialist. This program is divided into five teaching sections involving the elements of rhythm, melody, harmony, form and expression and gives guided instructions to teachers in listening, singing, moving, playing, exploring sound and reading and writing music. However, there is no

assessment program and teachers usually use a checklist, devised by themselves, to identify students' mastery of these elements. There are no guides to the checklist and many primary school teachers evaluate students on their 'enjoyment of music', 'participation' or 'behaviour'. *Music in Schools (Ministry of Education, 1989)* was written before the development of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) and there is no inclusion of aesthetics, criticism or past and present contexts in this program. Cultural diversity is also limited. The series, however, contains an excellent explanation of the elements of music and would be a useful document for planning classroom music programs that could be used, in conjunction with the outcomes framework to monitor student outcomes.

Specialist teachers frequently enhance their teaching repertoires with commercial teaching programs such as *Upbeat* (Leask, 1989). This program consists of a series of songs, with teaching points and related activities, including aural activities, based on the elements of music, and are designed to be used from early childhood through to Year 7. There are pieces from a range of cultures as well as pieces designed for special religious or other occasions which are designed to elicit aesthetic response and to strengthen the enjoyment of music. The *Upbeat* (Leask, 1989) series has a form of evaluation at the end of each section and teachers are encouraged to administer this test before moving on. In his Foreword of *Upbeat*, Leask (1989, p.6) describes the evaluation technique and suggests that teachers use a 'pretest posttest' technique to evaluate students. He suggests that test scores should be systematically tabulated on individual students' progress charts and interpreted by teachers. A criticism of this method is that there is no explanation to teachers as to how students' test scores should be interpreted and there is no framework or benchmark on which teachers can

gauge progress. Another problem is that, although *Upbeat* (Leask, 1989) emphasises the importance of developing aesthetic sensitivity through contact with music in its philosophy (p.9) the multiple-choice style evaluation sheets test only the elements of music and knowledge of instruments. There are also open-ended inventive tasks, which encourage exploration and creativity. Unfortunately, however, there is no marking key to guide teachers in the evaluation of these tasks. Despite these criticisms, *Upbeat* is a well designed, 'user friendly' series that contains a wide range of strategies for investigation, acquisition of skills and the opportunity for aesthetic response. It has the advantage of an audio-tape to accompany each teacher's book and it is one of the most widely used series in Western Australian schools. While the evaluation forms in *Upbeat* (Leask, 1989), do not equate with the developmental levels of the *Student Outcome Statements* (Education Department of Western Australia, 1996), the activities and assessment programs could be adapted and related to levels in the various strands of the statements and used for reporting to parents.

The *Springboards* (Farmer, 1984) series is also widely used in Western Australian schools and, like *Upbeat* (Leask, 1989) this series aims to develop aesthetic sensitivity in students by providing a wide range of cultural pieces and material for religious and special occasions. The series is designed to progress in level of difficulty throughout primary school and contains ready-made programs under the headings of specific objective, learning activities and teaching methods, notes/resources and evaluation ideas. The ideas are based on instrumentation and music elements and there is no suggestion of evaluating aesthetic sensitivity or students' critical responses to music. The evaluation ideas section suggests that teachers write the 'specific objectives' in behavioural terms and use them as guidelines for assessment by making checklists and constructing rating scales or tally sheets to record results. The

evaluation ideas do not include any scale of measurement and so would not be useful, in their present form, in providing data on student or whole school progress. The series, however, contains some suitable material for use in primary schools and provides teachers with some very useful ideas on programming. It would be possible for teachers to match the content to the student outcomes framework in order to map student progress through the levels.

Performance assessment

Performance assessment is the gathering of information about student learning based on students demonstrating what they can do. It values the process of learning as well as the product and incorporates a variety of strategies, from observation to self-assessment (Ministry of Education, British Columbia, 1994, p.1; Lehman 1994, p.52). In the case of music performance, assessment has traditionally included singing or playing a musical instrument but educators in other disciplines are now realising the value of performance assessment (Education Department of Western Australia, 1994b; Education Department of Western Australia, 1995) and it is increasingly being used in subjects such as science, physical education, mathematics and oral language. It is usually an on-the-spot evaluation and the only manner in which a 'concrete' product can be judged at a later stage is through the use of tape or video recording (Forster & Masters, 1996, p.1).

There are some kinds of learning that can only be assessed through the observation of student performance and, although these are performances on specific tasks, the intention is to infer the student's achievement in that particular discipline; that is, to generalise about the students' overall performance in that discipline from a limited performance sample. It is

important that this observation of performance plays a key role in assessment within *The Arts*, as there are relatively few useful items in *The Arts* (such as knowledge of specific information) that can be machine-scored (Lehman 1994, p.52). However, it should be kept in mind that there needs to be objective validation and concrete evidence to support and justify conclusions of that observation. Some educators argue strongly for performance based assessment (Ogilvie 1992. p.205; Darling-Hammond, 1994, p.8; McGuire, 1993, p.12; Lehman, 1994, p.47; Forster & Masters, 1996, p.5; Ministry of Education, British Columbia, 1994, p.2; Marzano, Pickering & McTighe, 1993, p.26). However, the issue of validity, that is, the relevance and coverage of material and the issue of reliability will always need to be taken into account when assessing the usefulness of evidence for inferring students' levels of achievement. If the observational assessment of performance is properly undertaken and tasks are constructed to explicitly include selected standards, performance assessment can be valid and reliable.

There are three benefits of performance assessment outlined by the Ministry of Education, British Columbia (1994, p.2). First, there is the opportunity to integrate assessment and instruction which may occur at any point during an activity for the benefit of both teachers and students. Second, there is the opportunity for collaboration between students and teachers in the assessment and evaluation process which allows students to become aware of strengths and areas for growth through reflection and discussion, as well as to become risk takers, critical thinkers and problem solvers. Third, multiple means of assessment may be used, including informal observations, check-lists, video and audio taping, conversations and conferences between teacher and students and the promotion of a diversity of approaches and responses to learning situations. Darling-Hammond (1994, p.8) adds to this by stating that the equitable use of performance assessment is dependent, not only on the

assessments themselves, but how they are used in relation to the goals of authentic school reform and effective teaching. She also maintains that when assessment is used for decision-making purposes it can exert powerful influences on curriculum and instruction and can drive instruction in ways that mimic the format and cognitive demands of tests.

Many researchers and Arts educators today are of the opinion that it is no longer valid to test students on content or knowledge only. They argue strongly that it makes no intellectual sense to test only for "knowledge" in *The Arts* (Armstrong, 1994, vii; Lehman, 1994, p.48; Wiggins, 1994, p.202). There is a belief that contextualised performance assessment, although more difficult to score reliably, is the only valid way to test, not only in *The Arts*, but throughout all facets of education (Wiggins, 1994, p.202). Wiggins cites test-makers who are more concerned with the precision of scores than with the intellectual value of the challenge and emphasises the need for ongoing negotiation, in relation to the conflict between validity and reliability. He goes on to say:

Modern, professionally designed tests intended for national and state use tend to sacrifice validity for reliability. In other words, test-makers generally end up being more concerned with the precision of scores than with the intellectual value of the challenge....While this conflict between validity and reliability must never be construed as an either/or choice, it remains a design problem to be carefully negotiated (Wiggins, 1994, p.202).

While Ogilvie (1992, p.205) agrees with the strategy of performance assessment and strongly supports its advantages, it is his opinion that there are still grounds for including a certain amount of objective testing which produces concrete written evidence from students. Concrete evidence of this nature and precision of scores is currently a priority in system-level testing and, according to Boughton, Eisner and Ligtoet (1996, p.200), "in

contemporary United States, a certain form of examination - the standardised, multiple choice, machine-scored test - has come to be regarded as almost sacrosanct". Indeed, in most states of Australia as well, there appears to be a politically driven thirst for quantitative test results to substantiate accountability in education systems. The New South Wales Basic Skills Program currently tests the entire cohorts of Year 5 (9 year olds) and Year 3 (7 year olds) students in Government schools, as well as around 100 students per cohort per year from self-selected non-government schools in numeracy and literacy using mainly multiple choice, machine-readable items (Lokan & Ford 1994, p.7). A similar format is used at the Year 10 level in numeracy, literacy and science (Lokan & Ford 1994, p.9). System level testing is now undertaken in every State of Australia in the literacy and numeracy skills in which students are tested for knowledge using single response-type or multiple choice-type items that are machine scorable. The Monitoring Standards in Education program in Western Australia is the only Australian system-level testing program which assesses the performance of students and the only one which has tested in eight learning areas (Lokan & Ford 1994, p.6). However, pressure from the Federal Government has been exerted to include whole of population testing in literacy and numeracy in this state. The only economically feasible way to achieve this is by using multiple choice, machine scorable testing methods. It is to be hoped that this will not compromise the rich, performance-based testing programs in other subject areas that are currently in place.

Portfolio assessment

Portfolios can be described as a purposeful collection of the student's work or artefacts that show the student's effort, progress and achievement over a period of time (Ministry of Education, British Columbia, 1994b, p.1).

Portfolio assessment has become widely accepted as a valid and reliable method of gathering data for student assessment (Forster & Masters, 1996, p.1; Ministry of Education, British Columbia, 1994b, p.2; Marzano, Pickering & McTighe, 1993, p.41). This method of assessment is seen by some educators (Lehman, 1994, p.50; Darling-Hammond, 1994, p.21; Forster and Masters, 1996, p.1) as a preferable alternative to isolated testing within *The Arts*, as it offers a broad view of student learning by documenting processes in the learning as well as presentation of the final products. Portfolio assessment provides direct evidence of students' skills and learning, by documenting creative work, performances and responses which demonstrate growth over an extended period of time as well as providing concrete evidence for evaluation and goal setting.

In discussing performance-based assessment and educational equity, Darling-Hammond (1994, p.22) emphasises its usefulness as a form of teacher development because analysis of student results by teachers leads them to further development and improvement of their pedagogy. She cites the practice at New York's International High School (where the entire student population is 100 percent limited English proficient immigrants) of using portfolios, projects and oral debriefings on the work of cooperative learning groups to judge the effectiveness of both students' progress and teachers' own teaching strategies. The portfolios are evaluated by the students themselves, their peers and their teachers (Darling-Hammond, 1994, p.21). While student evaluation of their own work is a valuable strategy in the learning process and in their overall development, care should be taken in judging its reliability in terms of measurement or the opportunity to make comparisons between students or groups. Their untrained judgements could be influenced by many factors such as high or low self esteem, lack of

knowledge of benchmarks on which to base results, different expectations between teachers and schools and lack of standardisation of criteria.

Portfolio assessment can be used in many different educational contexts and for many different purposes (Forster & Masters, 1996a, p.1). There is no one portfolio; there are many portfolios (Forster & Masters, 1996a, p.2). However, all portfolios, whatever their purpose, contain pieces of evidence and the relevance of those pieces is the degree to which they address the knowledge, skills and understandings of the learning area. For example, a portfolio that services the assessment needs of a classroom teacher may not necessarily be the most appropriate form for use in a system level assessment program. A recent survey of assessment methods in *The Arts* undertaken in Canada and reported by Roberts (1994, p.4) indicated that, of all Arts teachers surveyed, 81.5 per cent felt it was important to assess students using an "individual project developed to conclusion".

Portfolios are used in Visual Arts in Western Australian secondary schools in Years 11 and 12 for Tertiary Entrance Examinations. There is a structured set of five assessment criteria: one, organisation, which relates to the arrangement and layout of the folio; two, discernment, which relates to the students' abilities of self expression and discrimination; three, visual language, which refers to their visual and verbal understanding and design concepts; four, inter-relationships, which integrates notions such as art history, visual enquiry and critical analysis; and five, drawing skills, which includes the development of ideas and sensitivity to the discipline. Students are made aware of the assessment criteria and, because of the high stakes involved, in tertiary entrance, marking is centralised, with markers undergoing formalised training in an attempt to standardise levels. A quantitative score from 1 to 20 is allocated for each of the five assessment criteria, giving a possible total

score of 100. Markers are given descriptors of 'excellent' relating to scores from 17 to 20, 'high' relating to scores from 13 to 16, 'sound' relating to scores from 9 to 12, 'limited' relating to scores from 5 to 8 and 'inadequate' for scores from 1 to 4 (Education Department of Western Australia, 1993, p.20). These descriptors equate with the scores 'A' through to 'F' which are traditionally used in secondary schools and marking relies on the marker's 'on-balance' judgement according to definitions of the assessment criteria. For 'organisation', markers look for authenticity - the student's own work should demonstrate a personal expression of ideas, concepts, processes and product; for 'consistency' - evidence of equal effort and time allocated to each project; for 'sequence' - evolution of ideas presented in a logical order, commencing with a student brief and concluding with a studio photograph; and for 'layout' - an appropriate standard of presentation that readily links all parts of the visual diary [portfolio].

Marker reliability is obtained by holding marker-training meetings where marking procedures are demonstrated by expert markers and each form is marked by four markers. Marks are then compared and differences of four marks or greater for any criterion forms the basis for discussion and reconciliation within the group. There is also opportunity for identification of points needing further clarification. After the training process each portfolio is marked by two markers with markers being paired in overlapping patterns. Again, if the difference in scores for a criterion is four points or more, then that criterion must be remarked and reconciled. If the difference is three or less then reconciliation may be determined by averaging the two scores. If consensus cannot be reached, markers must refer to the Supervising Examiner. Random checks for further reliability are made by comparing scores allocated with scores which have previously been obtained from class teachers. Given the nature of a visual arts portfolio, it would seem difficult to

achieve a more reliable system of marking. However, a criticism of the criteria is their lack of relationship to *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) levels.

A less formalised strategy for allocating marks is used in Western Australian secondary schools in Years 8, 9 and 10 in Visual Arts, as part of the High School Certificate assessment, where it is a requirement that a portfolio, referred to as 'workbook/folio', forms part of the assessment strategy. The 'workbook/folio', which includes drawing, design and other visual inquiry, is kept by students and evaluated by teachers using five criteria. These are: one, use of media, which refers to the skill to select the most appropriate materials to solve problems; two, research/design development skills, which refers to the ability to gather, organise and assess information; three, knowledge and understanding about art, which refers to the ability to describe and apply facts, principles and concepts; four, critical and evaluative skills, which is the ability to use correct vocabulary and the correct categories of description and analysis; and five, understanding of processes, methods and techniques, which is the ability to use appropriate art processes, methods and techniques in their personal work.

Marking is done by the teacher, rather than centrally, and a score of A, B, C, D or F is allocated according to a marking key supplied (Ministry of Education, 1989a, p.39). A criticism of the marking key is the level of interpretation required by the teacher. For example, for the criteria, research/design development skills, students must collect and use information from a "wide variety of sources" to gain an A, and to gain a B they "collect and use information from several sources". Unless teachers were given specific guidance as to the difference between "a wide variety" and "several" their interpretations could be quite different. To gain an A for the criteria 'critical

and evaluative skills', a student "identifies problems in order to improve performance" and to gain a B a student "assesses performance to identify possible improvements" (Ministry of Education, 1989a, p.42). The difference between these two may be difficult to discern. Although teachers do, from time to time, moderate their marking by setting and marking common tasks within their schools, there is no formal moderation strategy at a system level so interpretations could vary considerably from one school to another. However, the criteria for assessment cater for a broad range of skills and an attempt was made in 1989 to obtain consistency at a system level by supplying criteria and a marking key for teachers. If moderation strategies were put into operation at a system level, discrepancies in grade allocations between schools could be minimised.

Although this type of portfolio assessment has been in use in the discipline of Visual Arts for some time, it is only in recent years that some educators have recognised its value in other areas of the curriculum (Ministry of Education, British Columbia, 1994, p.2; Forster & Masters 1996, p.10). It is seen as providing a broad view of student learning by providing a place to document processes in learning as well as involving students in taking responsibility for their own learning. The structure depends upon the educational purpose for which the portfolio has been designed. Forster and Masters (1996, (p.2) suggest that the first stage of the portfolio design process is deciding on the portfolio purpose by describing the assessment purpose and the instructional purpose, and reviewing these descriptions against important objectives for mapping against outcomes. The second stage is deciding on content by describing the types and range of evidence sought and reviewing these against the outcomes. The third step involves deciding on the portfolio selection by describing the portfolio selection procedure and the management system and reviewing these against the outcomes. The fourth

stage involves deciding on what will be assessed and the assessment criteria by describing criteria clearly, ensuring they do not favour a particular gender or cultural group and reviewing these descriptions and criteria against the portfolio purpose and the outcomes. The fifth stage involves deciding on a method for estimating and reporting locations on a progress map by describing the method for reporting locations and reviewing these descriptions against the portfolio purpose and audience.

The Ministry for Education, British Columbia (1994b, p.8), describes four basic elements for setting up portfolio assessment as; setting purposes for portfolio collections; identifying guidelines for portfolio collections, developing evaluation criteria; and providing for student reflection. The portfolio strategy mentioned earlier, which is used in Western Australian secondary schools as part of the assessment procedure at years 8, 9 and 10 (Ministry of Education, 1989a), is an example which includes these elements. If the process of marking and assessment were updated and mapped against *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) and steps were taken to ensure more reliability among markers, this would provide an improved method for assessing music achievement over current practices in Western Australia.

Student self-assessment

Student self-assessment provides students with the opportunity to take more responsibility for their own learning and to set themselves appropriate learning goals and plan ways to achieve them. They are able to reflect on past work, how successful it was, how it could be improved and how they can use it to their benefit in the future. It also raises their awareness of how their thinking has changed over time. The process of student self-assessment is

being used widely in many facets of education today (Forster & Masters, 1996b, p.14; Marzano, Pickering & McTighe, 1993, p.35). For example, in Western Australia, the *Technology and Enterprise Student Outcome Statements* (Education Department of Western Australia, 1998) include the sub-strand 'Evaluating' which is a very significant part of the technology process. At this stage, students “evaluate intentions, plans and actions with a view to modification and improvement. They develop and apply criteria to assess how well they have responded to the design challenge.”

Because the nature of *The Arts* is subjective, there is no right or wrong answer in appreciating or expressing *The Arts*. It is essentially a self reflective process and the process of student self assessment is considered an important strategy by Arts educators (Knight, 1992, p.26; Ogilvie, 1992, p.205; Forster & Masters, 1996b, p.14; Marzano, Pickering & McTighe, 1993, p.35). One reason for this is it enables students to share in ongoing evaluation and makes them aware of their achievements, as well as empowering them in the overall process with the opportunity to grow through self-reflection and documentation of the learning process in collaboration with the teacher. More recently, educators in other subject areas have also advocated using student self-assessment as part of the evaluation process. For instance, The Ministry of Education, British Columbia (1994b, p.1) includes students' self assessment as an essential component in their portfolio assessment strategies as it involves students in reflecting on their performances, products, thinking, and learning, as well as evaluating the quality of their work and knowledge and setting realistic goals for themselves. Forster and Masters (1996, p.8) also emphasise the importance of students having the skills, knowledge and confidence to evaluate their own thinking processes, work and progress if they are to develop as independent learners.

The student journal has been identified as being one of the most powerful tools to do this (Marzano, Pickering & McTighe, 1993, p.35). For these reasons, student self-assessment is a valuable strategy in the learning process and a useful method of assessment within the classroom. However, care should be taken in judging its reliability in terms of measurement or the opportunity to make comparisons between students or groups and, before adopting any form of this strategy, it would be necessary to look for correlations between results of self-assessment, written evidence and observational methods. The different criteria students use to assess their work, together with their different capabilities in making judgements, will create difficulties in standardising results. That is, student self assessment results cannot be used to make comparisons between student achievements because common criteria, standards or scales have not been used. There is value, however, in using student self-assessment of their own performances in *The Arts* , not as a part of the measurement process but as a method of gaining more insight into student knowledge and their abilities to identify areas where they could improve on performances.

Computer-based assessment in The Arts

Traditional methods of testing usually involve pen and paper tests using multiple-choice-type items which focus on 'atomistic' knowledge. In other words, they test for knowledge, portrayed in parts which is out of context and not always related to the body of knowledge as a whole (Fetherston, 1995, p.2). Tests which deconstruct important knowledge and skill into disconnected, unmeaningful sections are unable to validly assess complex activities (Resnick, 1994, p.523). In *The Arts*, students usually know and can demonstrate, much more than can be tested using a pen and paper test of this kind. To demonstrate their ability to analyse or appreciate a piece of

music or art work, they need to be able to view or listen to a stimulus and respond to it. A multimedia assessment scenario provides the opportunity for innovative ways of presenting items, using simulated environments involving representations of two-dimensional and three-dimensional objects, which can be manipulated by students and animated where appropriate. These simulation techniques, that offer students tasks which are more realistic and closer to those they encounter in everyday life, are likely to be more authentic and valid (Fetherston, 1995, p.4) and have the potential to test application and analysis rather than merely recognition. A frequently cited disadvantage of pencil and paper tests is the delay in feedback on scores (Fletcher & Collins, 1987 in Fetherston, 1995 p.4). Multimedia assessment techniques have the potential to provide immediate feedback to the test-taker thus making results more meaningful and lowering test anxiety (Fetherston, 1995, p.4).

The National Arts Education Research Center in Illinois reported on a Computer-based program which had been designed to 'Assess the Development of Music Listening and Rhythmic Performance Skills of Secondary School Students' (Leonhard, 1990, p.11). Not only did the computer control the presentation of screen graphics to display music notation and text, it also accepted musical input from students using an electronic keyboard linked to the computer through a digital interface card. One of the advantages of the system is the number of items that can be generated on a developmental scale. Another important application lies in the program's capacity to assess musical skills (Leonhard, 1990, p.11). "For the first time, the computer can be used to assess student musical performance or reactions to musical stimuli rather than their cognitive abilities or ability to recall information about music" (Leonhard, 1990, p.11).

Webster's (1995, p.22) optimism is equally encouraging in his descriptions of the technology as being an 'especially powerful way to get kids to compose, improvise and really listen to music in ways never before thought possible.' He believes that, for the first time, we can offer students many exciting new ways to think creatively about music because of the technological tools available and that, as researchers, we can study creative behaviour in more effective ways through technology. Webster (p.26) discusses the ever-increasing availability of music drill and practice software and the flexible options that allow students and teachers to take control of their learning. There are at least six key features for flexible practice. They include; stimulus/response items with flexible performance criteria; a comprehensive approach with multiple tasks; intelligent branching tailored to individual need; realistic music examples within context; flexibility in designing learning environment; and on-line tutoring for music concepts. Students are able to practise music decision making in real world music settings. This allows them to interact with the technology in a similar manner to the four ways humans interact with creative music experiences; listening, performing, composing and improvising. Although he does not discuss assessment strategies for his programs, Webster (1995, p.32) points out the advantage of students' appraisal and criticism of their work, and personal record keeping. An advantage of a program such as this is the opportunity for students to take control of their own learning. There is a strong push in Western Australian schools at present for this strategy. Another advantage is the capacity to provide opportunity for students to perform a variety of tasks, including stimulus/response items, which could be used as assessment tools within the classroom, providing a valuable alternative to pencil and paper testing. It is possible that computerised testing could be a preferred alternative to pencil and paper testing in the future. During experiments carried out by the researcher in Western Australian schools, indications were that, as long as

students were computer literate, relationships between the same items generated on a pencil and paper test and in computer form remained stable.

A comparison of relative item difficulties of the pencil and paper version and a computerised form of the Raven's Progressive Matrices was carried out by Styles & Andrich, (1993) using a Rasch latent trait model. The consistency of the responses across the two modes was evaluated by comparing the relative item difficulties of the computerised form with those from the pencil-and-paper version and to convert scores between the advanced and standard forms using the two modes of testing (Styles & Andrich, 1993, p.905). Results within each analysis showed high conformity of the data to the model. When a simple logistic model was used to produce conversions from the Standard Progressive Matrices to the Advanced Progressive Matrices, from both the computerised version and the pencil and paper version, there was stability of the relationships between items on the two versions which produced virtually identical conversions (Styles & Andrich, 1993, p.923)

Students' reading and comprehension skills are an ongoing problem in the administration of written tests designed to assess performance in subject areas other than reading and comprehension. The use of digitised sound in computerised, multi-media testing means that students can hear questions read aloud as many times as necessary, enabling assessment of musical skills and knowledge, rather than assessment of reading and comprehension skills. There is also potential for results to be collated and analysed within a program, thus eliminating many hours of scoring, recording and analysis of results by teachers.

Norm-based Assessment vs Criterion-based Assessment

Norm-based assessment provides a measure of performance that is interpretable in terms of an individual's relative standing in some known group (Ebel & Frisbie, 1986, p.275; Lehman, 1994, p.51; Gronlund & Linn, 1990, p.16; Kubiszyn & Borich, 1987, p.29). It has traditionally been the most widely used type of group-referenced interpretation. A standardised test used in the discipline of music in Western Australia is the *Aural Foundations of Music Reading* test (Bentley, 1966). It is used to identify musically competent students in primary schools and is norm-based on Australian students. Raw scores on the test are tallied to give the number of correct responses in the test and students are ranked, in order of aptitude for music, to be considered for selection for instrumental music tuition by an instrumental music teacher from the School of Instrumental Music. Students with the highest scores are selected, regardless of which items they got right or wrong.

Special music secondary schools in Western Australia also currently use norm-based assessment. They administer the *Advanced Measures of Music Audiation* test (Edwin & Gordon, 1989) to identify students with an aptitude for music and acceptance in special music schools is based on performance in this test. The formalised testing used for tertiary entrance testing in Western Australia is norm-referenced in that a process of standard distribution is carried out to ensure that a mean of 58% is obtained. According to Kubiszyn & Borich (1987, p.28), although an advantage of norm-referenced tests is that "you get an estimate of ability in a variety of skills in much shorter time than you could through a battery of criterion-referenced tests," it has some disadvantages in that it does not always provide information on what is learned, how it is learned or what individuals can or cannot do (Glaser & Nitko, 1971, cited in Griffin & Nix, 1991, p.88). It provides comparisons

between individuals, schools and systems that have been seen as unnecessary and even inhibiting learning (Griffin & Nix, 1991, p.88).

Criterion-based assessment yields information that is directly interpretable in terms of specified performance rather than where an individual stands in relation to others. In other words, it means that an individual has met a particular standard or pre-specified criterion in an explicit content area and is described in terms of some specifically defined set of skills or sphere of knowledge (Lehman, 1994, p.51; Griffin & Nix, 1991, p.91 ; Ebel & Frisbie, 1986, p.27; Gronlund & Linn, 1990, p.56; Withers, 1991, p.13; Kubiszyn & Borich, 1987, p.29). In order for criterion-based assessment to take place, it is necessary to have a set of defined standards of increasing competence and defined measures on a growth continuum. The performance of a task is interpreted by the relative position of the task on this continuum which indicates the development of competence. Some educators believe criterion-based testing has the advantage of decreasing the likelihood of excessive competition among students (Lehman, 1994, p.51; Griffin & Nix, 1991 p.88). It is important that the difference between the notions of standards and criteria are clarified however, as judgement of a work against pre-specified criteria does not necessarily mean a good standard has been achieved, even if all criteria have been addressed by the work (Boughton, 1995 p.3). Boughton clarifies the two notions by saying "criteria express the qualities we value in an object or performance, and standards express the degree to which they should exist" (Boughton, 1995 p.3). In Western Australian schools, the level descriptor at the relevant level of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) expresses the qualities to which Boughton refers. The standards to which he refers will be set by establishing benchmarks with the assistance of the testing described in this study.

Increasingly, systematic assessment is becoming criterion-based with tasks being ordered in coherent sets that lead to an overall interpretation of proficiency or competence (Lehman, 1994, p.51; Griffin & Nix, 1991, p.91). However, care should be taken that the overuse or inappropriate use of criterion-referenced assessment does not lead to incoherent sets of skills which may have little relevance to instruction (Griffin & Nix, 1991, p.92).

A precedent for large-scale assessment using criterion-based assessment was the music portion of the National Assessment of Educational Progress which was conducted in 1971-72 in the United States of America. This type of testing is more expensive and time-consuming than machine-scoreable norm-based testing and it is significant that, in the 1978-79 round of National Assessment of Educational Progress testing, performance was eliminated for economical reasons. National Assessment of Educational Progress testing in *The Arts* in 1997 will be performance based (Lehman, 1994, p.52).

There are problems with most norm-based and criterion-based methods of assessment in *The Arts*. If we are to use only norm-based assessment, we have no description of what students can do or the skills and abilities they possess, although we do know how they are progressing compared with other students. For instance, in music we may know that John possesses more knowledge than Jane in identifying rhythms but we don't know the extent of John's accomplishment. Can he identify only simple three or four beat rhythms, or is he capable of identifying complex sixteen beat compound rhythms? On the other hand, if we use only criterion-based assessment we have descriptions of their abilities but no framework on which to base them and so it is not possible to supply feedback on the progress of student populations so often required by Governments and other stakeholders. For

instance, if John can identify simple three or four beat rhythms, how does he compare with other students in his Year level? By using a Rasch analysis (Rasch, 1980), arising from the development of item response theory, it is possible to simultaneously scale item difficulties and student abilities on the same scale and to link them with a framework such as *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), thus allowing us to report on the criteria as well as the placement of students on a learning continuum. The use of item response theory and Rasch analysis, together with the development of the outcomes framework, could provide an important and interesting improvement in Arts assessment.

Rasch measurement in The Arts

The literature reveals limited examples of research using Rasch measurement strategies within The Arts. Myford (1989) uses a Rasch model of measurement (Rasch, 1980) to determine whether expertise in making aesthetic judgements exists and, if so, what is the nature of expertise in performing the task of judging aesthetics? Myford, (1989, p.1) found that, in the past, researchers' attention was focused nearly exclusively upon agreement between expert judges as being the sole criterion for detecting expertise in aesthetic judgement. Experts, according to researchers, showed stronger agreement in their aesthetic responses to works of art than less experienced judges or novices. Although researchers found that experts reproduced their ratings with a high degree of accuracy, it was not known whether the same applied to novices, as they had not undergone the same test-retest reliability trials, until Beard (1978) produced data to suggest that experts had higher test-retest reliability than novices. Myford suggests that perhaps the ability to reproduce one's ratings may be a useful criterion for identifying expertise in aesthetic judgement, but asks the question; are there

other criteria beyond this that might differentiate the aesthetic judgement of experts from those of novices? (Myford, 1989, p.3). The rating data were analysed to establish item difficulty for 36 items on a Judging Acting Ability Inventory (Myford, 1989, p.11) for each of three judge groups. These were experts, who were casting directors and experienced high school drama teachers; theatre buffs, who were not formally trained but attended professional theatre regularly and had some knowledge of the criteria used to evaluate acting; and novices, who attended theatre very infrequently, and had little training or experience in drama beyond high school.

Results indicated that Experts consistently rated performances more harshly than buffs and novices. Experts were able to employ multiple criteria in judging and were better able to replicate their ratings than were buffs and novices. That is, the amount of change for buffs was nearly twice that for experts, while the amount of change for novices was nearly four times that for experts (Myford, 1989, p.145). Myford suggests that an effective means of tackling the problem of differences between judges is to convert the raw score ratings to measures and then to correct these measures for differences in calibrated judge harshness. This results in measures that are 'judge free', that is, adjusted for the calibrated harshness of the individual judge supplying the ratings (Myford, 1989, p.148).

This study, which employed Rasch rating scale analysis methodology to facilitate the construction of measures of both within-judge and between-judge variation in ratings, has implications for marking procedures employed to assess *The Arts* in education. It may be necessary, in the future, to assess differences between markers, and to make allowances for those differences. Measurement in such subjective notions as aesthetics has long been one of the reasons for the reluctance in assessing *The Arts*.

Fairhall (1989) used Rasch models of measurement in a study of response to paintings to determine the relationship between dogmatism and aesthetic judgement. Subjects for the study consisted mainly of Western Australian students of design, craft, art education and fine art with comparisons being made with music students who represented an aesthetic, but not visual, discipline, and accounting students who represented a non-aesthetic discipline. The study addressed the question of the relationship of dogmatism to the aesthetic judgement variables. Subsidiary questions regarding the relationship of dogmatism to particular aspects of art judgement behaviour, the way people would value art works of differing kinds, and how they would make other, non-evaluative judgements (Fairhall, 1989, p.82) were also addressed.

Two Rasch psychometric models (Andrich, 1982a; Andrich, 1983a, 1985) were used to analyse Western Australian test data to characterise the response of a person to a polychotomously scored item as being governed by a person parameter, representing attitude or ability, and by properties of the item. The first model, requiring data from polychotomous items with at least three response categories or from subsets comprising at least two dichotomously scored items, was applied for all data analysis except for the Dogmatism Scale where skewness was of special interest as a manifestation of possible response bias. Here, a second model, with the skewness parameter requiring a minimum of four response categories or three dichotomously scored items in subtests, was used. Results of the study bore out its main predictions; that there would be negative relationships between dogmatism on the one hand, and all of aesthetic evaluation and boldness of judgement variables on the other (p.159). That is, more dogmatic people could be expected to fail to understand and hence, react negatively to novel forms or abstract forms of art as against more naturalistic paintings, and,

consequently they could be expected to place a lower value on these forms than would more open minded people (p.159). Interestingly, this line of reasoning was extended to music, where there were no differences in responses of dogmatic and open minded people to more familiar, older romantic musical systems, but they would exhibit different responses to less familiar forms such as contemporary musical systems (Fairhall, 1989, p.159). This study further highlights the problems with which Arts educators are confronted when assessing students' aesthetic judgements.

Ethnographic strategies

Ethnographic strategies of assessment in *The Arts* usually involve observation techniques which enable educators to measure student performance in its social context. It is a qualitative strategy used to search for an understanding of phenomena which are not usually quantifiable as opposed to the quantitative strategy of research traditionally used in the past. Ethnographic methodologies can vary to include systematic observation where the incidence of selected behaviours are usually recorded on a graph of some description; case studies where the researcher examines the profile of an individual; or participant observation where the researcher is an active participant and, at the same time, makes descriptions and evaluations (Swanwick, 1984, p.202).

Ethnographic methods of participant observation and interpretation are frequently used in *The Arts* in early childhood education where teachers observe children's experiences in physical activities such as play, drawing and moving where they express themselves through pure aesthetic response and through use of symbols (Wright, 1994, p.28). These methods of assessment are particularly suited to early childhood education where assessment is

closely linked to program evaluation (Cartwright, 1989; Salvie & Ysseldyke, 1992; Wright, 1987 in Wright, 1994, p.30) with the teacher programming and making daily changes that are based upon the children's learning (Wright, 1994, p.30).

Some music educators believe that no music has value except in a social context of some sort and that ethnographic strategies are needed to evaluate it (Bannister, 1992, p.133; Wright, 1994, p.28). Other educators identify potential weaknesses of ethnographic strategies, questioning their validity and warning that subjectivity in assessment must be controlled (Swanwick, 1984, p.202). However, in early childhood education, where formal testing of very young children is inappropriate, participant observation procedures derived from qualitative-naturalistic traditions (Alexander, 1982; Almy & Genishi, 1979; Dyson, 1988; Spodek, Saracho & Davis, 1987 in Wright, 1994, p.30) would seem to be appropriate to the needs of the teacher in program evaluation and in the assessment of children's accomplishments.

Ethnographic strategies are also useful for the individual researcher in collecting evidence about the ways in which students accomplish musical tasks or develop musical awareness. They enable teachers to keep anecdotal records and profiles of individual children's progress in processes such as discovery, pursuit, perception, communication, skill use, creativity, analysis and critique (Wright, 1994, p.28). However, it would seem to be a practical and economical impossibility at a system level to use ethnographic methodologies of participant observation techniques to gather valid data specifically related to student outcomes. Teacher judgements and the criteria for assessment would vary from school to school unless lengthy and expensive training of teachers was undertaken and the alternative, to send judges from one school to another, would be time-consuming and expensive.

Wright (1994, p.32) advocates ethnographic strategies of assessment in pre-schools, using a strategy of formative assessment, which refers to the evaluation of learning outcomes that derive from the effectiveness of the teaching program, and that of summative assessment, which refers to the collection of data that provides information about the learner's progress and, to this end, Wright has designed a summary of the hierarchy of skills ranging from 'discovery' through to 'critique' which teachers could use as an evaluation guide (Wright, 1994, p.35). However, the guide is purely qualitative and teachers would use it as a checklist as there is no measurement scale or allowance for quantitative data.

Outcomes based education

The National Standards for Arts Education (Music Educators National Conference, 1994) developed for use by schools in the United States of America are described as being 'concerned with the *results* that come from a basic education in the arts, *not with how those results ought to be delivered*' and the standards framework developed for use in Victoria (Board of Studies, 1995) are described as "a set of standards which it is expected students across the State will attain at different stages in their schooling." A more complex definition of outcome-based education by Willis and Kissane (1995, p.2) in Western Australia is :

For a school or education system (which may be a nation, state or district) to adopt an outcome based philosophy means, in effect, that the system believes there are certain things that all students should learn as a result of attending its school(s), that it is prepared to say publicly what these things are, and that it is prepared to stand accountable in terms of them.

Rather than focusing on what systems and schools have provided and what teachers have taught, outcome based education focuses on what students

have learned and this must, of necessity, involve a shift in curriculum policy, practice and evaluation. Decisions about what to teach and how to teach it should be driven by the outcomes we expect students to demonstrate at the end of their educational experience. It is desirable, therefore, that teachers develop a shared and improved understanding of what the outcomes are in order to judge students' learning validly and reliably. One of the advantages of outcome based education is that it provides an improved approach to accountability within the classroom, schools, and within the system (Willis & Kissane, 1995, p.2). Improved accountability, however, will only be achieved if outcomes are measured against the outcomes framework using reliable methods of measurement that establish benchmarks and allow teachers, schools and the system to map students' progress.

Statements which describe behaviour patterns indicating levels of learning, rather than general statements such as 'doing well' or 'should work harder' are important if teachers, parents and the community are to be kept informed on how students are progressing (Griffin, 1991, p.4). Student outcome statements provide these indicators, which can be used for the collection of evidence about goal attainment for teachers and curriculum developers (Griffin, 1991, p.5).

The *Arts Student Outcome Statements* (Education Department of Western Australia 1996), which describe the expected levels of achievement of students along a developmental continuum from level one to level eight, are divided into the four strands. These are *Exploring, developing, creating and communicating ideas; Using skills, techniques and technologies; Responding, reflecting and evaluating; and Understanding and using historical, social and economic contexts*. These outcome statements, together with descriptive level statements, provide the 'set of criteria' which give the basis for criterion-based

assessment (Education Department of Western Australia, 1996). These statements have been adapted from *The Arts - a curriculum profile for Australian Schools* (Curriculum Corporation, 1994b) which were developed collaboratively between education systems in Australian States and Territories to assist in the improvement of teaching and learning in Australian schools. All states in Australia have now adopted these statements, or adaptations from them, describing levels of achievement through the compulsory years of schooling (Curriculum Corporation, 1994b). A similar set of standards has been implemented in the United States of America (Consortium of National Arts Education Associations, 1994) and in England (British National Curriculum, 1993).

Whilst statements of student outcomes are useful in providing a growth continuum which indicates the development of competence and which provides an ordered, coherent set of indicators for criterion-referencing for teachers and curriculum developers, there may be a disadvantage for parents and other members of the community who are not familiar with statement levels and for whom they may not have the same significance (Griffin & Nix, 1991, p.92). Detailed explanation of levels and their meanings would be needed for reporting at a system level. There could also be a disadvantage for inexperienced teachers in that student outcome statements are designed to describe student progress and hence, do not contain detail of curriculum content. The development of curriculum material describing resources appropriate to various levels could overcome this problem.

Summary

Assessment procedures used in Western Australian schools in The Arts have been limited. In the discipline of music, they have included the

commercially-produced, multiple-choice formatted *Aural Foundations of Music Reading* test (Bentley, 1966) and the *Advanced Measures of Music Audiation* test (Edwin & Gordon, 1989) used in primary and secondary schools. In secondary school music, the standards-referenced *Unit Curriculum* (Education Department of Western Australia, 1987a) grading, which includes portfolio assessment, is used in Year 10 and the formal standardised testing procedure for tertiary entrance, which includes performance, is undertaken in Year 12. The assessment of arts disciplines has included performance assessment, portfolio assessment and student self-assessment. Assessment procedures have been norm-based or criterion-based, providing students with a 'mark' or 'grade' related to their knowledge of arts elements or performance. They have not provided descriptive information about what the student can actually do and they have not included the assessment of student skills and abilities in relation to aesthetics.

Evidence indicates that, in the area of *The Arts*, Australian education systems, as well as many overseas education systems, have adopted outcome levels or continua describing progress of student achievement (Education Department of Western Australia 1996; Curriculum Corporation, 1994; Consortium of National Arts Education Associations, 1994; Department of Education and Science, 1989). Although the current emphasis on the development of performance indicators may have its genesis in economic rationalism and accountability, educators persevere with directing the use of indicators and assessment to educational purposes and they need reassurance that information gathered is used justly, with due respect for differing contexts (Hewton, 1991, p.vii).

Many Arts educators believe that authentic, performance-based assessment, including the process of developing performance to conclusion,

responding to and analysing art work, and the use of portfolios is more valid in assessing *The Arts* than the use of multiple-choice, machine-scorable pencil and paper tests, and that assessment should be criterion-based, using developmental indicators (Ogilvie, 1992, p.205; Darling-Hammond, 1994, p.8; McGuire, 1993, p.12; Lehman, 1994, p.47; Forster & Masters, 1996, p.5; Ministry of Education, British Columbia, 1994, p.2; Marzano, Pickering & McTighe, 1993, p.26). Arts educators also believe that *The Arts* have traditionally been relegated to a secondary status in schools (Gordon, 1992, p.24; Jorgensen, 1994, p.26; Carlton, 1987, p.45; Kemp & Freeman, 1988, p.21) and that the necessary support strategies must be provided to ensure *The Arts* become a part of any school-wide or district-wide assessment program undertaken in the future (Lehman, 1994, p.53; Fehrs-Rampolla, 1994, p.31). It is impossible for teachers to find more time in the school day to include more subject areas and teachers need to be encouraged to use arts forms in combination with each other or to integrate them into other learning areas. *The Arts* provide the opportunity for students to develop creative ways of expressing themselves and to develop criteria for critically appreciating, analysing and making informed judgements about their own and others' work. The integration of *The Arts* into learning areas such as Studies of Society and Environment, English Literature, Technology and Enterprise and Physical Education will provide the opportunity for these skills to be demonstrated and consequently assessed, within context, across all curriculum areas. The inclusion of *The Arts* as one of eight compulsory learning areas in Western Australian schools and the development of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), together with the Education Department's decision to include *The Arts* in their monitoring standards project in 1996, are positive steps in the direction of more authentic performance measures of arts, in general, and music, in particular, in Western Australian schools.

Research questions

The review of literature for this study has lead to the following research questions:

1. Can a music test be devised to assess appreciation and expression, that uses an outcomes focus and that can be administered and marked by classroom teachers?
2. Can a reliable marking key be generated to provide multiple categories of responses at different levels of achievement for open-ended tasks?
3. Is it possible to match the music achievement scale to *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) and determine level cut off points?
4. Will it be possible to generate descriptive profiles of student performance on a scale of achievement that assists classroom teachers meet the demands for accountability in schools?

These research questions are linked closely to the aim of this study (p.12).

CHAPTER THREE

THE MODEL AND THEORETICAL FRAMEWORK

Measuring the achievements of students in a subjective and non-quantitative area such as The Arts is likely to be difficult and certainly, in the past, Arts educators have been somewhat reluctant to apply quantitative measures to Arts disciplines that include such things as aesthetics and non-verbal arts language. It is possible, however, to measure student knowledge and abilities on a continuum of achievement, with the use of a descriptive profile or reporting framework which describes a progression of knowledge, skills and understandings against which student achievement can be measured. To gain a measure of students' knowledge of any art form, including music, it is essential to observe both the ability to practice it, that is, to 'do it' or 'make it' and the ability to understand and appreciate the discipline (Mercer & Church, 1998). It is, therefore, necessary for any conceptual framework on which assessment is based, to address these two criteria.

The Arts Student Outcome Statements (Education Department of Western Australia, 1996) are divided into two sections which address the concept of "doing" and the concept of "understanding". The first section is related to *expressing* and comprises two strands entitled *Creating, exploring and developing ideas* and *Using skills, techniques, technologies and processes*. The second section is related to *appreciating* and comprises two strands entitled *Responding, reflecting and evaluating* and *Understanding the role of the Arts in society*. Each strand is a developmental continuum which maps student achievement within a defined area of The Arts across eight levels. These levels indicate the progression of student learning from simple experiences drawn from play and imagination to complex tasks demonstrating sensitivity, artistry, technique and cultural and historical knowledge. At each

level of development, activities may appear to have some similarities, but are more demanding, requiring students to work with increasingly complex and challenging ideas, more refined skills, and in wider social contexts. Students also experience more complex aesthetic responses as they progress through the levels (Curriculum Council, 1997, p.53). This approach has been referred to as a spiral curriculum (Bruner, 1960). It should always be kept in mind, however, that students progress at different rates depending on a large number of variables reflecting attitudes and experiences that reflect different cultural and geographical settings (Curriculum Council, 1997, p.53). It is, therefore, unrealistic to attribute certain levels of achievement to specific school Year levels.

It should also be remembered that divisions between levels in a learning continuum of this nature are somewhat arbitrary because the progress of a student from one level to the next is a gradual process rather than something that happens suddenly. For example, students may be at a stage where they are achieving most of the Level 3 outcomes and some of the Level 4 outcomes. The sequence [of levels and their indicators] does, however, provide a useful framework for describing and discussing students' progress within a specific area of learning (Masters, 1994, p.5).

Indicators of learning have always been used, either intuitively, or by design, by teachers to analyse and describe students' work (Griffin, 1991, p.3). In the case of classroom arts disciplines, this has often been in an informal way. Using the common set of statements or indicators described at each achievement level, it is now possible to develop tasks and items which allow students to display behaviour typically found at each achievement level on a strand, providing a basis for the descriptive interpretation of that behaviour in a formal way.

The model involves the identification of goals, the delineation of appropriate outcomes, as described by the indicators at each level, and the methods of assessment used in gathering the information. The tasks, or items, are matched to outcomes for each goal. This chapter will outline, briefly, the skills and knowledge students are expected to display at each level, the model used in the derivation of the tests, and the generation of the scores.

This model is proposed as a general model for use in measuring outcomes in any of the Arts disciplines, for, although they are described as the five discrete disciplines of dance, drama, media, music and visual arts, the student outcome level descriptions are common to all, requiring the use of a variety of aural, kinaesthetic, tactile, spatial/visual and verbal symbols in the relevant discipline which progress along a continuum of development. The first two strands necessitating the display, by students, of behaviours related to expression in the relevant discipline and the second two strands necessitating the display, by students, of behaviours related to appreciation in the relevant discipline. The model assumes that there is a need for a context around which tasks and items are designed. To provide this context, a stimulus or prompt is provided for each test.

The model assumes that there will be a number of situation variables relating to schools and students that may influence the measurement of student achievement in music in a system-level testing situation. These include the experience of the teacher administering the tests, the quality of audio equipment in the school, the quantity and quality of musical instruments available and the literacy skills of students. An attempt has been made to address these differences in four ways. First, through the inclusion of detailed guidelines and instructions for teachers administering the

tests. Second, by recommending the use of 'found' and 'made' sound sources as well as traditional musical instruments. Third, by collecting data on the performance of sub-groups of girls and boys, non-Aboriginal and Aboriginal Torres Strait Islander students, and English speaking and non-English speaking background students. Fourth, by taking measures, such as instructing teachers to read questions aloud, in an effort to reduce the effect of students' literacy skills on results. However, the inclusion of open-ended tasks necessary to measure the progression of student achievement, may have handicapped students with poor literacy skills in responding to tasks where their written skills are required to interpret the musical language of the stimulus.

A further influencing factor on the measurement of achievement in music will be teachers' attitudes to music and to the use of outcome statements to measure student progress. For instance, those who believe that music is an unimportant learning area may feel that it is not worth measuring. Those who believe in the old traditional methods of classroom teaching and are unreceptive to change in teaching or assessment methods may prefer a more structured, right or wrong answer, style of test, rather than a test where students are required to display their use of musical language through performance. These attitudes could lead to teachers being negative in their administration of the tests and this negativity could affect student results.

Teachers' receptivity to the use of outcome statements to measure student progress has, in some cases, been negative. However, the statements are now a part of government school policy and it will be compulsory for teachers to adopt them. The collection of data in the eight learning areas, which include *the Arts*, is also now a compulsory requirement in government schools. This test design will provide a model that will assist

teachers in using an outcomes-based approach to the assessment of classroom music.

The model allows for the collection of data across all strands and sub-strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). The two *Expressing* strands provide the opportunity for students to explore, develop, create and communicate ideas through their activities in music.

The Model

Expressing as part of the model

There are two strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) that relate to the expressing of The Arts and two strands that relate to the appreciation of The Arts. The *Expressing* strands describe the use of skills, techniques and technologies in music in exploring, developing, creating and communicating through students' musical activities and musical works. The first strand of the model is *Creating, exploring and developing ideas*, and the second is *Using skills, techniques, technologies and processes*. The two strands are developed into eight ordered levels of achievement from low (level 1) to high (level 8). These two strands from level 1 to level 8 are summarised as Appendix xxvii.

Appreciating as part of the model

The 'appreciating' strands require students to respond to, reflect on, and evaluate their own musical works and the work of others, using their

aesthetic understanding. Students understand that music is shaped by historical, social and economic contexts and use this understanding both in their own work and when responding to the work of others. The 'appreciating' strands are, *Responding, reflecting and evaluating* and *Understanding the role of the Arts in Society*. These two strands from level 1 to level 8 are summarised as Appendix xxviii.

In order to test the strands related to both *Expressing* and *Appreciating*, it was necessary to ensure that the students had the opportunity to display their understanding of music through the use of an array of symbol systems. This includes both the non-verbal language of the arts to express an idea and the most common system of language, the spoken or written symbol (Mercer & Church, 1998).

To display their knowledge and skills in the strands of *Expressing*, students had the opportunity to use both non-verbal arts language in the performance of their musical compositions, and written language to illustrate their planning and reflection. To display their knowledge and skills in the strands of *Appreciating*, it was necessary for students to receive and read the specific language of music within the stimulus, and then to translate it into written language (Mercer & Church, 1998, p.2). It must be appreciated that, while students might be constrained in their use of written language to fully interpret the subtleties of the art message, this is difficult to avoid in a testing situation.

Test design

The importance of evaluating students' skills in both practising and appreciating the discipline is emphasised by the fact that *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) encompass outcomes in both the expression and the appreciation of The Arts. The conceptual framework developed for this study addresses these two criteria. The outcomes framework is developmental, providing a progression of expected outcomes from Level 1 to Level 8. The instrument developed in this study allows for the observation of the development of student skills from Year 3, through Year 7 to Year 10 through the use of open-ended questions and link items across the three Year groups.

The Analysis test was designed to address the 'appreciating' strands of the outcome statements and the Process test was designed to address the 'expressing' strands. The Analysis test was a paper and pencil test consisting of a combination of multiple choice and extended answer question types and, where possible, tasks were open-ended. Linking of items through Years 3, 7 and 10 was achieved through the use of common stimulus material and common tasks. Year three students were given only one stimulus piece entitled *Ballet for Children* (Bliss, 1995). Year 7 students were given this piece, in addition to a second piece entitled *Dharpa* (Kellaway & Yununpingu, 1992), and Year 10 students were given these two pieces as well as a more complex, contemporary piece entitled *Earthcry Kakadu* (Sculthorpe, 1989). The structure of the tests were similar with the stimulus being played in parts

on a tape and students answering the questions between the playing of each part. For Year 3 groups, teachers were instructed to read questions aloud, whereas for Years 7 and 10, students read the questions, with teachers clarifying comprehension problems where necessary. Simple question types were used at Year 3 whereas Year 7 students had the additional tasks involved with comparing and contrasting the two pieces. Year 10 students were presented with a combination of these, in addition to more complex items that provided the potential for students to respond up to Level 8.

The structure for the Process tests was the same for Years 3, 7 and 10 with test duration times being adjusted to encompass two normal class 'periods' for that Year level. Students participated in a warm-up activity before being presented with the stimulus material, the brainstorming activity and the opportunity to develop and present their composition in a group situation. This provided the opportunity to map student progress across the three Year groups. The stimulus presented to Years 7 and 10 was a painting entitled *Heaven and Earth* (Pericles, 1978). However, because of the difficulty in interpreting a painting at the Year 3 level, the Year 3 students were provided with a taped storm sequence, which provided visuals but no sound.

First, an examination was made of the *Arts Student Outcome Statements* (1996) to determine the type of testing that would be necessary to gather the maximum amount of data on student performance in the strands related to both *Expressing* and *Appreciating* and, at the same time, be feasible in a classroom testing situation. Second, in order to cover both sections, two test forms were designed; a Process form and an Analysis form.

Third, a marking key was generated for each of the forms and music teachers were trained to mark them (see Figure 3.1: Music test design using an outcomes framework). Third, raw scores were obtained for the Analysis test only, the Process test only, and for a combination of both tests. Fourth, a Rasch analysis of the scores and an estimate of fit was undertaken for the Analysis test, the Process test, and for a combination of both. Ability estimates revealed that the fit was better by combining the results of both tests. Fifth, an examination of scores, together with individual student scripts was carried out to obtain level cut-off points and to estimate student levels (see figure 4.1: Process for student ability level estimates, p.120).

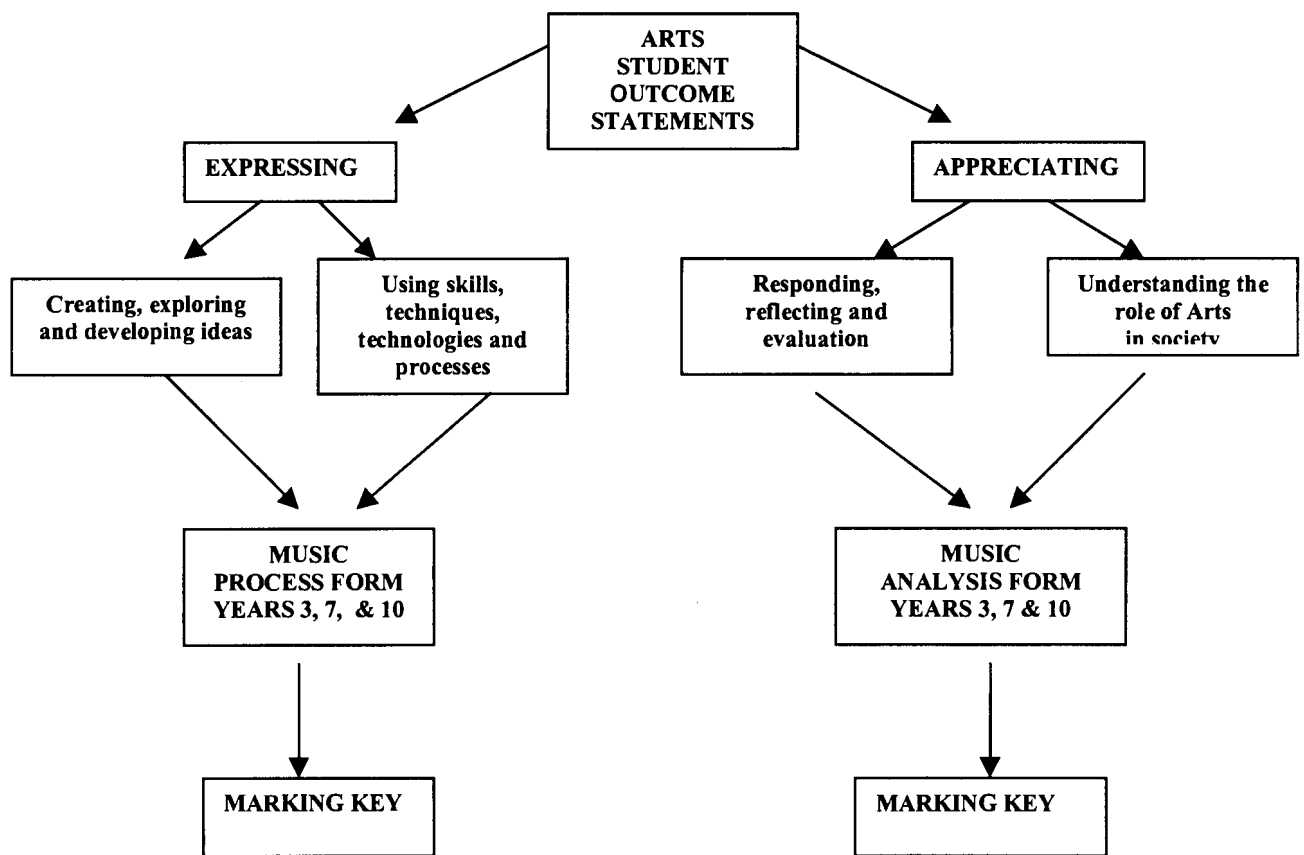


Figure 3.1: Model of music test design using an outcomes framework

A Rasch Analysis was undertaken to establish student ability levels and item difficulty levels. In four cases, where categories were not discriminating sufficiently, categories had to be collapsed and the items rescored. This will be discussed more fully in Chapter Six.

Development of the test items

Four main considerations were taken into account when developing test items for this study. The first was the identification of what should be evaluated and obtaining a balance between expectations and what could realistically be achieved, together with clear guidelines as to how judgements would be made. The second was the issue of social justice, and the moral, ethical and legal implications involved, when selecting stimulus materials and wording items. The third consideration was the issue of time allocations and striking a balance between what would be ideally suitable at each of the three year levels, and what could be realistically expected in terms of teacher and classroom timetables. Fourth was the issue of manageability of administering the tests for teachers who are not music specialists (The Joint Committee on standards for educational evaluation, 1980).

The development of assessment tasks in music is based on Western Australian Education Department policy on teaching and learning and is informed by resources such as *Student Outcome Statements* (Education Department of Western Australia, 1996), current exemplary models of teaching and assessment practices in music (Curriculum Corporation, 1994a, 1994b; Education Department of Western Australia, 1993; Ministry of Education, 1989a, 1989b), recent and current national and international assessment literature and research (Music Educators National Conference

Committee on Performance Standards, 1996; Ministry of Education, British Columbia, 1994a, 1994b, 1994c, 1994d; Forster & Masters, 1996a, 1996b, 1996c), collaboration with classroom teachers, interest groups, education department consultants and the trialling of materials in schools. A Student questionnaire designed to gain information on students' background and experience in music was included in the trialling process in order to further clarify expected abilities of students in various situations. This was necessary because some primary schools in Western Australia have specialist music teachers; others have no specialists and music is taught by classroom teachers. At the Year 10 level, only students undertaking music options were tested.

An examination of the *Student Outcome Statements* (Education Department of Western Australia, 1996) was undertaken to extract a wide spread of outcomes within each strand, and to establish which levels of outcomes should be aimed at each Year level. This was necessary before deciding which testing strategies would be appropriate and manageable within the practical and financial constraints of trialling the materials.

Structure of the Analysis test

A combination of multiple choice and extended answer question types was included in the tests and, where possible, tasks were open-ended in order to provide students with the opportunity to demonstrate their maximum levels of ability. As this was an assessment of music, student responses were not assessed for spelling or writing skills. Through the use of common items and common stimulus material, tasks allowed for linking of items through

Years 3, 7 and 10, thus providing valuable information on student progression through the outcome levels. Where subjective questions, asking for students' opinions or reflections were asked, they were used as prompts for further justification and were not scored.

At Year 3, teachers were provided with an audio tape of the *piece Ballet for Children* (Bliss, 1995), which was recorded in parts, as well as containing verbal instructions for teachers on where to pause the tape. Teachers were then requested to: read the questions for part 1, play the passage of music for part 1, and read the questions one at a time, giving the students reasonable time to answer before going on to the next question. When part 1 was completed, they then repeated the procedure for parts 2 to 7.

The test contained thirteen questions that were designed to address the outcome levels primarily in the strand of *Appreciation* from Level 1 to 5. The test is included as appendix ii. All questions in the test, apart from multiple choice items, had the capacity to earn partial credit for students who answered below the targeted Level. (see appendix iii; *Year 3 Music Analysis Marking Key*).

Question 1 demonstrates a Level 1, multiple choice item. Students were asked; "Where would you be most likely to hear this piece of music?" They chose their answer from the selection provided which was; birthday party, orchestral concert, street parade, rock concert. This item addresses the Level 1 statement; *Identifies arts experiences in their own lives* in the sub-strand

Understanding the role of the Arts in Society (Education Department of Western Australia, 1996, p.3).

Question 2 demonstrates an extended answer item type and asks students to “Explain what you heard in the music that made you pick this answer” [referring to their answer to question 1]. This question provided the opportunity for students to provide a range of responses from Level 2; that is, *Outlines features of their own and others’ arts works and activities using simple arts terminology relating their responses to these features*, to Level 5; that is, *Uses arts terminology and critical frameworks to analyse and express informed opinions about arts works and activities* in the sub-strand *Responding, reflecting and evaluating* (Education Department of Western Australia, 1996).

Question 5 represents an example of a subjective question asking for students’ personal responses. Students were asked for their interpretation of the mood of the piece by selecting from the answers; “sleepy,” “happy,” “sad,” or “angry.” Where students were asked for a personal response such as this, answers were not assessed. However, this type of question was always followed up by asking for a justification of their response as demonstrated by question 6; that is “Explain what you heard in the music that made you pick this answer.” This question required an extended answer that demonstrated students’ knowledge of the elements of the music and allowed them to respond up to Level 5 in the sub-strand *Understanding the role of the Arts in society*, that is: *Identifies and discusses distinguishing features of arts works*

which locate them in a particular time, place or culture (Education Department of Western Australia, 1996).

At Year 7, teachers were provided with an audio-tape of the same stimulus piece as that for Year 3, with an additional piece entitled *Dharpa* (Kellaway & Yununpingu, 1992). The format was similar to that of the Year 3 test with the test being presented in parts, from part 1 to part 9, containing a total of 15 questions. Teachers were instructed to: ask the students to read the questions for part 1 (or read aloud if you think that it is necessary), play the passage of music for part 1, give the students reasonable time to answer all the questions in part 1. When Part 1 was completed they were then asked to repeat the procedure for parts 2 to 9.

Question types were similar to those in the Year 3 tests with the addition of a “compare and contrast” item, as demonstrated by question 14, which allowed the students to compare and contrast the two stimulus pieces in the areas of instrumentation, expression and rhythm. This question addressed the Level 5 statements; *Identifies and discusses distinguishing features of arts works which locate them in a particular time, place or culture*; and *Identifies and discusses the distinguishing features of arts works and activities in contemporary Australian society* (Education Department of Western Australia 1996, p.3) from the sub-strand *Understanding the role of the Arts in Society*. This test is included as Appendix iv.

Questions 3, 4, 5, 6, 7 and 10 were linked to the Year 3 test. This

provided the opportunity for comparisons to be made, and progress to be mapped, between Years 3, 7 and 10 students. Items were coded so that the same item was given the same code name across the three levels. For instance, Year 3 item 7, Year 7 item 3 and Year 10 item 3 was coded MU07 (see Table 4.1, *Analysis test Item Links and Levels*). As for the Year 3 test, answers to questions in the Year 7 test which were not multiple choice item types earned partial credit for lower level responses (see Appendix v; *Year Seven Music Analysis Marking Key*).

Table 3.1: Analysis test item links and levels

Code	Max score	Year 3	Year 7	Year 10	SOS Level
Mu01	1	1			App 2.1
Mu02	2	2			App 2.3, 2.4
Mu03	1	3			App 2.1
Mu04	2	4			App 1.3, 2.3, 2.4
Mu05	No score	5			App 1.1, 2.1
Mu06	2	6			App, 1.2 - 1.4
Mu07	1	7	3	3	App 1.2
Mu08	2	8	4	4	App 1.4, 1.5
Mu09	1	9	5	5	App 1.2, Exp 2.2
Mu10	1	10	6	6	App 1.2
Mu11	1	11	7	7	Exp 2.3
Mu12	4	12	10	10	App 1.2 - 1.5
Mu13	3	13			App 1.2 - 1.5
Mu14	1		1	1	App 1.2
Mu15	3		2	2	App 1.3 - 1.5
Mu16	No score		8	8	App 1.1 - 2.1
Mu17	3		9	9	App 1.3 - 1.5
Mu18	1		11	11	App 2.2
Mu19	3		12	12	App 2.2 - 2.5
Mu20	3		13	13	App 2.3 - 2.5
Mu21	4		14a	14a	App 1.3 - 1.6
Mu22	4		14b	14b	"
Mu23	3		14c	14c	"
Mu24	4		15		App 1.2 - 1.5
Mu25	3			15	App 1.2 - 1.8
Mu26	4			116	App 1.2 - 1.7, Exp 1.6
Mu27	4			17	App 1.4 - 1.5

Key:
Mu14: Music coded item 14

Key:
App: Appreciating strand
Exp: Expressing strand
1.3: sub-strand 1, level 3
2.5 sub-strand 2, level 5
SOS: Student Outcome Statements

The structure for the Year 10 Analysis tests was similar to that of the Year 3 and 7 tests. Both of the stimulus pieces used at Year 7 were provided, together with an additional, more complex, contemporary piece entitled *Earthcry Kakadu* (Sculthorpe, 1989). The test consisted of 17 questions and the audio-tape was played in 10 parts.

Item types similar to those of the Year 3 and 7 tests were used, with the addition of more complex items, providing the potential for students to respond as high as Level 8; *Researches arts works from a variety of contexts, understanding how histories are constructed in the arts and how their own expression and appreciation of the arts is shaped by them*; and *Critically examines the ways the arts challenge and shape values and are influenced by prevailing values* (Education Department of Western Australia, 1996, p.3). An example is Question 13, which asks; "What effect has this style of music had on Australian culture?" (see Appendix vi; *Year 10 Music Analysis Test*).

It should be emphasised that, while items in all tests at Years 3, 7 and 10 were targeted to address particular outcome levels, all, apart from multiple choice items, allowed for partial credit to be awarded and the analysis of the data, using a Rasch model, provided item difficulty estimates which enabled outcome Levels of achievement to be established. This process is described in detail in Chapter Six. Partial credit item categories for the Year 10 tests are outlined in the *Year 10 Music Analysis Marking Key* (see Appendix vii).

It was possible to make comparisons among the three Year levels, and to map progress from Year 3, through Year 7 to Year 10 through the use of link items. Questions 3, 4, 5, 6, 7 and 10 in the Year 10 tests are linked to both the Year 3 and Year 7 tests. An example of a successful link item is Question 10 in the Year 10 test which refers to the stimulus piece *Ballet for Children* (Bliss, 1995) which asks students to “Explain how the music ends.” This question enabled students to provide responses varying from a simple Level 1 answer such as “It ended very loud” to high level responses where they aurally identified and described distinguishing features and used musical language to describe and discuss elements such as harmonic and rhythmic tension (Refer to Table 4.1; *Analysis test item links and levels*).

Structure of the Process Test

The structure for the Process tests was the same for Years 3, 7 and 10. First, students participated in a directed music warm-up that was intended to focus students' thinking on the creative use of sound and different musical elements. They were presented with a stimulus that they examined before participating in a class brainstorming activity to discuss the stimulus. They were then instructed to: write down their own ideas about different sounds that could be used to represent the stimulus, join a small, pre-determined group to plan a composition to reflect the stimulus and notate the composition in either traditional form, or their own style. Groups then rehearsed their pieces before performing them for the class. Teachers videotaped the group performances for central marking. Specific instructions were given for the videotaping

process to avoid differences in the quality of productions. After all groups had presented their items, students were asked, individually, to complete a critique of their groups' performances.

Links were achieved through Years 3, 7 and 10 by using the same procedure, the same items and the same marking key across the three Year groups. Tasks were developmental so that, potentially, it was possible for students at all levels to achieve as high as Level 8. (see Table 3.2: *Process test item links and levels*).

Table: 3.2 Process test item links and levels

Code	Max score	Year 3	Year 7	Year 10	SOS Level
MuP01	4	1	1	1	Exp 1.1 - 1.8, App 1.1 - 1.8
MuP02	4	2	2	2	Exp 1.1 - 1.8
MuP03	4	3	3	3	"
MuP04	4	4	4	4	Exp 1.1 - 1.8
MuP05	4	5	5	5	Exp 2.1 - 1.8
MuP06	4	6	6	6	"
MuP07	4	7	7	7	"
MuP08	4	8	8	8	"
MuP09	4	9	9	9	Exp 1.1 - 1.8, Exp 2.1 - 2.8
MuP10	3	P			Exp 1.1 - 1.8
MuP11	4	C1			App 1.1 - 1.8
MuP12	4	C2			App 1.2 - 1.8
MuP13	3		P	P	Exp 1.2 - 1.8
MuP14	4		C1		App 1.2 - 1.8
MuP15	4		C2		App 1.2 - 1.8
MuP16	4			C1	App 1.2 - 1.8
MuP17	4			C2	App 1.2 - 1.8
MuP18	4		1	1	Exp 1.1 - 1.8, App 1.1 - 1.8
MuP19	4		2	2	Exp 1.1 - 1.8
MuP20	4		3	3	"
MuP21	4		4	4	Exp 1.1 - 1.8
MuP22	4		5	5	Exp 2.1 - 1.8
MuP23	4		6	6	"
MuP24	4		7	7	"
MuP25	4		8	8	"
MuP26	4		9	9	Exp 1.1 - 1.8, Exp 2.1 - 2.8

Key:
MuP11: Music Process coded item 11

Key:
App: Appreciating strand
Exp: Expressing strand
1.4: sub-strand 1, level 4

There were differences between the groups in time allocations, as primary school students can not stay on task as long as Year 10 students. The stimulus material used at Year 3 was different from that used at Years 7 and 10 as the interpretation of a painting, which was required from the two higher Year groups, was considered too difficult for Year 3 children.

The stimulus used at Year 3 was a videotaped excerpt from a newsreel depicting the calm before a storm, the build-up and climax of the storm and the stillness of the devastation after the storm. This structure was intended to guide the students into using basic form; that is, beginning, middle and end, in their compositions. In order to acquaint students with the points for assessment, they were supplied with information entitled, "Ideas to help you make your composition." The time specified for the Year 3 test was approximately 85 minutes, comprising of approximately 40 minutes for the warm-up, viewing the stimulus, brainstorming, group planning and group rehearsal. Following a short recess or lunch break, the remaining 45 minutes was used for the final rehearsal, the group performance, the student critique of their performance, and collection of materials (see Appendix viii; *Year 3 Process test Administration Procedures*).

The structure for the Year 7 Process test was similar to that used at Year 3, except that the time allocation for the Year 7 Process test was 110 minutes. The first 55 minutes was allocated to the warm-up, brainstorming and discussion, group planning and rehearsal. After a short break, the second 55 minutes was used for the final rehearsal, group performance, student critique

and collection of materials (see Appendix ix; *Year 7 Process Test Administration Procedures*). The stimulus for Year 7 was a painting entitled *Heaven & Earth* (Pericles, 1978) which was selected to provide some contrast, intended to assist students in their use of form. Year 7 students were supplied with a more detailed guide than that provided at Year 3, to acquaint them with points for assessment. This guide, entitled "Ideas to help you make your composition," used musical terminology to describe the elements students were expected to include in their compositions. This terminology, however, was accompanied by explanations of meaning; for instance, "harmony – two or more sounds heard together" (see Appendix ix; *Year 7 Process Test Administration Procedures*).

The structure for the Year 10 test was similar to that used at Years 3 and 7, except that, at Year 10, the time allocation was 115 minutes. There was no break in the time allocation as, unlike primary school children, Year 10 students are expected to work for this period of time without a break. The "Ideas to help you make your composition" page described the same musical elements as those for Year 7 except that there was no explanation of the musical terminology (see Appendix x; *Year 10 Process Test Administration Procedures*). The stimulus for Year 10 was the same painting, *Heaven & Earth* (Pericles, 1978), as that used for Year 7.

CHAPTER FOUR

MEASUREMENT AND INSTRUMENTS

Assessment

The term 'assessment', in this study, is used for the process of making judgements about student performance, both in relation to the acquisition of an understanding and in the detection of that understanding in order to facilitate further learning. Importantly, the focus of assessment today has shifted from teacher input to student outcome and teachers are expected to use observational skills to collect appropriate evidence of these outcomes (Lehman, 1996, p.1; School Examinations Assessment Council, 1991, p.1; Masters & Forster, 1996, p.1; Hanley, 1992, p.22; McArthur, 1987, p.xiv). Subjectivity, which was regarded as completely unacceptable in assessment strategies in the past is now more accepted as a part of any assessment and this is relevant to measurement of achievement in *The Arts*. However, there is need for caution when allowing for subjectivity and teacher judgement and it will be important, when using more subjective assessment techniques, to develop more objective criteria to be used as descriptors of student development. For example, if students are asked to "Explain how the music ends," they have the opportunity to discuss the mood, the elements and/or the instrumentation. The marking key must allow for partial credit for simple explanations of the music becoming louder and using more instruments, as well as allowing more credit for discussing elements, instrumentation and orchestration in a technical way using music-specific language. Descriptors of student development along a continuum of proficiency in *The Arts* involved the development of clear, precise marking keys based on information gathered during extensive trialling of instruments. This was an important step in the process of test development.

These marking keys provide information on students' knowledge from a qualitative point of view. However, the data must be aggregated or used to create a scale and then interpreted to form a scale of measurement. Seven criteria for measurement have been suggested by Wright and Masters (1981). They are first, an evaluation of whether each item functions as intended; second, an estimation of the relative position (difficulty) of each valid item along the scale; third, an evaluation of whether each person's responses form a valid response pattern; fourth, an estimation of each person's relative score (attitude or achievement) on a scale; fifth, the person scores and the item scores must fit together on a common scale defined by the items and they must share a constant interval from one end of the scale to the other so that their numerical values mark off the scale in a linear way; sixth, the numerical values should be accompanied by standard errors which indicate the precision of the measurements on the scale; and seventh, the items should remain similar in their function and meaning from person to person and group to group so that they are seen as stable and useful measures.

Evaluation

There are two basic ways of interpreting student achievement. The first, Norm-referencing, as described in pp 50-51 of this study was considered to be an unsuitable method of reporting because the test score does not indicate a student's ability. In other words, it does not tell you what he/she can or cannot actually do (Glaser & Nitko, 1971 in Griffin & Nix, 1991, p.88; Ebel & Frisbie, 1986, p.27). It was also considered unsuitable as there was no intention to make comparisons between individuals, schools and school systems (Griffin & Nix, 1991, p.88). Traditionally, items used for system level norm-referenced testing have been multiple choice or single response, machine scoreable items which test purely for knowledge of content and which do not actually

demonstrate student performance. Although this type of testing has been used for many years in Western Australian primary and secondary schools for identifying students for specialist music programs (Bentley, 1966; Edwin & Gordon, 1989), they have been of no value in describing what students can actually do in the area of music.

“Criterion-referenced” testing, which enables specific skills and processes to be described, was used in this study. Items were constructed that are relevant to the learning outcomes to be measured so that a description of students' specific skills could be obtained (Gronlund & Linn, 1990, p.15; Ebel & Frisbie, 1987, p.27). An advantage of criterion-referenced testing is that it provides a description of the breadth and depth of the competency and thus aids the teaching and learning process at both the classroom and system level. When using criterion-referenced testing, however, care needs to be taken to ensure that tests are not just a meaningless, unrelated series of trivial tasks. They must be relevant to the prespecified goals of the discipline and must be ordered along a continuum of proficiency which demonstrates students' progress (Griffin & Nix, 1991, p.77).

Item response theory

Item Response theory models the relationship between a person's level on the trait being measured by a test and the person's response to the test item or question (Lord, 1980). Item response theory makes assumptions about a person's behaviour when responding to test items. For example it is possible to describe an item independently of any sample of people who might respond to the item. Second, it is possible to characterise a person independently of any sample of questions administered to the person and, third, it is possible to predict properties of a test in advance. Item response

theory assumes that it is possible to describe mathematically the relationship between a person's trait level and their performance on a test item (Keeves, 1997, p.836).

Information gained from the item is used to locate students on a continuum of developing proficiency or possession of a skill. Test items can vary from multiple-choice, short answer, extended answer, or essays, to a task such as playing an instrument or producing a piece of artwork as long as it is related to an attribute or trait along the developmental continuum and item response theory is concerned with the probabilistic relationship between the student's performance on the task and the amount of the attribute that the student possesses (Andrich, 1978a, p.451, 1978b, p.561, 1995, p.57; Griffin & Nix 1991, p.97; Ebel & Frisbie, 1987, p.335; Gronlund & Linn, 1990, p.467).

In a probabilistic model, allowance is made for the possibility of a person being wrong on an easy item and right on a hard item, so that if the person's ability is greater than the item's difficulty we would expect that the probability of the person being correct would be greater than 0.50. Similarly, if the person's ability is below the item's difficulty we would expect the probability of a correct response to be less than 0.50. If a person's ability and the item's difficulty are at the same point on the scale, the probability of a successful response would also be 0.50. This analysis provides the opportunity to relate the probability of a correct response to the difference between the person's ability and the item's difficulty (Andrich, 1978b, p.561).

Item response models are being used more and more in national and statewide assessments (Beaton 1985; Eggen, 1990; Masters, 1990; McGaw et al., 1989; Phillips et al., 1989 in Masters, 1993). Application of item response theory include John Keeves' (1991) comparison of international

performances on the First and Second Science Studies and Warwick Elley's (1992) report of the Reading Literacy Study (Masters, 1993 p.1) and Monitoring Standards in Education assessments in English; mathematics; science; studies of society and environments; and health and physical education (Education Department of Western Australia, 1995, p.3).

Rasch models

The Simple Logistic Model, the original and simplest form of item response theory, was developed by Danish mathematician Georg Rasch (1960/1980; Andrich, 1978a, p.451; 1978b, p.561; Griffin & Nix, 1991, p.90; McArthur, 1987, p.111). A basic requirement central to a Rasch model of item response theory is that of specific objectivity, which means that the difference between person abilities on the scale is independent of the difficulties of the items, and the difference between item difficulties is independent of the abilities of the persons. This ordering permits a parameterization of people and tasks that fits the simple logistic model of Rasch (Andrich, 1978b, p.561; 1979, p.188; McArthur, 1987, p.111). Rasch's (1980) simple logistic model is used where dichotomous responses are required, that is, answers are right or wrong. Andrich (1988a, p. 636) developed this model to an extended logistic model which allows for polychotomous or partial credit responses and it is Andrich's (19881) extended logistic model that is used in this study. Both Rasch's (1980) model and Andrich's (1988a) extended logistic model are one parameter models which means there is neither a guessing factor nor a discrimination parameter included.

The basic assumption of Rasch's (1980) model has some important implications, one of which is local independence, that is, that a person's

success in responding to an item should not be influenced by his/her having successfully responded to another item. Similarly, local independence assumes that the response of a student to an item is not affected by responses given by others to the same item.

Another implication of the basic assumption of the model is equality of discrimination, that is the ordering of the items in terms of difficulty must be the same for persons of lower ability as for persons of higher ability. Unidimensionality is also a consequence of the basic assumption (Andrich, 1978b, p.562; McArthur, 1987, p.112). This means that, in order to represent the interaction of person and task, a unidimensional test measures only a single trait or ability. According to Andrich (1989, p.14), these specifications for the Rasch (1980) model should no longer be referred to as assumptions as they are, in fact, requirements for the data to produce measurements, and not assumptions about the model.

If data conform to the model, it is possible to use different overlapping sets of questions with different groups of students or to delete questions which are problematic in some tests while retaining them in others without compromising the comparability of student achievement measures (Masters, 1994, p.1; Honeyman 1996, p.3).

The Extended Logistic Model of Rasch (Andrich, 1978b, 1988a, 1988b; Wright, 1985) was used to create a scale based on the 159 items in the music tests. The scale is based on the log odds (called logits) of students' responses to the items; that is, item difficulty levels. The items are ordered along the scale at interval measurement levels from easiest (that is, those with negative logit values), to most difficult (that is, those with positive logit values). The Rasch method produces scale-free student ability measures and sample-

free item difficulty measures (Andrich, 1988b; Wright and Masters, 1982). This means the differences between student ability measures and item difficulties are expected to be sample independent (Waugh, 1996, p.9). The consistency of student ability measures were checked and the scale score needed for a 50 per cent chance of a student getting an item correct was calculated. These scale scores are the threshold values. The thresholds are conceptualised as a set of boundaries between categories and indicate the change in probability of a response occurring in two adjacent categories. The threshold values are calculated in logits and are ordered to represent the increasing difficulty of items. The scale score for this study was further modified to show a 70 percent chance of a student getting an item correct. The reason for this modification was to gain a more reliable indication that a student was working at the level of performance calculated in accordance with *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996).

The Extended Logistic Model of Rasch

An extended logistic model of Rasch was developed by Andrich (1988a, p.363) to move item response theory beyond dichotomous responses to include more meaningful tasks. Models for graded items with three or more response categories, referred to as polychotomous response items, are developed and used as a basis for comparing and interpreting performances across groups, and from one instrument to another (Andrich 1988a, b; Samejima, 1969; Andrich, 1978b; Masters, 1982, in Griffin & Nix 1991, p.99; Masters 1994a, b).

The polychotomous response model, allowing for partial credits (Andrich, 1985, 1985a 1988a), is appropriate for assessing *The Arts*, which

incorporates performance, analysis and response to stimuli. Partial credit allows for identification of what students know or can do rather than just allowing for a right or wrong answer. For example, it is possible for students to respond to an open-ended item which requires their interpretation of an excerpt of music, in several categories from a simple reference to the dynamics, to an extremely complex explanation of the elements, the mood created, and the instrumentation, using music-specific language.

The differences between raw scores and person ability measures

Raw scores cannot reliably be used to compare ability among students because they do not constitute any standardised form of measurement (Andrich, 1985, 1985a, 1988a, 1988b; McArthur, 1987, p.99). For instance, it is not always true that the ability difference between two students scoring 15 and 20 is the same as the ability difference between two students scoring 95 and 100. It is also difficult to accept that there is a true zero on a scale of ability. It may be argued that some ability exists but the test failed to ask a sufficiently easy question. It is also difficult to accept that a perfect score indicates that a student has perfect or complete attainment of the subject being measured. A score of zero may not imply a complete absence of ability and a perfect score does not imply perfect ability (Honeyman, 1996, p.2). For these reasons, to compare the performances of students on the basis of ability, raw scores must be transformed onto an interval scale. A Rasch model, by producing estimates of difficulty of items on the same scale as estimates of ability, allows calculation of the probability that a particular student will correctly answer a particular item (Andrich, 1978b, p.561). Item difficulties and person abilities are arranged from negative values through to positive values and there are no theoretical upper or lower limits, but in practice the values usually fall within six units of the origin (Honeyman, 1996,

p.4). This unit of measure, for both item difficulty and person ability, is called the logit (log odds of answering correctly).

The choice of a unit for reporting is an arbitrary matter and, for this study, logits were translated to a scale from 1 to 800 for purposes of reporting. This scale was selected to coincide with the eight levels of achievement in *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996).

The RUMM program

Data collected for this study were analysed using a computer program entitled RUMM: Rasch Unidimensional Measurement Models (Andrich, Luo & Sheridan, 1996). RUMM. Rasch Unidimensional Measurement Models is a sophisticated data analysis program in which data for analysis is considered in two broad sections; Item responses and Person ability, and allows for analysis of single response and extended response category data, that is, dichotomous and polychotomous items. It also provides test-of-fit in the form of Item-trait interaction and Item-person interaction. Other features of the program include an estimation of item and person parameters, threshold estimates, category response frequencies and individual person fit (Andrich, Luo & Sheridan, 1996). One of the special features of the program is its ability to provide a range of graphical displays including a Guttman pattern, category probability curves, item characteristic curves, person frequency distribution and person-item frequency distribution. Application and use of these features were extremely useful in examining trial data for the modification of items and marking keys before final item selection.

Item fit

Performance on a particular task on a particular day can be affected by many variables such as a student being tired or unwell, an environment being noisy and distracting, or a task being badly constructed in a way to disadvantage a particular group, and so there are times when, based on past performance, we expect a student to correctly answer an item but we observe an incorrect response. We say that an item does not 'fit' when student responses to items have been examined for their consistency with the notion of a single trait and it is found that an item does not work together with the other items to define a single variable. This anomaly may exist for a number of reasons, including an inaccurate marking key, a badly worded item or a source of bias that is advantageous to the less able student. It may also be that the item simply does not measure the same construct that all other items are measuring. A reversal of thresholds may be an indicator that an item does not 'fit' and when this occurs, the source of this anomaly is sought. If it cannot be located or rectified, the test developer usually discards the item.

The RUMM program has the capacity to check that student responses fit the measurement model according to strict criteria as described by Wright & Masters (1982) and Wright (1985). The fit statistics comprise of weighted and unweighted mean squares that can be approximately normalised. The normalised statistics are called *infit t* and *outfit t*, and the mean is zero with a standard deviation close to 1, with values of *t* outside +2 and -2 indicating significant departure from the expectation of the model. A fit mean square of 1 plus *x* indicates 100*x* per cent more variation between the observed and predicted response patterns than would be expected if the data fit the model. Similarly, a fit mean square of 1 minus *x* indicates 100*x* per cent less variation

between the observed and predicted response patterns than would be expected if the data fit the model.

For this study items were examined individually, initially by identifying the least fitting items from the fit order and, for items for which this statistic was extreme, the Item Characteristic Curve was examined. In cases where any anomalies appeared, the observed and expected values were examined for that item. A few items showed relatively low discrimination, although still positive. In each case, in the context of the test validity, it was decided that its exclusion would only minimally affect the test reliability and overall fit, and, because the item contributed to the integrity of the test, it was decided not to delete any items.

All polychotomous items were examined for reverse thresholds and Category Probability Curves were examined to facilitate decisions about whether categories needed rescoring or collapsing. A total of four items from the analysis tests and four items from the process tests were rescored. None were discarded. After rescoring of these items all items from both the Analysis test and the Process test fit the model, that is, threshold values were correctly ordered from a simple response through to the most difficult response.

An advantage of a Rasch model of analysis is that, providing test items are derived from the bank of items valid for the test, different students can answer different sets of questions and still be placed on a common scale. This is useful in cases of missing data where not all students have completed every item (Honeyman, 1996, p.4).

Development of the Marking Keys

Analysis Marking Key

In order to ascertain categories for the partial credit model to be used to mark the analysis items, it was necessary to trial the items with children in Western Australian classrooms. This was done by asking teachers to volunteer to administer the tests to their classes. After collection of the materials, the extended-answer test items were examined one by one to determine what types of responses students were likely to give. These were then collapsed into three or four general categories for each question, examined against *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) and categorised in order of difficulty. Answers which were wrong, made no sense, or were tautological, were given '0' marks; answers which provided little information were given 1 mark; those which provided more were given 2 marks and so on. Items usually had between two and four categories. In four of the items, after the analysis of the data, some categories were not discriminating sufficiently from each other. In these cases, categories had to be collapsed and the items rescored. This will be described in detail in Chapter Six (see Appendices iii, v and vii for Music Analysis Marking Keys).

Process Marking Key

Experimentation was carried out to establish the most effective structure for marking keys. As mentioned previously, each group's performance was

videotaped so markers could watch it as often as necessary to allocate the appropriate mark. Again, the trial material was used to finalise the most effective method of marking.

First, to reflect the development of skills, a line of continuum was developed in a style similar to a Likert scale. For instance, the marker was prompted with the question, "How effectively has the student's artwork communicated his/her ideas?" Along a continuous line across the page were three vertical marks. Under the first mark was the indicator, "not very effectively" with the descriptors; *no mood evident, no evidence of form, no use of musical elements, and lacks confidence*. Under the middle mark was "somewhat effectively" with the descriptors; *suggests a mood, some evidence of musical elements, and some confidence shown*. Under the third mark was "very effectively" with the descriptors; *clearly shows mood, makes use of musical elements such as harmony, rhythm, makes good use of instruments, music has a form and confident music*. The problem with this method was the tendency for markers to be inclined to allocate a level in between the indicators. An attempt was then made to divide the line into smaller degrees with 20 marks along the continuum so that levels between the descriptors could be measured. This resulted in markers tending to count the marks and give a score out of 20. This was detrimental to the notion of assessing and describing what students can actually do, and reverted back to the old method of allocating a numerical score. It appeared that using this style of marking did not fit with the concept of the vertical progression of student achievement

described in the outcomes framework, and so experimentation was carried out to design a marking key in a vertical, rather than a horizontal format.

Finally, a method known colloquially as a 'marking tree' was developed. A prompt question to the marker, such as "How effectively has the group used expression?" was followed by a sequential, vertical list of competency levels matched to a mark allocation. For instance, '0' mark for no evidence – *no expression – even sound, all loud or all soft*, 1 mark for beginning to develop – *slight changes in dynamics – loud/soft*, 2 marks for sound development – *obvious variation in dynamics, tempo and/or melody in an attempt to reflect mood*, 3 marks for well developed – *effective use of dynamics, tempo, rhythm, melody, harmony, tone, etc to reflect mood – some evidence of organisation in planning as well as performance*, 4 marks for highly developed – *exceptional use of elements to create a pleasing sense of expression which clearly conveys mood – inclusion of appropriate variety of dynamics, tempo, rhythm, melody, harmony, tone, texture, legato, staccato etc – evidence of organisation/leadership in planning and performance*. Using this structure, markers could not mark between the descriptors and had to allocate the one which most closely reflected the student's performance. (see Appendix xi; Music Process Marking Key).

Training the markers

Marker reliability was considered to be of vital importance in the process of gathering data. The volume of the completed tasks required a marking complement of seven markers for the Music Analysis and twelve markers for

the Music Process tests and a full day's training was undertaken in these two areas. Expressions of interest were called in schools for markers and those selected were experienced music teachers from either the primary or secondary sector who indicated a strong interest and enthusiasm for participating in the marking process. Markers were required to mark at all of the three year groups, Year 3, Year 7 and Year 10, so they could get a 'feel' for the development of skills and levels from Year 3 through to Year 10 in order to accurately assess link items on a developmental scale.

Marking the Music analysis tests

In the case of Music Analysis tests, exemplars demonstrating a variety of abilities were selected from student trial materials and photocopied for each marker. Markers were issued with marking keys which had been developed for each of the Year levels (see "Development of the marking keys"). It was explained to markers that some items were dichotomous, meaning they were right or wrong, thus there were only two categories; zero or one. Other items, however, were polychotomous, meaning there were several levels at which students could respond. This meant there could be up to five category levels in an item. Each marker examined the first item in test one and allocated an item category level that they interpreted from the marking key. A discussion of the allocated category level then occurred and justification of category levels was verbalised and discussed by markers until agreement was reached. It was emphasised that levels were not related to *Student Outcome Statements* (Education Department of Western Australia, 1996) levels. They had been designed to identify varying levels of skill that could be demonstrated in a student's response to a single item. The process of allocating category levels was continued for each item over a series of tests for each of the Year groups and where necessary, after consensus among the

markers, modifications, alterations, additions or deletions of category levels were effected.

When markers had reached reasonable consensus as to the category levels allocated to items using this method, they were asked to select an individual test paper at random and to mark the whole paper, allocating a category level to each item according to their interpretation of the marking key. They then passed the paper to the next person for marking and when each marker had marked all seven papers, a check and discussion of category levels allocated to each item took place until consensus of category levels was reached.

At the completion of the day, markers had marked a series of common tests and were confident they had reached a unified understanding of the item categories and their levels and a clear understanding of interpretation of the marking key. However, they exchanged telephone numbers in order to make contact to discuss any unusual or difficult examples that may not have emerged during marker training. Markers were requested to rotate Year 3, 7 and 10 tests when marking in order not to lose track of the development of levels of skill through the year groups.

Marking the Music Process tests

Marking of the Process tests for the three performing arts of dance, drama and music involved viewing of students' performances on video in relation to individual plans, group planning sheets and student appraisals. Exemplars of students' planning sheets and appraisals were extracted from the sample and photocopied, and markers worked in pairs with all markers viewing and marking the same video-taped performance and student material.

It was necessary to play the video several times during which time each pair of markers viewed and discussed students' plans, performances and appraisals of their performances. All items were polychotomous as there were several levels at which groups were performing and marking key categories gave descriptions of the levels. When the marking of each student group was completed, discussion took place among all markers, with each pair giving explanations to justify category levels allocated. The next video-taped student performance was then marked in the same way with markers working in pairs. This was followed by a whole group discussion and any alterations, modifications, deletions or additions to marking key categories were effected if and where necessary. This procedure was repeated throughout the day, rotating through Year groups 3, 7 and 10. At the end of the day markers felt confident they had reached an understanding of 'levelness' of performance as outlined in item categories.

Although markers were confident about their understandings, more reliability in relation to markers' use of the marking keys would have been attained if some or all of the tests could have been 'double marked'. Unfortunately, budgeting constraints did not provide the opportunity for double marking of tests.

CHAPTER FIVE

SAMPLE, DATA COLLECTION AND PRELIMINARY ANALYSIS

The sample for Years 3 and 7

As mentioned previously, the Analysis and Process tests for all five disciplines of *The Arts* were designed to reflect good classroom practice. Both tests were administered to all students so that levels of achievement could be measured in their responses to the musical works of others (Analysis test), as well as in the exploration and creation of their own (Process test). In Western Australian classrooms, activities which involve exploring and creating in the performing arts are primarily undertaken by students in group situations. To achieve these group situations for the Process test, it was necessary to sample whole classes, rather than individual students. Although the Analysis test was a paper and pencil test undertaken on an individual basis, the same whole classes of students completed both tests in order to achieve a measure across all strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). It was also less disruptive to schools to administer the tests to whole classes, rather than to individually selected students withdrawn from classes.

Forty classes of students from each of Years 3 and 7 were selected, in a stratified random sample, to be tested in each of *The Arts* disciplines from government primary schools. Five steps were taken to achieve this. First, a list of Western Australian Government primary schools and student numbers

from the Education Department's Information directorate was obtained. Second, the list was ranked in size according to student numbers, and schools with minimum class numbers of less than six in Year 3 or Year 7 were withdrawn from the list. This was done so that the recommended student group size of three, four or five for the Process test could be achieved, and group interaction could occur. Third, in order to obtain the maximum sample size, the largest 400 schools were selected from the list. Fourth, the list of 400 schools was divided into two halves by selecting every second school. A decision had been made not to ask schools to test in *The Arts* at more than one Year level, so one half of the list was allocated to Year 3 and the second half to Year 7. Fifth, the 200 schools on the Year 3 list were randomly divided into five sections; that is, one for each of the disciplines; dance, drama, media music and visual arts. This was done by counting 1-5 repeatedly down the list to ensure that the school sizes were distributed evenly across the five disciplines. This process was then repeated for the 200 schools on the Year 7 list.

The final sample for music testing at Year 3 consisted of 40 classes, providing a total of 946 students. This compares with a total number of Year 3 students in Government schools in 1996 of 20,661. Of the 946 students tested, 426 were identified as girls and 486 were identified as boys. There were 34 students who did not state their gender. Other sub-groups identified in the sample were Aboriginal Torres Strait Islander students, of which there

was a total of 59 at Year 3, and Non-English Speaking background students, of which there was a total of 122 at Year 3.

The final sample for music at Year 7 consisted of 40 classes, providing a total of 921 students. This compares with a total number of Year 7 students in Government schools in 1996 of 20,524. Of the 921 students tested, 397 were identified as girls and 487 were identified as boys, with 37 students not stating their gender. The total number of Aboriginal Torres Strait Islander students identified in the sample at Year 7 was 44, and the total number of Non-English speaking background students was 114. The overall total of primary school music tests submitted for marking at Years 3 and 7 was 1,867.

The sample for Year 10

Because *the Arts* are treated as optional subjects in Western Australian secondary schools, it was not possible to draw a truly random sample. To avoid excessive disruption to schools, a decision was made to test only one discipline per school, and limited numbers of Year 10 students undertaking arts options reduced the number of Year 10 students available for sampling. The following table demonstrates the limitations to the Year 10 students available for the Arts sample:

Table 5.1: Arts enrolments for Year 10, Semester one, 1996

<i>Year 10 Arts Enrolments, 1996</i>				
<i>Discipline</i>			Total enrolled	Total population
Dance	n		1 974	17 140
	% of population		11.5	
Drama	n		3 519	17 140
	% of population		20.5	
Media	n		1 938	17 140
	% of population		11.3	
Music	n		1 374	17 140
	% of population		8.0	
Vis. Art	n		5812	17 140
	% of population		33.9	

Twenty classes of Year 10 students, for testing in each of the five disciplines were selected from Government secondary schools, in a stratified sample, which was selected as randomly as possible within the constraints detailed above. This was done by obtaining a list of Western Australian Government secondary schools, including numbers of students studying each of the five disciplines. It became apparent that the numbers of students studying *The Arts* varied significantly from one discipline to the other, which obviated the option of simply dividing the list into five equal parts. There was an added problem in that the number of disciplines being offered by schools varied, and an alarming number of schools offered none at all. The following table demonstrates the limited numbers of secondary schools offering the complete range of Arts disciplines at Year 10.

Table 5.2: Numbers of schools offering the range of Arts disciplines

<i>No. of arts disciplines offered in Year 10</i>	<i>No. of schools</i>
5	32
4	25
3	19
2	20
1	17
0	58

Some large senior high schools offer specialist courses in arts disciplines and it was decided that, in order to ensure a full range of abilities, some of these students should be included in the sample. Two specialist schools were therefore included in the sample for each discipline. These schools were then removed from the selection list.

The number of schools offering the discipline of dance was lowest so the sample for testing in dance was drawn first by selecting every fifth school from the list. If the fifth school did not offer dance, the next one on the list that offered dance was selected. Because each school was only being tested in one discipline, these schools were then withdrawn from the list, even though many of them offered one or more other arts options. The second lowest enrolment for arts subjects was music, and so the same procedure was used to draw the music sample, with these schools also subsequently being withdrawn from the list. This process was repeated for media, drama and visual arts in that order, with visual arts being the most commonly offered arts option at Year 10.

It is realised that the constraints in student access to arts resulted in some limitations to the Year 10 sample. However, an effort was made to make the sample as random as possible within the circumstances. A further limitation in the Year 10 arts sample was the low proportion of Year 10 boys undertaking arts options. This meant that it was not possible to obtain a gender balance in the sample. The following data demonstrates the significant difference between the proportion of girls and boys undertaking arts options at Year 10 during 1996.

Table 5.3. Arts enrolments by gender, Semester one, 1996

<i>Year 10 Arts Enrolments</i>					
DISCIPLINE		Females enrolled	Males enrolled	Total enrolled	Total population
Dance	n	1944	30	1974	17140
	% of enrolled	98.5	1.5		
Drama	n	2516	1003	3519	17140
	% of enrolled	71.5	28.5		
	%of population	14.7	5.9	20.5	
Media	n	1084	854	1938	17140
	% of enrolled	55.9	44.1		
	%of population	6.3	5	11.3	
Music	n	791	583	1374	17140
	% of enrolled	57.6	42.4		
	%of population	4.6	3.4	8	
Vis. Art	n	3759	2053	5812	17140
	% of enrolled	64.7	35.3		
	%of population	21.9	12	33.9	

NOTE: Population refers to the number of Year 10 students currently enrolled in government schools in Western Australia.

The final sample for music testing at Year 10 consisted of 20 classes, providing a total of 324 students. Of these, 172 were identified as girls and

139 were identified as boys, with thirteen students not stating their gender. There were 17 Aboriginal Torres Strait Islander students and 41 Non-English speaking background students identified in the sample.

Data Collection

The tests were administered in school classrooms that reflected students' usual learning environments. In primary schools where a specialist music teacher normally taught music classes, the music teacher administered the tests. In primary schools where there was no music specialist, the teacher who normally taught music to the class administered the tests. This was usually the classroom teacher. In secondary schools, the specialist music teacher administered the tests.

In order to reduce variability in administration of the tests, explicit administration instructions were distributed to teachers. These included the overall time allocation for the tests, as well as times to be apportioned for specific sections of the tests. Instructions were also given as to what the teacher was required to prepare before administering tests. For the Process test, this included the viewing of a Teacher Training Video demonstrating the warm-up and group work.

Teachers were instructed to help students who were having difficulty following instructions or reading questions, but were asked, emphatically, not to help them with the actual task. Standardised wording for the teacher's

verbal instructions to the students were provided and teachers were instructed not to deviate from this, except to clarify understanding. At the Year 3 level, teachers were asked to read questions aloud while students followed, whereas at Years 7 and 10 they gave students time to read the questions themselves, assisting only when requested (See Appendix xii, Administration Procedures Year 10 Music Analysis).

As mentioned in Chapter three, 'Structure of the Analysis Test,' The Analysis stimulus audio tapes were divided into parts to correspond with the parts in the test paper, with the voice on the tape instructing when to pause the tape.

For the Process test, teachers were instructed to organise the students into groups of four prior to testing. If numbers were uneven, groups of three or five were allowed. Some control over group selection was exercised by providing teachers with a numbered list on which an asterisk had been placed beside every fourth number. Teachers were then asked to copy students' names directly from their classroom attendance roll onto the list. Each 'asterisked' student became the nucleus of a group and teachers then organised groups around these students to create the most suitable working groups.

Guidelines for the administration of the Process test were very explicit and teachers were asked to adhere rigidly to the verbal instructions provided during the time prior to the group planning and rehearsal session. During the

group planning and rehearsal, teachers were asked to move around the room, supervising as they would in a normal classroom situation, dealing with questions or problems, or clarifying understanding when necessary, but without actually helping students with the task.

It was important to have good quality videotapes for the central marking of performances. To ensure that teachers supervised classes adequately during videotaping of performances, they were requested to work in collaboration with a support teacher or student to operate the video camera. Clear instructions as to the positioning of the camera, the background, the size of the performing area and identification of groups were provided (See appendix ix Year 7 Process test Administration Procedures). These instructions minimised the potential for markers being influenced by either professionally produced videotapes or poor quality ones.

Process for student ability level estimates

After raw scores had been obtained for both the Analysis test and the Process test, the RUMM (1996) program was used to analyse the data from the Analysis and Process tests separately, and to examine the test validity, reliability and the power of the test of fit. An analysis combining items from the two tests was then undertaken giving detailed consideration to individual item statistics, including fit statistics and the item characteristic curves. The reliability and the power of test of fit were maximised by combining the Analysis and Process items in a single test, with the Power of

test of fit being recorded as 'Excellent', and with a reliability of 0.928 (proportion of observed variance considered to be true). A decision was made, therefore, to combine the items from both tests into one, thus providing the opportunity to have all items on a single scale, and hence a single ability estimate for each person. An examination of scores, together with individual student scripts was carried out to obtain level cut-off points and to estimate student levels (see figure 5.1: Process for student ability level estimates).

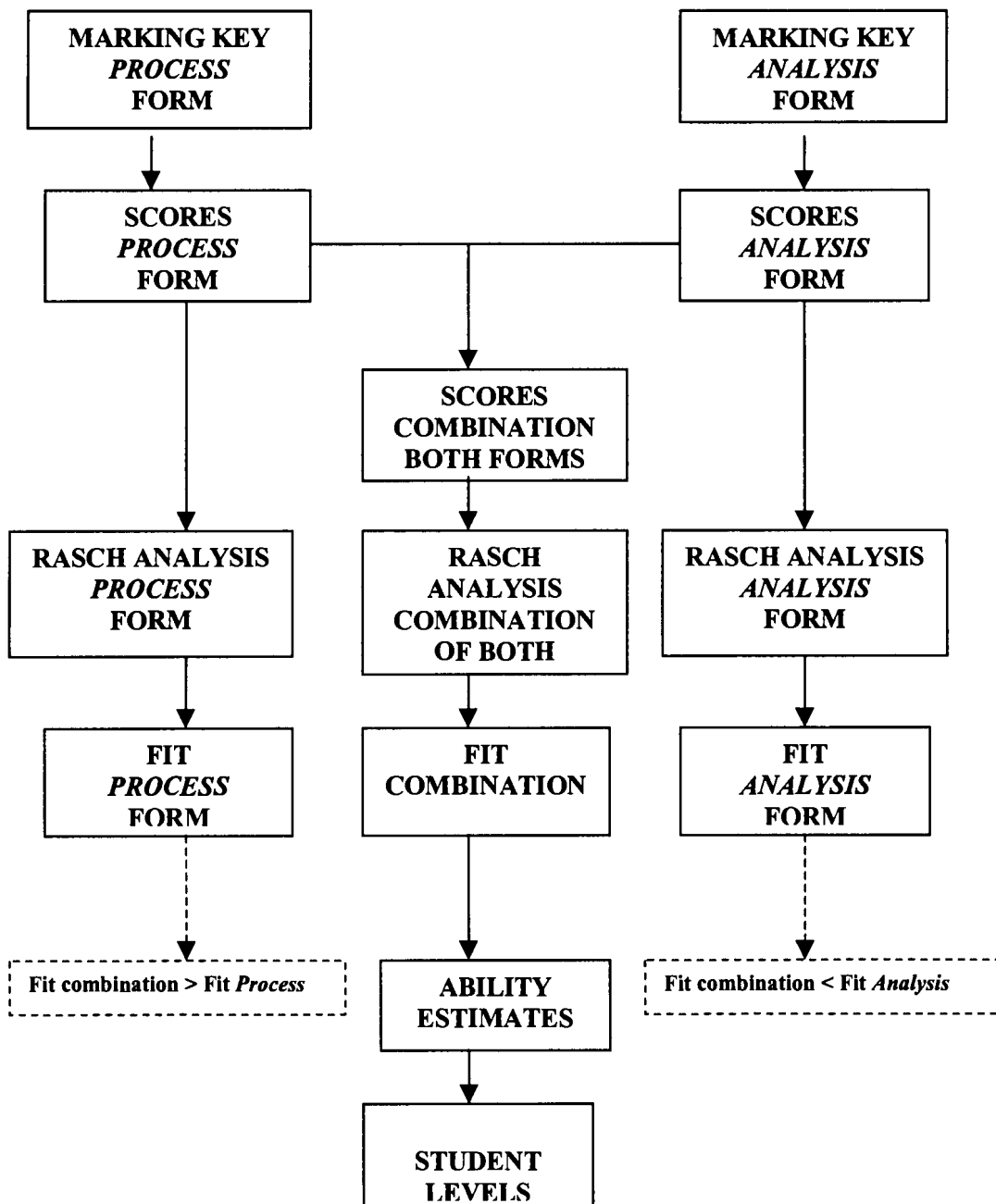


Figure 5.1: Process for student ability level estimates

Next, the combined Analysis and Process items were examined individually for item fit. A test-of-fit was used to assess the residual between the expected estimate and actual values for each student-item combination across all items or, conversely, across all students for a given item. The chi-square statistic was used as a general guide, although, in view of the large sample size, this statistic was treated with caution (Mercer, 1997, p.2). Where the chi-square statistic was extreme, the Item Characteristic Curve was examined and where anomalies existed, the observed and expected values were examined for individual groups for that item. In cases where items displayed relatively low discrimination, these items were viewed in the context of the test validity. Although discrimination was low, it was positive, and because the items contributed to the integrity of the test, it was decided not to delete them from the final analysis.

The polychotomous items, that is, items with more than two categories, were examined for reverse thresholds or thresholds which had minimal discrimination. The Category Probability Curves were used to assist in the rescoring, or collapsing, of categories. Rescoring was carried out on four Analysis items and four Process items before the final item and person parameters were estimated. The RUMM program was then used to reanalyse the data with the rescored items, and person and item parameters were extracted for calculation of a numerical score from 0 to 800, which would be known as the *Performance scale*. Descriptive statistics and inferential statistics were also calculated (Mercer, 1997, p.4).

Student performance

Students generally engaged well with the tasks, providing a range of rich information on their knowledge of musical elements and their responses to musical works. They interpreted works in terms of style and mood and did not hesitate to express their preferences.

At the Year 3 level, students recognised music heard in their everyday lives and were able to identify it in relation to its purpose and how its purpose affects the way it should be performed. They were able to respond to music in terms of associated movement, and recognised basic sound characteristics such as ascending/descending, pitch or tempo changes. They reflected upon music works, noting particular features including melody, instruments used, form and expression. They expressed their like or dislike of the music and justified their opinions in relation to the instruments used or basic music elements such as pitch or tempo.

They worked in groups, choosing their own sound sources, to plan, rehearse and produce simple soundscapes to reflect the stimulus provided. In their planning they invented simple signs and symbols to represent basic characteristics of sounds and used them to notate their musical creations. Year 3 students generally did not use conventional musical notation. Most groups attempted to include basic form, that is; a beginning, middle and end, in their pieces. They applied simple critical reflections to their group performances.

Year three students generally used non music-specific language, apart from simple terms such as 'beat'. They talked in emotive terms, such as the music was 'happy', or from personal experiences, such as "I know it was orchestral music because I went to an orchestral concert once and it sounded like that." Their planning and performances rarely displayed organisation or leadership and cues to begin or end were generally given by the teacher.

Year 7 students were able to aurally recognise and describe musical features such as simple rhythmic and melodic patterns, tempo, instrumentation, timbre, dynamics and structure and use and interpret signs and symbols representing pitch, duration of sound and dynamics. They described obvious features such as repetition, form, changes in dynamics and texture, as well as identifying music from another culture and associating characteristics of the music with the style. Some were able to discuss the effect of the music style on Australian culture. They were able to compare and contrast two pieces of music in relation to some aspects of instrumentation, expression and rhythm. They generally used a combination of non music-specific and basic musical language in their discussions.

They co-operated in group situations to plan, rehearse and perform a short, simple, structured musical composition, selecting their own conventional or non-conventional musical sound sources to reflect the stimulus provided. These generally consisted of tuned or untuned percussion instruments, recorder, sounds from the environment, voice and body percussion. As well as non-conventional signs and symbols to represent

sound characteristics, some used conventional musical notation in their planning. An attempt to use texture was evident in some of the works and most groups attempted to include basic form, that is; a beginning, middle and end, in their pieces. There was evidence of organisation and planning to include some form of leadership in many of the performances, with a group member giving cues to begin and end. Planning generally indicated an attempt to reflect the stimulus, and there was some evidence of the characteristics of the stimulus in the use of tonality and/or texture. There was limited evidence of consideration of audience in performances.

Year 10 students were able to listen to a piece of music and discuss how an identified musical element, such as the key, was important and how it was used to create musical mood, tension or purpose. Their discussion of musical elements was more technical and contained more sophisticated use of musical terminology, including chord structure, key, metre and stylistic conventions. They discussed the manner in which musical elements were used to create unity and contrast and the effect of the music in achieving the desired purpose. They compared music from different times, places and cultures, identifying notable differences in musical characteristics. They discussed the effect of specific music styles on Australian culture as well as using stated criteria to give reasons for their preferred performances. Although most students successfully identified and interpreted the musical works to which they listened, very few students actually used the processes of critical analysis to support their judgements.

Year 10 performance, as part of a group, was generally more sensitive and there was evidence of more organisation within the Year 10 groups than within the Year 3 and Year 7 groups. Year 10 students created musical works that captured characteristics of a given stimulus and interpreted elements of pitch, rhythm, dynamics and phrasing in composition. They explored major and minor tonalities, textures, forms, media, and invented a score related to the theme. In most instances, they used conventional notation in their scores. They explored combinations of sounds from the environment, chords, ostinati, and incorporated known structures such as ternary or binary form. There was evidence of the use of a variety of structural devices such as repetition, variation and contrast to create the desired effect in reflecting the stimulus provided. Their selection of combinations of sounds, both conventional and unconventional, included the exploration of electronic sounds. There was obviously more consideration of the audience and purpose and leadership within most groups was evident.

The mean level for each Year group shows a clear pattern of development from Year 3 through Year 7 to Year 10, although there is considerable overlap in performance between the Year groups. For example, the highest achieving 10 per cent of Year 7 students performed above the level demonstrated by approximately 25 per cent of Year 10 students.

Over 80 per cent of Year 3 students demonstrated skills associated with Level 2 outcomes or above (*The Arts Student Outcome Statements*, 1996, pp. 1-3). Around 20 per cent of these demonstrated skills associated with Level 3

outcomes and a small percentage, less than 10 per cent, demonstrated skills above Level 3. Around 20 per cent of students were still demonstrating skills associated with Level 1 outcomes. This indicates that these students, while showing an awareness of music in everyday life, are still drawing on their play and imagination and responding to music in their own way.

Over 55 per cent of Year 7 students demonstrated skills associated with Level 3 outcomes or above (*The Arts Student Outcome Statements*, 1996, pp. 1-3). Of these, approximately 20% demonstrated skills associated with Level 4 outcomes, and less than 10% demonstrated skills above Level 4. Approximately 45 per cent of Year 7 students were still working at Level 2 or below. In view of the fact that, in other learning areas, benchmarks for Year 7 are being set at Level 3, this is cause for some concern. It is, indeed, indicative of the fact that many Western Australian schools are not adequately catering to the needs of students in this learning area. Many are still not using an outcomes-based framework in their learning programs.

Over 80 per cent of Year 10 students demonstrated skills associated with Level 4 outcomes or above (*The Arts Student Outcome Statements*, 1996, pp.1-3). Of this 80 per cent, around 50 per cent of students demonstrated skills associated with Level 5 outcomes or above and, of these, around 30 per cent demonstrated skills associated with Level 6 outcomes or above. Only 10 per cent of Year 10 students demonstrated skills below Level 3. In other learning areas, Level 4 is being set as a minimum benchmark for performance at Year 10. However, as at November, 1998, no decision has

yet been made as to what the minimum benchmarks will be for music in Western Australian government schools. The significant improvement in performance from Year 7 to Year 10 is to be expected in view of the fact that the primary school study sample was drawn randomly from all students and, in secondary schools, a stratified sample was drawn to include only those students undertaking music options.

Patterns of development

The link items in the Analysis tests demonstrate the clear pattern of development from Year 3, through Year 7 to Year 10 that is evident in the overall results. An example of a link item is Analysis item MuA12 (see Table 3.1, p.108) which asked students to “Explain how the music ends”. Students listened to an excerpt of the ending of a piece of music and identified the prominent musical features of the ending in relation to elements such as dynamics, pitch, tempo and instrumentation. The item was open-ended. Of the Year 3 students, 32.9 per cent scored zero, compared with 17.3 per cent at Year 7 and 9.7 per cent at Year 10. Only 0.1 per cent of Year 3 students provided a higher order response to score 3, compared with 0.8 per cent of Year 7 students and 16.0 per cent of Year 10 students (see Tables 5.4, 5.5 and 5.6 hereunder).

Table 5.4: Year 3 Frequency table for Item MUA12

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	330	31.5	32.9	32.9
1.00	588	56.2	58.6	91.5
2.00	83	7.9	8.3	99.8
3.00	1	.1	.1	99.9
4.00	1	.1	.1	100.0
9.00(missing cases)	44	4.2	missing	
Total	1047	100.0	100.0	

Table 5.5: Year 7 Frequency table for Item MUA12

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	166	16.6	17.3	17.3
1.00	569	57.0	59.5	76.8
2.00	214	21.4	22.4	99.2
3.00	8	.8	.8	100.0
4.00	0	0	0	100.0
9.00(missing cases)	41	4.1	missing	
Total	998	100.0	100.0	

Table 5.6: Year 10 Frequency table for Item MUA12

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	34	9.4	9.7	9.7
1.00	106	29.2	30.4	40.1
2.00	142	39.1	40.7	80.8
3.00	56	15.4	16.0	96.8
4.00	11	3.0	3.2	100.0
9.00(missing cases)	14	3.9	missing	
Total	363	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

Analysis item MuA08 (See Table 3.1, p.108) relating to the element of 'beat' (another link item), further demonstrates this pattern of development. While this item is also open-ended, it is more directed than item MuA12. Item MuA12 invited discussion on a variety of elements, where MuA08 was confined to characteristics associated with the change in beat from a 4/4 metre to a 3/4 metre. Many of the lower achieving students confused 'beat' with tempo or dynamics and referred to the music 'getting slower' or 'getting

softer.’ 73.2 per cent of Year 3 students, 59 per cent of Year 7 students and 34.3 per cent of Year 10 scored zero on this item. The top score of three was achieved by 17.6 per cent of Year ten students, 1.2 per cent of Year 7 students and 0.4 per cent of Year three students (see Tables 5.7, 5.8 and 5.9 hereunder).

Table 5.7: Year 3 Frequency table for Item MUA08

Value	Frequency	Percent	Valid percent	Cumulative Percent
.00	737	70.4	73.2	73.2
1.00	139	13.3	13.8	87.0
2.00	127	12.1	12.6	99.6
3.00	4	.4	.4	100.0
9.00(missing cases)	40	3.8	missing	
Total	1047	100.0	100.0	

Table 5.8: Year 7 Frequency table for Item MUA08

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	564	56.5	59.0	59.0
1.00	158	15.8	16.5	75.5
2.00	223	22.3	23.3	98.8
3.00	11	1.1	1.2	100.0
9.00(missing cases)	42	4.2	missing	
Total	998	100.0	100.0	

Table 5.9: Year 10 Frequency table for Item MUA08

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	119	32.8	34.3	34.3
1.00	77	21.2	22.2	56.5
2.00	90	24.8	25.9	82.4
3.00	61	16.8	17.6	100.0
9.00(missing cases)	14	3.9	missing	
Total	363	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

Item MuA11 represents an example of a closed, multiple-choice link item. Students were asked to aurally recognise a rhythm pattern and to make

a selection from four visual patterns supplied. The correct pattern was selected by 56.8% of the Year 3 students, 81.5% of Year 7 students and 92.8% of Year 10 students (see tables 5.10, 5.11 and 5.12 hereunder). While there is still a clear pattern of development here, it is interesting to note that the differences between results from Year 3, through Year 7 to Year 10 are not as marked as in the previous two items. This could be an indication of the limitations in the ceiling effect of closed items and the advantages of open-ended items which provide the opportunity for students to achieve to the maximum of their abilities. If students had been able to interpret and describe the rhythm pattern demonstrated, a more significant difference in performance may have been witnessed.

It is also interesting to note that, in most cases, the incorrect responses selected, decreased towards the end of the selections provided, that is, the number of students who chose 'D' in every case is significantly less than the number who chose 'A'. This is possibly an effect of the location of distracters (see tables 5.10, 5.11 and 5.12 hereunder).

Table 5.10: Year 3 Frequency table for Item MUA11

Value	Frequency	Percent	Valid percent	Cumulative percent
A	258	24.6	26.3	26.3
B (correct response)	558	53.3	56.8	83.1
C	113	10.8	11.5	94.6
D	53	5.1	5.4	100.0
9.00(missing cases)	65	6.2	missing	
Total	1047	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

Table 5.11: Year 7 Frequency table for Item MUA11

Value	Frequency	Percent	Valid percent	Cumulative percent
A	102	10.2	10.7	10.7
B (correct response)	775	77.7	81.5	92.2
C	53	5.3	5.6	97.8
D	21	2.1	2.2	100.0
9.00(missing cases)	14	3.9	missing	
Total	998	100.0	100.0	

Table 5.12: Year 10 Frequency table for Item MUA11

Value	Frequency	Percent	Valid percent	Cumulative percent
A	10	2.8	2.9	2.9
B (correct response)	321	88.4	92.8	95.7
C	12	3.3	3.5	99.1
D	3	.8	.9	100.0
9.00(missing cases)	17	4.7	missing	
Total	363	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

The Process task items were the same for Years 3, 7 and 10 and the marking key was common across the three Year groups. The first item on the marking key, coded MuP01 and MuP18 (See Appendix xi Music Process Marking Key), related to the Student Outcome strand "Using Skills Techniques, technologies and processes" which is one of the two strands of *Expressing*. This item related wholly to the performance of the group. It was the marker's on-balance judgement of the group's overall performance and their ability to communicate their composition/message in the language of music. The first category of "Beginning to develop" was achieved by 59.8 per cent of Year 3 students, 40 per cent of Year 7 students and 8.2 per cent of Year 10 students. It was obvious to the marker that students who achieved this category were making an attempt to reflect the stimulus and convey their

message through soundscape. To achieve the category “Beginning to develop,” the piece had to be more than just a series of isolated, non-musical sound effects, although it did not display qualities of expression or mood and the use of musical elements was limited. The highest category of “Highly developed” was not scored at Year 3 and was scored by only 1.0 per cent of Year 7 students compared with 17.3 per cent of Year 10 students. A “highly developed” performance displayed evidence of sensory experiences to effectively communicate a message to reflect the stimulus through the use of a variety of musical elements such as melody, rhythm, dynamics, tempo and texture. From these results, a clear pattern of development of the students’ overall performance can be seen (see Tables 5.13, 5.14 and 5.15 hereunder).

Table 5.13: Year 3 Frequency table for Item MUP01

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	216	20.6	21.3	21.3
1.00	608	58.1	59.8	81.1
2.00	171	16.3	16.8	97.9
3.00	21	2.0	2.1	100.0
9.00(missing cases)	31	3.0	missing	
Total	1047	100.0	100.0	

Table 5.14: Year 7 Frequency table for Item MUP18

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	35	3.5	3.7	3.7
1.00	378	37.9	40.0	43.7
2.00	428	42.9	45.3	89.0
3.00	95	9.5	10.1	99.0
4.00	9	.9	1.0	100.0
9.00(missing cases)	53	5.3	missing	
Total	998	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

Table 5.15: Year 10 Frequency table for Item MUP18

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	4	1.1	1.2	1.2
1.00	28	7.7	8.2	9.4
2.00	105	28.9	30.8	40.2
3.00	50	39.9	42.5	82.7
4.00	14	16.3	17.3	100.0
9.00(missing cases)	22	6.1	missing	
Total	363	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

The second item on the Process marking key, coded MuP02 and MuP19 (See Appendix xi Music Process Marking Key) related to the Student Outcome *Expressing* strand of "Creating, Exploring and Developing Ideas" and focused on the group's planning of the performance. Students were required to work with their groups to explore ideas from the stimulus provided and to capture the mood of the stimulus in their planning of a performance. Planning was to include interpretation of the stimulus, through the selection of instruments to be used, and through the use of elements such as pitch, dynamics, rhythm and phrasing. The planning should also have included organisational procedures within the group; that is, who would play which instruments, where they would be placed and who, if anyone, would lead the group. Planning should also have included a score of some description. Scores could have been shown in either unconventional or conventional musical notation. Category 1, "Beginning to develop" was achieved by groups who showed some attempt to relate performance to the stimulus. For instance, they may have made lists of the sounds and the sound sources or instruments they intended to use, without actually describing the sounds and linking them to the instrument sounds. They may have assigned tasks or

roles to individuals within the group but with no indication of leadership or placement of players. There was no attempt to produce a score of any kind. Category 1 was achieved by 65.6 per cent of Year 3 students, 33.3 per cent of Year 7 students and 10.5 per cent of Year 10 students. Students who achieved categories 2 and 3 moved slightly further along the continuum, attempting to produce a score and making stronger links between the stimulus and sound sources. For instance, they may have said they would use “triangle tinkles to make rain” or “clarinet for a rusty windmill squeaking.” They also attempted to show more organisation of cues to start and finish performances. Category 4, “Well developed,” was achieved by groups who made strong links between sound sources or instruments and descriptions of sounds. These groups also produced a score in conventional form that, although not strong in the use of musical terminology, provided clear structure of composition that correlated with the final performance, and that could have been played or followed by others. This category was not achieved at Year 3 and was achieved by only 1.7 per cent of Year 7 students, compared with 19.6 per cent of Year 10 students. An additional category 5, “Highly developed,” for groups which achieved all of the requirements for category 4 as well as a well-developed musical score using correct, sophisticated musical terminology and a variety of musical elements such as melody, rhythm, dynamics or tempo, was scored by 4.4 per cent of Year 10 students. This category was not scored at all at Years 3 or 7 (see Tables 5.16, 5.17 and 5.18 hereunder).

Table 5.16: Year 3 Frequency table for Item MUP02

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	213	20.3	21.0	21.0
1.00	626	59.8	61.9	82.9
2.00	156	14.9	15.4	98.3
3.00	17	1.6	1.7	100.0
9.00(missing cases)	35	3.3	missing	
Total	1047	100.0	100.0	

Table 5.17: Year 7 Frequency table for Item MUP19

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	29	2.9	3.1	3.1
1.00	313	31.4	33.3	36.3
2.00	305	30.6	32.4	68.8
3.00	278	27.9	29.5	98.3
4.00	16	1.6	1.7	100.0
9.00(missing cases)	57	5.7	missing	
Total	998	100.0	100.0	

Table 5.18: Year 10 Frequency table for Item MUP19

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	22	6.1	6.4	6.4
1.00	36	9.9	10.5	17.0
2.00	33	9.1	9.6	26.6
3.00	169	46.6	49.4	76.0
4.00	67	18.5	19.6	95.6
5.00	15	4.1	4.4	100.0
9.00(missing cases)	21	5.8	missing	
Total	363	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

Process item C1, code MuP11, MuP14 and MuP 16 (See Appendix xi Music Process Marking Key) related to the Student Outcome strand “Responding, reflecting and evaluating” which is one of the strands of *Appreciating*. This item was the individual student’s reflection of the group’s performance and the justification of the importance of selected musical elements used in their composition. The majority of Year 3 students (68.4 per cent) answered in the first category, “Beginning to develop,” compared with 40.5 per cent at Year 7 and 9.0 per cent at Year 10. Students at the

higher Year levels answered less in the first category and more in the second, third and fourth categories. Students who answered in the first category provided almost tautological or irrelevant justification of the use of elements that did not demonstrate understanding. For instance, “We did good melody.” Those who achieved Category 2 displayed limited understanding of the relevant musical element but did not link it to the composition. For instance “I liked our melody because it had a good tune.” Those who achieved the third category, “Well developed,” provided justification that indicated understanding of the relevant musical element and which was linked to the composition. For example, “We used high notes in the melody to reflect the sunshine and low notes to show the storm.” Only 1.5 per cent of Year 3 students scored in the third category, “Well developed,” compared with 10.4 per cent at Year 7 and 49.4 per cent at Year 10. No Year 3 students achieved the fourth category, “Highly developed.” To achieve this category, students, as well as demonstrating understanding and linking to the composition, had to discuss elements in the context of the whole work. For example, “We used a minor key for our melody with a slow tempo to create a mood of eeriness for the deserted farm.” Only 0.1 per cent of students at Year 7 and 9.3 per cent of students at Year 10 achieved this category (see tables 5.18, 5.19 and 5.20 hereunder).

Table 5.19: Year 3 Frequency table for Item MUP11

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	77	7.4	8.1	8.1
1.00	654	62.5	68.4	76.5
2.00	211	20.2	22.1	98.5
3.00	14	1.3	1.5	100.0
9.00(missing cases)	91	8.7	missing	
Total	1047	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
Valid percent: the percentage of students achieving that score
Frequency: the number of students achieving that score

Table 5.20: Year 7 Frequency table for Item MUP14

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	20	2.0	2.2	2.2
1.00	375	37.6	40.5	42.7
2.00	434	43.5	46.9	89.5
3.00	96	9.6	10.4	99.9
4.00	1	.1	.1	100.0
9.00(missing cases)	72	7.2	missing	
Total	998	100.0	100.0	

Table 5.21: Year 10 Frequency table for Item MUP16

Value	Frequency	Percent	Valid percent	Cumulative percent
.00	12	3.3	3.8	3.8
1.00	28	7.7	9.0	12.8
2.00	89	24.5	28.5	41.3
3.00	154	42.4	49.4	90.7
4.00	29	8.0	9.3	100.0
9.00(missing cases)	51	14.0	missing	
Total	363	100.0	100.0	

NOTE: Value: the partial credit raw scores allocated to the item
 Valid percent: the percentage of students achieving that score
 Frequency: the number of students achieving that score

Conclusions

This study shows that, despite the beliefs of some arts educators in the past, it is possible to measure a non-quantitative subject such as *The Arts*, using quantitative measurement techniques such as a Rasch model. It is, however, necessary to have a developmental framework of achievement such as *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) upon which to measure levels of achievement. If all facets of *The Arts* in the classroom, that is, all strands of the outcome statements are to be covered, it is necessary to test students' abilities in the exploration and creation of the art form, as well as their appreciation and analysis of it. The

data collected from both tests needs to be combined in order to assess the overall ability of the student in the discipline.

In accordance with the original aims of the study, the music assessment instrument and marking keys have been successfully developed to reliably measure music outcomes for Year 3 (8 year olds), Year 7 (12 year olds) and Year 10 (15 year olds), using The Arts Student Outcome Statements (Education Department of Western Australia, 1996) as a framework. The analysis, using the Extended Logistic Model of Rasch (1988) has created interval level measurements for the instrument, as well as benchmarks or standards of performance that can be used by teachers to make comparisons between their own students' performances and state means. Patterns of performance from Year 3, through Year 7 to Year 10 have also been achieved through the use of common or 'link' items in tests across the three Year groups.

The use of open-ended questions that allow for partial credit has highlighted the fact that, in most cases, student responses are not 'right' or 'wrong'. Students can often display some feature of the trait being tested, even if they cannot answer at a highly technical level. The use of open-ended questions also provided the 'link' items that showed patterns of development across Year groups, as well as allowing students to achieve to the maximum of their abilities. Caution needs to be exercised, however, when using open-ended questions, to ensure that students' levels of literacy skills do not affect results. Instructions to test administrators need to clearly state that questions

may be read aloud and comprehension questions answered to ensure that all students understand what is required. Spelling, grammar and sentence construction must not be a factor in the marking of tests.

The music achievement scale developed for this study has provided the opportunity to develop a useful set of music assessment instruments and marking keys appropriate for use at Years 3, 7 and 10 by teachers in Western Australian schools, as well as benchmarks and profiles for comparison with state means. The instruments will provide a useful model for teachers in developing tasks to monitor students' progress using The Arts Student Outcome Statements (Education Department of Western Australia, 1996).

Some significant issues that have arisen from the sampling method used in this study are the limited access to *The Arts* for many Year 10 secondary students in Western Australia, as well as the gender imbalance apparent in all arts disciplines. These issues are of major concern to the future of *The Arts* in Western Australian schools and should be the subject of further investigation.

CHAPTER SIX
DATA ANALYSIS (PART A): PSYCHOMETRIC
CHARACTERISTICS AND MEANING OF THE SCALE

Introduction to the Scale of Music Achievement

A Rasch analysis was undertaken to transform students' raw scores onto a scale in which the unit of measure is constant along the scale, and to produce estimates of the difficulty of items on the same scale as the measurement of music achievement for students. The model assumes that a score of zero does not imply a complete absence of music achievement and a perfect test score does not imply a presence of knowledge of all musical concepts. For example, a student who scored zero on the test may know something about music that was not asked in this particular test. Conversely, there may be many things about music that a student who gained a perfect score on this particular test does not know. For this reason, the items are located centrally around zero on the scale with no limit in score to indicate absence of ability or perfect ability. Technically, the mean difficulty of the items is calibrated as zero on the scale. The unit of measure, which is used for both item difficulty and student music achievement is called the logit – the log odds of answering the item correctly.

The parameter estimate for the items, that is, the item difficulties, range from -3.771 logits, to highest difficulty of 1.373 logits (see Figure 6.1). This means that the difficulty level of the items is spread over a range exceeding

five logits, which is a wide range of difficulty levels for a single testing situation where both time and resources are constrained. It means that there was opportunity for all students, from those with very limited ability levels to those with high ability levels to display skills and abilities across at least six achievement levels of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). The allocation of levels will be explained later in this chapter (page 184).

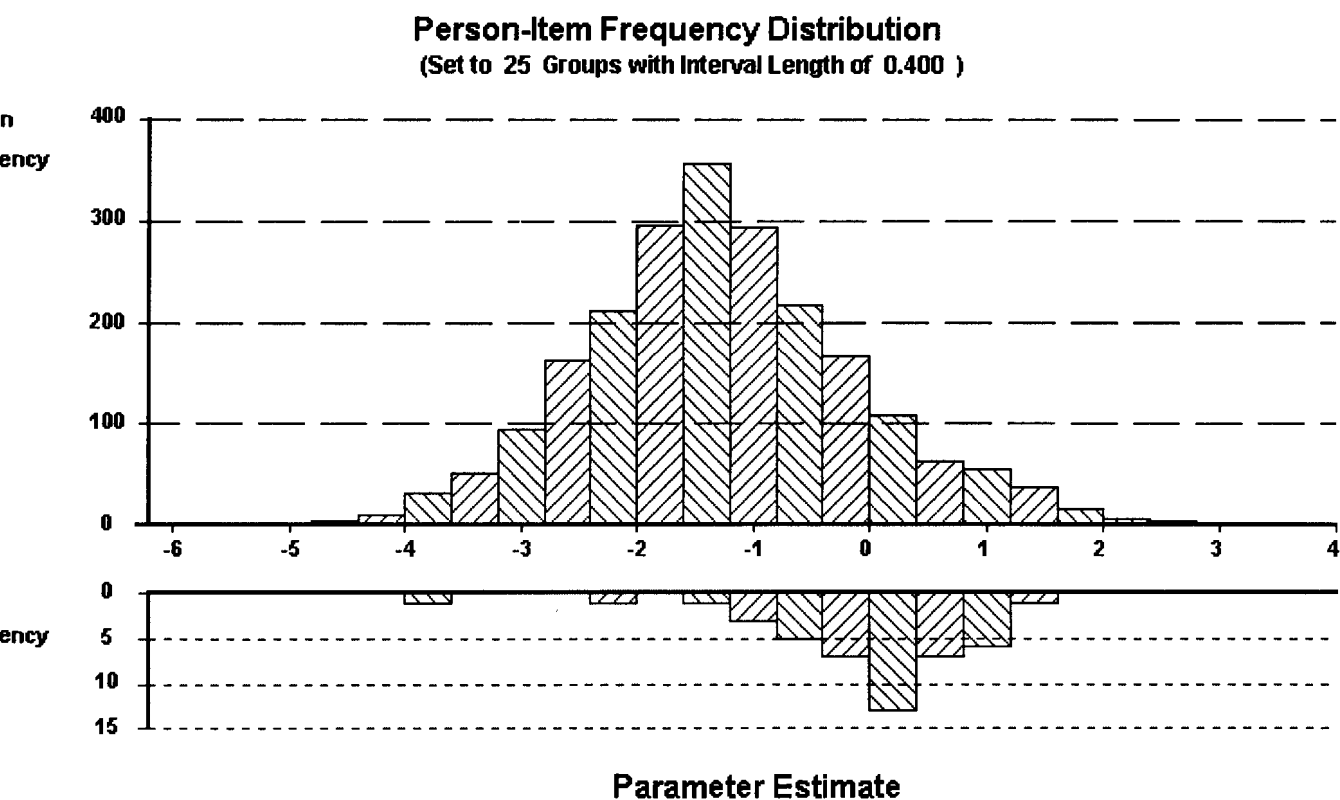


Figure 6.1: Distribution of Student Music Achievement and Item Location Estimates

NOTE: the top line represents the total number of Year 3, 7 and 10 students divided into 25 ability groups and distributed along the logit scale from lowest ability, -5.016 logits, to highest ability, 3.034 logits. Each bar on the graph represents the number of persons achieving that level along the continuum (Person frequency). Note that the five ability groups at the ends of the scale are not visible because of the small number of students in these groups. The bottom line represents the range of test items, distributed from easiest to most difficult, along the same scale. Each bar represents the number of items located at that level of difficulty along the scale from the easiest item, located at -3.771 logits, to the most difficult item, located at 1.373 logits (Item frequency).

The parameter estimate for the students, that is the student measures of music achievement, range from -5.016 logits to 3.034 logits. The student sample is biased towards the negative end of the scale in comparison to the item difficulties. This means that the students found the test to be difficult, or at least, music achievement was graded at a high standard. This is to be expected in view of the fact that the sample of students represents the three groups; Year 3, Year 7 and Year 10, and the number of students in the Year 10 sample was less than those for Years 3 and 7. The sample at Year 10 consisted of 324 students, compared with 921 at Year 7 and 946 at Year 3. As was expected, Year 3 and Year 7 students produced lower level responses than those of Year 10 and, subsequently, the person-frequency distribution curve is biased towards the negative end of the scale. Figure 6.1 indicates a spread of ability levels from -5.016 logits to 3.034 logits. This means that there was a broad spread of ability levels among students, allowing a sufficient range to provide the level of variance required for mapping student knowledge and abilities using a student outcome framework. The tail that can be observed at each end of the continuum indicates that there were a small number of students at each of the extremes of ability levels; that is, some with very little knowledge and ability in music and some with a high level of knowledge and ability.

The mean score for music achievement for Year 3 is -1.55 logits, the mean for Year 7 is -0.76 logits and the mean for Year 10 is 0.85 logits. This indicates a steady progress of music achievement from Year 3, through Year 7, to Year 10. The increase in music achievement from Year 7 to Year 10 is

greater than that from Year 3 to Year 7. There are three reasons for this. The first is that, in primary school, all students, regardless of their ability, participate in music programs whereas, in Year 10, only those students who have selected, and qualified for, a music option are participating. The second is that, in primary school, music is frequently taught by generalist teachers whereas, in Year 10 music option classes, music is always taught by music specialists. The third is that, in primary school, music classes are sometimes ad-hoc and lacking in continuity, whereas Year 10 music option classes are time-tabled on a regular basis.

General psychometric characteristics of the scale

Fit of students to the model

A test-of-fit was conducted to examine the degree to which students responded to items of differing levels of difficulty in a logical and consistent manner. The student test-of-fit relates to two aspects. The first involves the response pattern of individual students across all items and the second involves the pattern for each item across all students. The residual between the expected estimate and the actual data values for each student-item combination is examined, both across all items for an individual student, and across all students for an individual item.

These fit statistics approximate a t-distribution when the data fit the model. This means that, ideally, the overall distribution for both the item and student statistics should have a mean of zero and a standard deviation of one

(Andrich & Sheridan, 1980). Any individual item or student statistic which is greater than plus or minus four is considered to be digressing from the model in a way that is not by chance alone. Negative, or decreasing values indicate a student or item pattern response that is over-discriminating. In other words, there are dependencies present in the item that may benefit certain student groups. Positive, or increasing values indicate under-discrimination or poor fit to the model. This means that the item is contributing little towards the measurement of knowledge or understanding of music.

An example of an item that is over-discriminating is item MP18, which had a chi square of 124.529 and a fit statistic of -12.028 . This was a Process item for Years 7 and 10 and relates to the marker's overall on-balance judgement of each group's ability to communicate through music. There were larger than expected differences between lower ability groups and higher ability groups in the quality of their performances. A reason for this could be that, in Year 10, students focus much of their time on performance, whereas in Year 7, performance does not play a significant part in classroom music. This means the dependency, that is performance, favoured the Year 10, or higher ability group.

An item that is under-discriminating is item MA10, which had a chi square of 323.565 and a fit statistic of 10.546. This was a multiple choice Analysis item for Years 3, 7 and 10 in which students were asked to identify the instrument playing a short excerpt of music from *Ballet for Children* (Bliss, 1995). There was little difference between the performance of lower ability

groups and higher ability groups on this item. In other words, the Year 3 students were able to select “trumpet” almost as frequently as the Year 10 students. This item contributes little to the integrity of the test as it does not display sufficient differences in student knowledge and abilities. The table of Individual Item-fit statistics is displayed as Appendix xv.

Fit of items to the model

The RUMM (1996) program provides two statistics for the estimation of item fit to the model. The first of these is an item-student interaction statistic. The degree to which students respond to items of different difficulty value in a consistent manner is examined. The fit statistic distribution approximates a t distribution when the data fit the model and have a mean of zero and a standard deviation of one (Andrich & Sheridan, 1980). A negative value indicates a student or item pattern response fits the model closely, while a positive value indicates poor fit to the model (see Appendix xv) In the case of negative values, there is usually one or more dependencies present in the data. Where values are positive, it usually means that the item does not contribute in a meaningful way to measurement of the trait being tested and that it does not ‘work with’ the other items on the test.

The second statistic used for the assessment of item fit of data to the model is the item-trait interaction test-of-fit, which examines the consistency of the item parameters over the range of student estimates. The item-trait interaction test-of-fit indicates the degree of consensus of the items across students located at different ability levels. The initial RUMM (1996) analysis of

the combined results of the music Analysis and Process tests produced a mean of 0.000 which indicates excellent overall fit to the model, and a standard deviation of 0.898. The total number of students included in the analysis was 2192. The Person Separation Index, which is the proportion of observed variance considered true, for the final analysis was 0.896, which means that the proportion of observed variance considered true was 90%. The RUMM (1996) program measures the power of the tests of fit on a five point scale, that is: *Excellent, Good, Reasonable, Low and Too low*. The Power of the tests of fit for this analysis was rated as *Excellent*, and the total item chi-squared value for the test was 3577.143 ($p = 0.000$), indicating a strong fit to the model.

Individual items were examined for 'fit' to the model. When data are ideally ordered, students with high ability are expected to attain higher total scores and be more likely to correctly answer an item than students of lower ability. Where this does not happen and there is unexplained inconsistency of performance, the item does not fit the model. An example of this is when the number of low ability students correctly answering an item exceeds the number of high ability students correctly answering the same item. Where this occurs, a search is made for the source of the anomaly, which may be the wording of the item, an incorrect or unclear marking key, or a source of bias that advantages less able students. A poorly fitting item may be deleted from the final analysis.

The student sample was divided into ten ability groups and the data were examined for identification of the least fitting items. The chi-square statistic was used as a general guide but, in view of the large sample size, this statistic was treated with caution and the residual fit statistic was used as an indicator of item fit. For items where this statistic was extreme, the Item Characteristic Curve from the RUMM (1996) program was used to examine the observed and expected values for each group for that item. Although there were a few items that showed relatively low discrimination, they were all positive and contributed to the integrity of the test (See Appendix vx for table of Individual Item-fit). Therefore, a decision was made not to delete any items from the final analysis.

An example of a RUMM (1996) Item Characteristic Curve for Item MA21 is shown hereunder (Figure 6.2). This was a link item between Year 7 and Year 10. Students were asked to compare and contrast the two pieces, *Ballet for Children* (Bliss, 1995) and *Tribal Voice* (Kellaway & Yunupingu, 1992) in terms of instrumentation. The line indicates the expected score of ability groups, ranging from the lowest to highest achieving groups. Each black dot represents the observed score of a student ability group. When the observed scores closely follow the curve of expected values, the group is performing as expected on the item. Item MA21 shows good fit to the model with all groups achieving close to the expected scores, except that the highest achieving group did not perform quite as well as expected. This is indicated by the location, below the line, of the black spot representing the highest ability group.

Ex013 MA21: Locn = 1.060 Resid = 2.286 ChiSqProb = 0.042

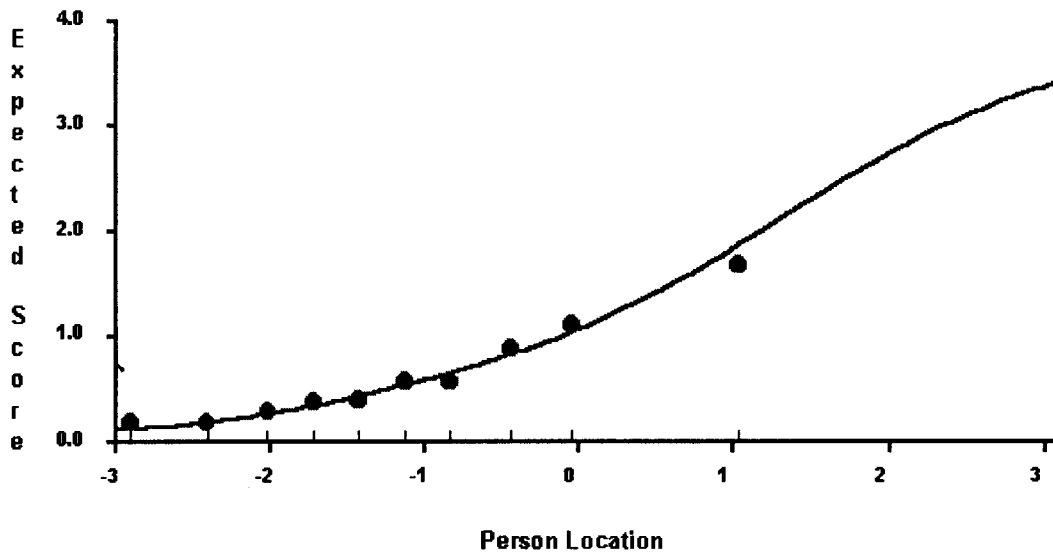


Figure 6.2 Item Characteristic Curve for Item MA21

One of the least fitting items is Item MA18. This was an Analysis link item between Year 7 and Year 10. Students were asked to identify the different styles of music combined by the composer in the piece *Tribal Voice* (Kellaway & Yunupingu, 1992). The five lower ability groups have performed higher than expected on this item, whereas the five higher ability groups have performed lower than expected. This is demonstrated by the Item Characteristic curve for Item MA18 where the black spots representing the five lower ability groups appear above the black line (which represents the expected score) and the black spots representing the five higher ability groups appear below the black line. (see Figure 6.3). A possible explanation for the poor fit of this item is the marking key. This was an Analysis link item from Year 7 to Year 10. Students received 0 marks for an incorrect response and one mark each for “rock music” and “Aboriginal music”. It appears that, even

those students who have a limited knowledge of music, knew that rock and traditional aboriginal music had been combined in this piece. There was no scope for the more able students to describe the music in more technical terms. This resulted in minimal discrimination between lower and higher groups on this item.

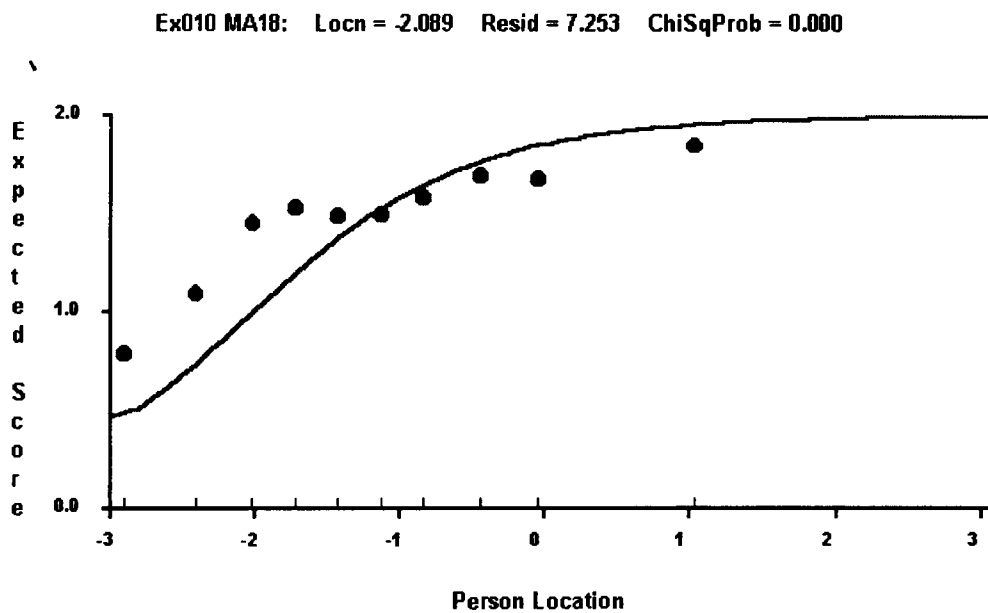


Figure 6.3: Item Characteristic Curve for Item MA18

Another of the least fitting items is Item MA08. This was an Analysis link item between Years 3, 7 and 10. Students listened to an excerpt from the piece *Ballet for Children* (Bliss, 1995) and were asked to explain what they heard in the music to justify their interpretation of a ‘change’ in the beat. The six lower ability groups all performed above expectations for this item,

whereas the three higher ability groups performed below the expected level. This is displayed in the Item Characteristic Curve for Item MA08 where the spots representing the lower ability groups are above the expected score line and the spots representing the higher ability groups are below the line (see Figure 6.4).

One cause for the poor fit of this item may have been the inadequacy of the marking key. This was a four-category response item with categories earning 0, 1, 2 and 3 marks. To gain the maximum result of three marks, students had to “make reference to rhythm change, eg beat changes from march to waltz time or from 4/4 to 3/4.” Compared to other item categories that attracted a three-mark credit, this is a fairly simple response. Many students with minimal knowledge and ability in music are able to discriminate between a march and a waltz rhythm. In fact, the thresholds between categories one and two in this item were disordered and so these two categories were collapsed into one, thus making it a three category item. (For further explanation of the disordering of thresholds for this item see page 146).

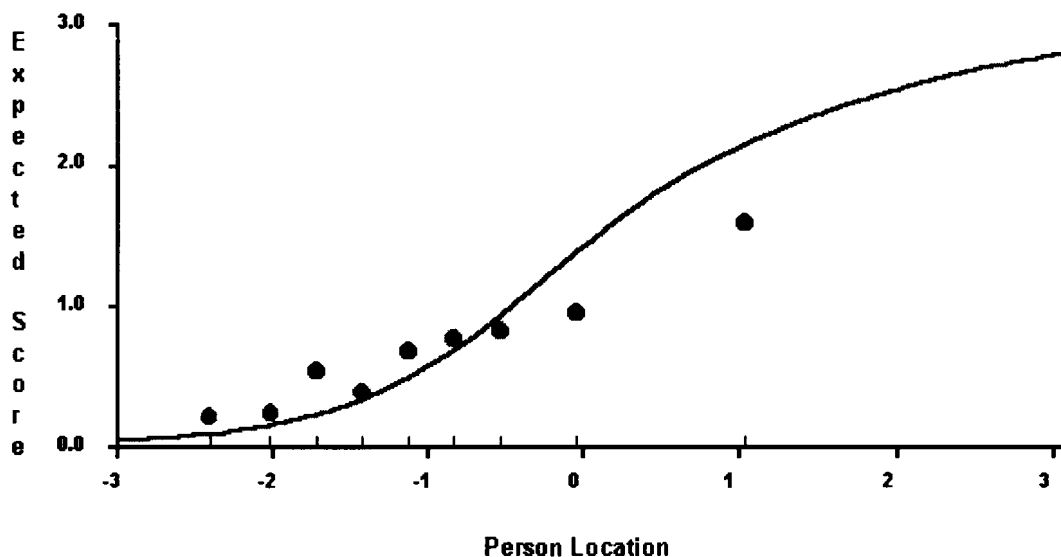


Figure 6.4: Item Characteristic Curve for Item MA08

An item in which the performance of ability groups was erratic is Analysis Item MA25 which was completed by only the Year 10 students. They listened to an excerpt from *Earth, cry kakadu, mangrove* (Sculthorpe, 1989) and were asked to describe the mood of the landscape created by the composer in the music. The lowest ability group performed at a slightly higher level than expected, while the second lowest ability group performed at a much higher level than expected; their performance almost matched that of the highest ability group. The third lowest ability group performed at the expected level, while the next five ability groups performed better than expected. The highest ability group performed below expectations. A possible explanation for the erratic performance of groups on this item is that students were being asked to put their own interpretation on the mood created by the composer, and to describe it in words. It is possible that this item related more to the literacy skills of students and their ability to write an appropriate description of the landscape, rather than their musical skills and knowledge. The Item

Characteristic Curve for Item MA25 demonstrates the erratic performance levels of ability groups (see Figure 6.5).

Ex017 MA25: Locn = 0.435 Resid = 2.517 ChiSqProb = 0.000

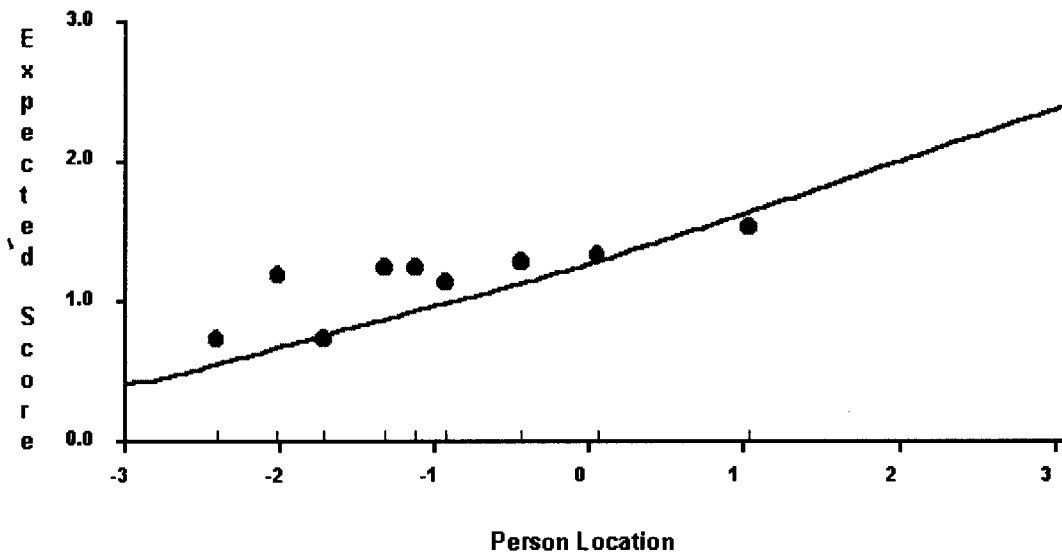


Figure 6.5: Item Characteristic Curve for Item MA25

In many instances where 'misfit' occurred in the Process items, it was because the highest ability groups performed better than expected. An explanation for this is that the Year 10 music specialist programs usually focus most of their time and attention on attaining high level skills in performance. An example of this feature is Item MP08 which related to the effective use of harmony during the groups' musical performances. While the lower ability groups performed within expectations, the top two groups performed well above the expected levels. This is demonstrated by the Item Characteristic Curve for Item MP08 (see figure 6.6).

Ex027 MP08: Locn = 0.270 Resid = -3.076 ChiSqProb = 0.000

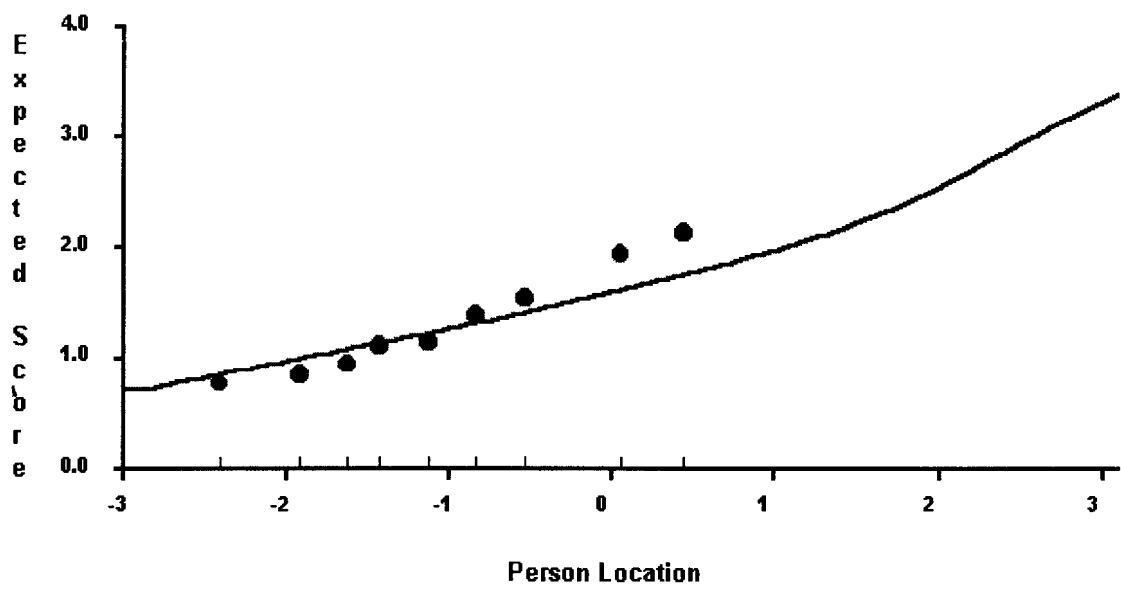


Figure 6.6: Item Characteristic Curve for Item MP08

A Process item in which ability groups performed to expectations is Item MP19. This is a Year 7 and Year 10 item and relates to the plan and musical score developed by class groups in preparation for their performances. The Item Characteristic Curve for Item MP19 demonstrates that all ability groups performed to expectations (see figure 6.7).

Ex038 MP19: Locn = 0.022 Resid = 0.382 ChiSqProb = 0.578

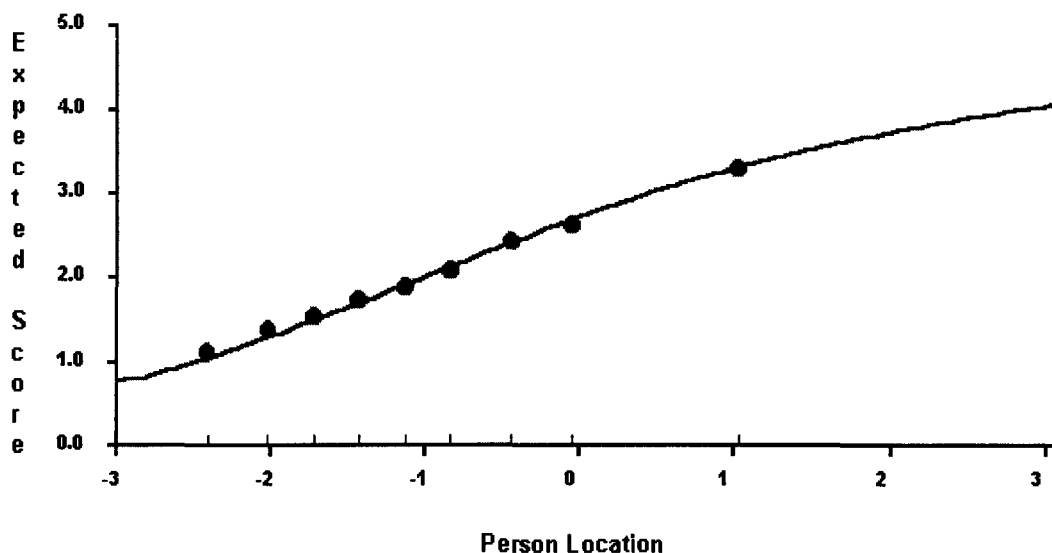


Figure 6.7: Item Characteristic Curve for Item MP19

Validity of the measure of music achievement

The measure of music achievements for this study consisted of 27 Analysis tasks, and 26 Process tasks. The partial credit method of marking the items provided a total of 165 item categories. These categories were all measured using *the Arts Student Outcome Statements* (Education Department of Western Australia, 1996) as a framework.

The two main criteria for measuring music achievement, that is, to 'do it' and to 'understand and appreciate it,' have been measured using both the Creating and Appreciating strands of *the Arts Student Outcome Statements* (Education Department of Western Australia, 1996) descriptive profile and reporting framework. The range of items has provided the opportunity for the

demonstration of skills and abilities along a continuum which maps student achievement levels from simple experiences to complex tasks. All items were created from the same conceptual framework. A Rasch model of analysis has converted students' raw scores to item logits and located student achievement measures on the same continuum as item difficulty. The music achievement scores range from -4.316 logits, to 3.057 logits and the item difficulties range from -4.316 logits to 6.102 logits with ordered thresholds. Although Figure 6.1 (p.151) indicates that many items were difficult for students at Year 3, this was intended so that the tests would display progression of student achievement from Year 3, through Year 7, to Year 10. The range and difficulties of items were appropriately targeted in each of the test instruments to provide Year 3, Year 7 and Year 10 students with the opportunity to perform to the maximum of their ability levels.

The separability (reliability) is high (see page 156), the thresholds are now ordered (see pages 166-179), the item-person interaction is good - students responding to items of different difficulties logically and consistently (see page 155), the consistency of item parameters is good (see page 156), item difficulty and person measures are calibrated on the same scale, and targeting of items to student abilities is satisfactory.

Threshold values

The model of analysis used in this study is an Extended Logistic Model of Rasch where partial-credit items were developed to allow for three or more

ordered categories of response. The model checks that the category responses are ordered by calculating threshold values between categories. These thresholds are conceptualised as a set of boundaries between the response categories, and indicate the change in probability of a response occurring in two adjacent ordered categories. When the item fits the model, the thresholds are ordered in correspondence with the ordered response categories. The threshold parameters are labelled, in hierarchical order and the first of these is location (δ), which is the position of an item on the music achievement scale. The second parameter is scale (θ), which is equal to the average half-distance between thresholds for an item and, ideally, is the same for each item, within the measurement error; the third is skewness (η) which represents the asymmetric deviation from equidistant thresholds; and the fourth is kurtosis (ψ) which represents the symmetric deviation from equidistant thresholds.

A threshold structure associated with the categories was parameterised to obtain an indication of the item structure which, if the data had fit the model, should have demonstrated a hierarchical progression from the easiest to the most difficult to achieve. Ideally, for a student possessing a low level of knowledge and ability in music, all response categories should be answered consistently, for all items, in order from easiest to most difficult. For example, the most probable response to a difficult item would be in the first category. A student possessing a high level of knowledge and ability in music would probably respond in a high order category for the same item. This means that students with a high level of ability in music would achieve successively

higher order categories than would be the case for a student of lesser music ability. If this pattern is observed for all items then the item thresholds will be in an ordered sequence. If thresholds are not ordered, the categories are not behaving as expected and this needs to be investigated. An examination of the Category Probability Curves provided by the RUMM (1996) program is a useful source for facilitating this investigation. An example of a Category Probability Curve showing the ideal ordering of sequence attained is Item MA19. This was an analysis link item for Years 7 and 10 in which students were asked to discuss elements of the music that helped them to decide what different styles of music were combined by the composer in the piece *Tribal Voice* (Kellaway & Yunupingu, 1992). The item was developed to contain four response categories; that is, 0 marks, 1 mark, 2 marks and 3 marks. The thresholds for the respective categories were ordered -3.398 , 0.051 and 3.347 . Figure 6.8 displays the curves that represents each item category and its' allocated mark. The even distribution of the curves indicates an evenly distributed order of thresholds.

Ex011 MA19: Locn = 0.881 Resid = 3.224 ChiSqProb = 0.000

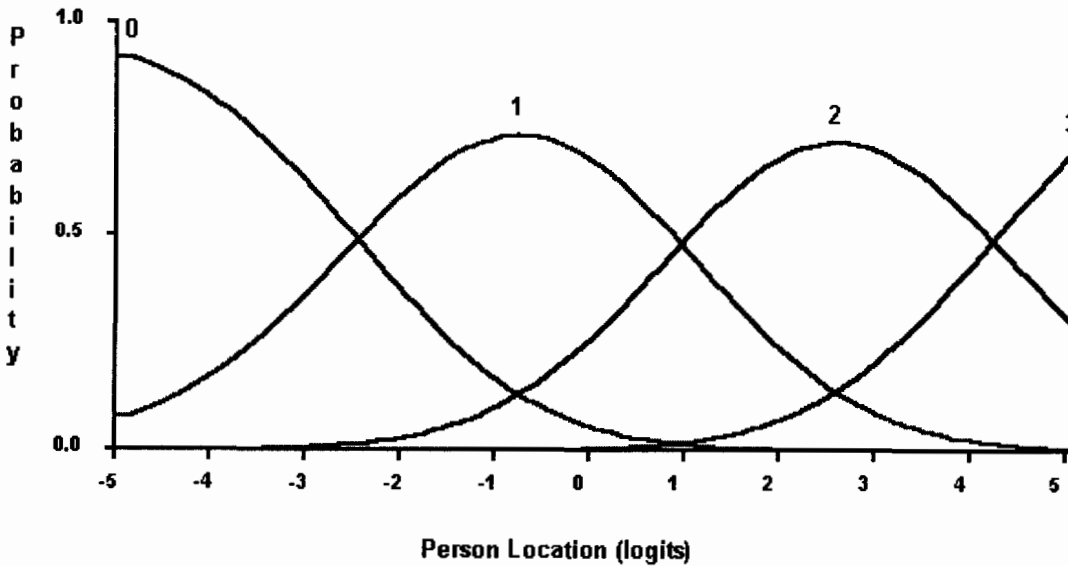


Figure 6.8: Category Probability Curve for Item MA19

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 is around 0.9 and if their music ability is located at 1.0 logit, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that, if a student's music ability is located at -5.0 , the probability of getting 1 is around 0.1 and if their music ability is located at -1.0 logits, the probability of getting 1 mark is around 0.8.
3. The Category 2 curve indicates that, if a student's music ability is located at -3.0 logits, the probability of getting 2 marks is 0 and if their music ability is located at 2.5 logits, the probability of getting 2 marks is around 0.8.
4. The Category 3 curve indicates that, if a student's music ability is around 1.0 logits, the probability of getting 3 marks is 0 and if their music ability is located at 5.0 logits, the probability of getting 3 marks is around 0.8.

Reversed thresholds

There were four Analysis items and four Process items in which the thresholds were either not strictly ordered or had minimal discrimination. In these cases, an examination of the category probability curve was undertaken to facilitate decisions regarding the rescoring, or collapsing of categories for these items. Each of these items is discussed individually.

The first Analysis item in which thresholds were disordered was item MA06. Students were asked to justify their interpretation of the mood of the

piece *Ballet for Children* (Bliss, 1995) by describing what they had heard in the music to suggest the mood. The item was developed to contain four response categories; that is, 0, 1, 2 and 3 marks. A response in the second category (1 mark) related to the sound of the music, as for instance, “the sound was fast and joyful.” The threshold location for this category was – 1.459 and was ordered. A response in the third category (2 marks) was intended to display a reference to musical elements or musical instruments and a response in the fourth category (3 marks) was intended to display a discussion of rhythm, instrumentation, texture and harmony. The threshold location for category three was 0.907 and for category four was 0.552, indicating that it was more likely for a student to achieve the fourth category than the third. In other words, rather than just making a passing reference to elements, students were more likely to discuss them in some way. Unlike Figure 6.8, the curves representing Item MA06 are not evenly distributed (Figure 6.9). Categories three and four were collapsed into one category for this item, thus making it a three category item. This means that students were allocated two marks for either reference to, or discussion of musical elements and instrumentation in their responses. A re-analysis found the thresholds to be ordered properly with three categories.

The second Analysis item in which disordering of thresholds was evident was Item MA08. Students listened to a short excerpt from the piece, *Ballet for Children* (Bliss, 1995), and were asked to explain what they had heard in the piece to justify their explanation of a change in the music. The item contained

four response categories earning 0, 1, 2 or 3 marks. Students received 1 mark (category two) for a reference to movement, or for reference to fast and

Ex003 MA06: Locn = 0.553 Resid = 2.449 ChiSqProb = 0.000

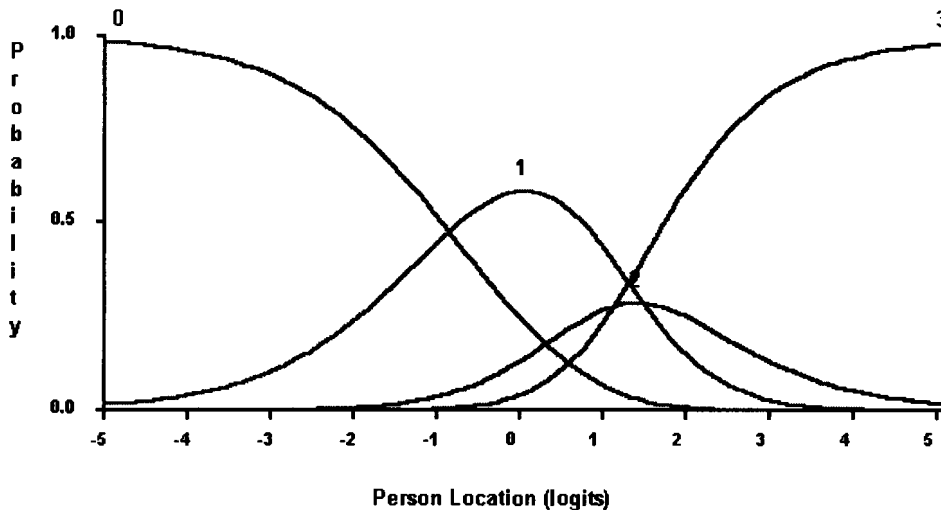


Figure 6.9: Category Probability Curve for Item MA06

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 is 1.0 and if their music ability is located at 2.0 logit, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that, if a student's music ability is located at -5.0 , the probability of getting 1 is around 0 and if their music ability is located at 0 logits, the probability of getting 1 mark is around 0.6.
3. The Category 2 curve indicates that, for no student group is there a probability of getting 2 marks.
4. The Category 3 curve indicates that, if a student's music ability is around -1.0 logits, the probability of getting 3 marks is 0 and if their music ability is located at 5.0 logits, the probability of getting 3 marks is 1.0 .

slow without actually using the term 'beat'. They received 2 marks (category three) for a reference to fast and slow, using the terminology, 'beat.' The threshold location for category two (1 mark) was -0.185 , and for category three (2 marks) the location was -1.204 . This indicates there was more likelihood of students achieving what was intended to be the more difficult category. Further examination revealed that, if students knew the difference between fast and slow music, they called it 'beat' and so very few students

scored in the category two range. This is demonstrated by the category probability curve for Item MA08 (Figure 6.10). Categories two and three were collapsed and 1 mark allocated to each. In other words, if a student referred to slow or fast music or slow or fast 'beat' of the music, they were allocated 1 mark. A re-analysis showed that the thresholds were then found to be properly ordered with three categories.

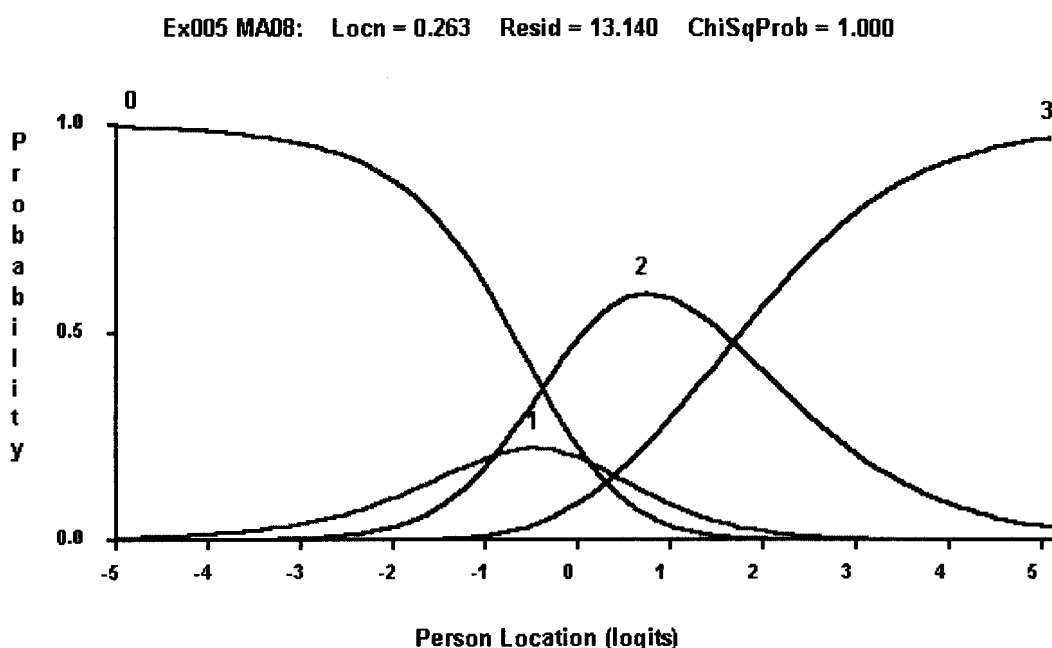


Figure 6.10: Category Probability Curve for Item MA08

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 is 1.0 and if their music ability is located at 2.0 logits, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that for no student group is there a probability of getting 1 mark.
3. The Category 2 curve indicates that, if a student's music ability is located at -3.0 logits, the probability of getting 2 marks is 0 and if their music ability is located at 0.6 logits, the probability of getting 2 marks is around 0.6.
4. The Category 3 curve indicates that, if a student's music ability is located around -3.0 logits, the probability of getting 3 marks is 0 and if their music ability is located at 5.0 logits, the probability of getting 3 marks is 1.0.

Analysis item MA18 was the third Analysis item with reversed thresholds. After listening to the piece *Tribal Voice* (Kellaway & Yunupingu,

1992), students were asked to name the different styles of music the composers had combined. Marking was designed with three response categories: category one (0 marks), category two (1 mark) and category three (2 marks). To gain 1 mark, students named either traditional Aboriginal music or Rock music. To gain two marks, students named both traditional Aboriginal and Rock music. Results showed that very few students named only one or the other. They either didn't name any or, if they knew their music, they named both (see Figure 6.11 below). Categories one and two were collapsed, thus making item MA18 a two-response category item. A re-analysis showed that the thresholds were now ordered properly.

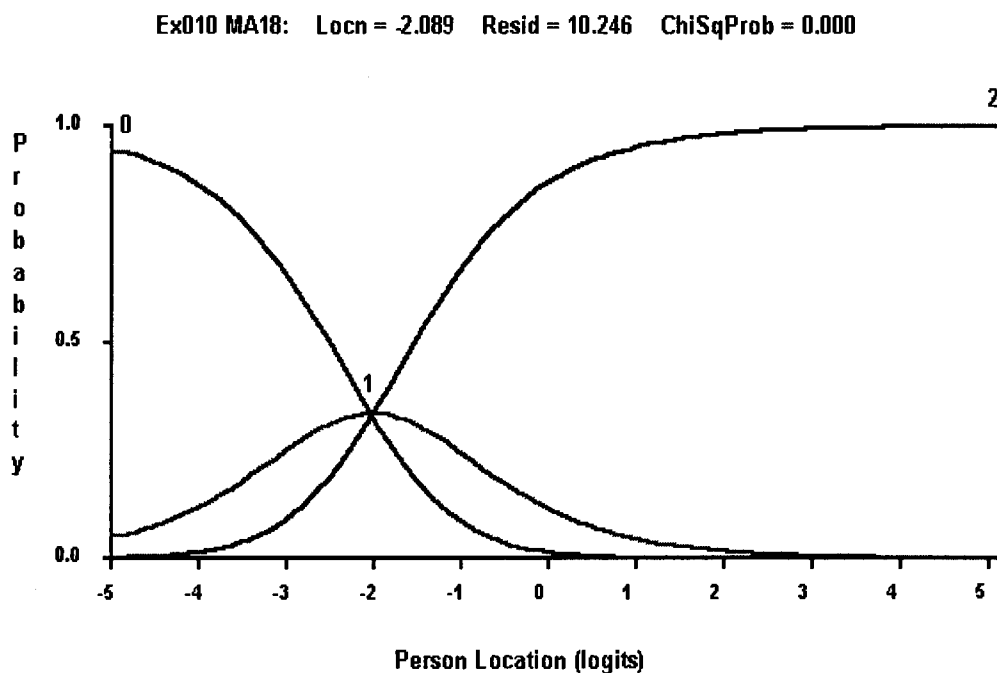


Figure 6.11: Category Probability Curve for Item MA18

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 is around 0.9 and if their music ability is located at 0.9 logits, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that for no student group is there a probability of getting 2 marks.
3. The Category 2 curve indicates that, if a student's music ability is located at around -5.0 logits, the probability of getting 2 marks is 0 and if their music ability is located at 5.0 logits, the probability of getting 2 marks is 1.0.

Item MA23 was the fourth Analysis item to display a reversal of thresholds. After listening to the two pieces, *Ballet for Children* (Bliss, 1995) and *Tribal Voice* (Kellaway & Yunupingu, 1992) students were asked to compare and contrast the two pieces in relation to rhythm. The marking key was developed to contain five categories, 0 marks, 1 mark, 2 marks, 3 marks and 4 marks. To gain 1 mark, students made basic reference to the type of rhythm in the first piece only. For instance, "it had light and bouncy rhythm." To gain 2 marks they made reference to the change, or lack of change of beat in the first piece. For instance, they said it changed from a march to a waltz. To gain 3 marks, students were required to compare the pieces, making "reference" to the effect of the changes in rhythm in both pieces. To gain 4 marks, students were required to compare the pieces in the same way "using music specific language." It seems that, if students are capable of making comparisons between two pieces of music in relation to rhythm, then they usually use music-specific language. Very few students referred to the effect of the changes in rhythm without using musical terminology and so there was no discrimination between categories three and four (see Figure 6.12).

Categories three and four were collapsed so that students received 3 marks for making reference to the effect of the changes in rhythm, with or without the use of music specific language, and a re-analysis showed that the thresholds were now properly ordered.

There are four Process items which had reversed thresholds. The first was item MP02. Apart from two critique items, the Process test, including

Item MP02, was completed by students in small groups of four. This item related to the plan developed by the group for their performance. After viewing a stimulus and being involved in a brainstorming and discussion session with the teacher, students, in groups of four, completed a written or illustrated plan for their performance. The item contained six response categories, so that a group could be allocated 0, 1, 2, 3, 4 or 5 marks for their plan. To gain 4 marks, groups developed a clear, easy-to-follow musical score

Ex015 MA23: Locn = 1.373 Resid = 2.792 ChiSqProb = 0.000

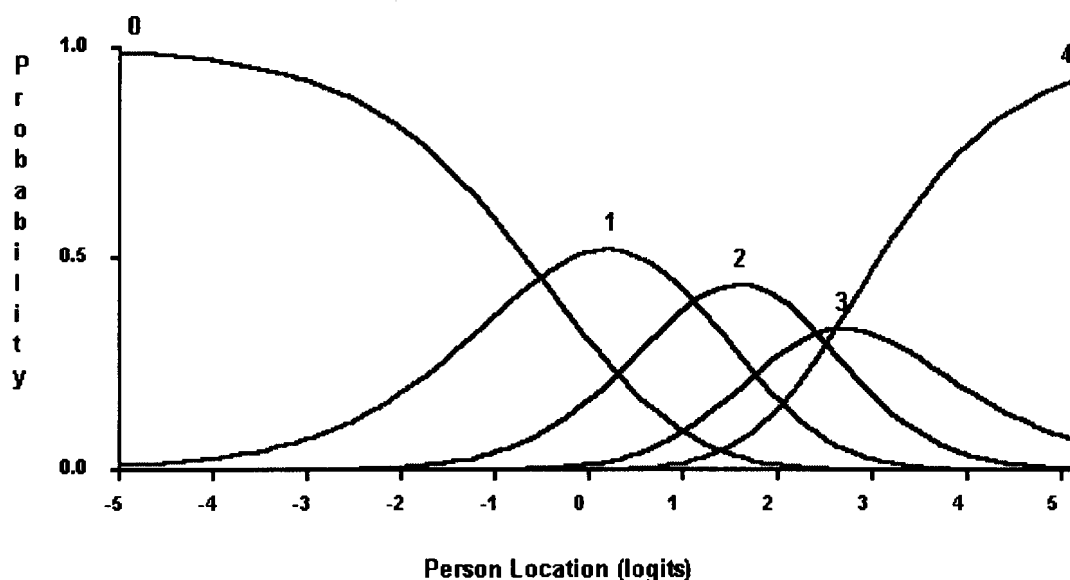


Figure 6.12: Category Probability Curve for Item MA23

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 marks is 1.0 and if their music ability is located at 2.0 logits, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that, if a student's music ability is located at -5.0 , the probability of getting 1 is 0 and if their music ability is located at around 0.0 logits, the probability of getting 1 mark is around 0.5.
3. The Category 2 curve indicates that, if a student's music ability is located at -2.5 logits, the probability of getting 2 marks is 0 and if their music ability is located at 1.5 logits, the probability of getting 2 marks is around 0.3.
4. The Category 3 curve indicates that, for no student group is there a probability of getting 3 marks.
5. The Category 4 curve indicates that, if a student's music ability is around 0.50 logits, the probability of getting 4 marks is 0 and if their music ability is located at 5.0 logits, the probability of getting 3 marks is around 0.9.

in conventional or unconventional form. In other words, they could use traditional musical terminology and symbols, or they could create their own.

The score should have provided a clear structure of composition that

correlated with their final performance and could be followed or played by others. (This correlation was judged by markers who viewed the group's performance and their score simultaneously). To gain 5 marks, the musical score contained all the same characteristics as the 4 mark category, as well as the use of "correct terminology" and a "variety of musical elements." Results indicated that, if groups were able to fulfil all the requirements to gain 4 marks, then they usually used correct terminology and included a variety of musical elements. This meant there was very little discrimination between categories four and five. (see Figure 6.13). Categories four and five were collapsed, thus making Item MP02 a five response category item. A re-analysis showed that the thresholds were now properly ordered.

The second Process item in which a reversal of category thresholds was evident was item MP04. This was another group item and it related to the group's expression of mood in their composition and performance. It was developed to contain five categories to gain 0, 1, 2, 3 and 4 marks. To gain 1 mark, the group demonstrated slight variation in dynamics (loud and soft) in an attempt to create some feeling of mood. To gain 2 marks, they demonstrated evidence in the use of instruments and a variety in dynamics, melody or tempo. To gain 3 marks, the group "provided evidence in planning and performance of organised structure to reflect mood through variety in such elements as dynamics, melody, tempo and rhythm." To gain 4 marks, they fulfilled all these requirements as well as "including musical terminology."

Results indicate that it was more likely for students to produce a response in the fifth category than in the fourth. This means that, if students were able to provide the evidence required to gain 3 marks, they usually included musical terminology (see Figure 6.15). Categories four and five were collapsed to a 3 mark response, thus making Item MP04 a four category item. A re-analysis showed that the thresholds were now properly ordered.

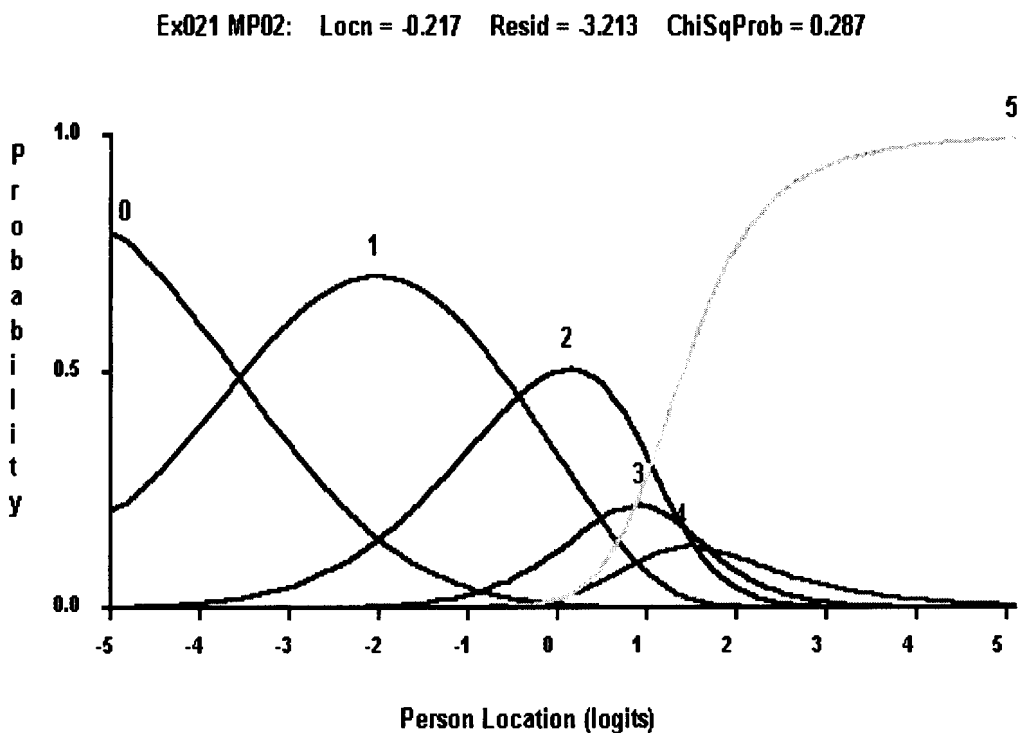


Figure 6.13: Category probability Curve for Item MP02

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 marks is around 0.8 and if their music ability is located at 0 logits, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that, if a student's music ability is located at -5.0, the probability of getting 1 is around 0.2 and if their music ability is located at -2.0 logits, the probability of getting 1 mark is around 0.7.
3. The Category 2 curve indicates that, if a student's music ability is located at -4.8 logits, the probability of getting 2 marks is 0 and if their music ability is located at 0 logits, the probability of getting 2 marks is around 0.5.
4. The Category 3 curve indicates that, for no student ability group is there a probability of getting 3 marks.
5. The Category 4 curve indicates that, for no student ability group is there a probability of getting 4 marks.
6. The Category 5 curve indicates that, if a student's music ability is -0.6 logits, the probability of getting 5 marks is 0 and if their music ability is located at 5 logits or better, the probability of getting 5 marks is 1.0.

The third Process item which had reversed thresholds was item MP05 which related to the extent to which the group expressed the element of form (or structure) in their planning and performance. Item MP05 was a five response category item for which students gained 0, 1, 2, 3 or 4 marks. Students received 1 mark for an attempt to show a beginning, middle and end (as for a narrative). They received 2 marks if there was evidence of some organisation within the group to show obvious changes in expression and instrumentation to demonstrate form. They received 3 marks for “displaying a well organised structure to show form through a variety of elements such as

Ex023 MP04: Locn = 0.981 Resid = -13.273 ChiSqProb = 0.000

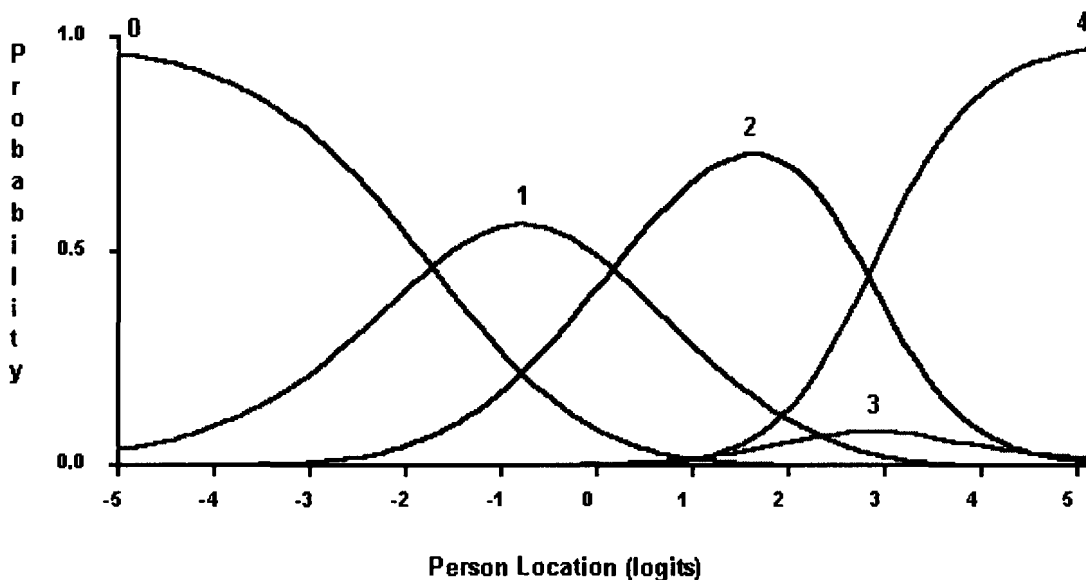


Figure 6.14: Category probability Curve for Item MP04

NOTE:

5. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 is around 0.9 and if their music ability is located at 1.0 logit, the probability of getting 0 marks is 0.
6. The Category 1 curve indicates that, if a student's music ability is located at -5.0, the probability of getting 1 is around 0.1 and if their music ability is located at -1.0 logits, the probability of getting 1 mark is around 0.5.
7. The Category 2 curve indicates that, if a student's music ability is located at -2.5 logits, the probability of getting 2 marks is 0 and if their music ability is located at 1.5 logits, the probability of getting 2 marks is around 0.7.
8. The Category 3 curve indicates that, for no student ability group is there a probability of getting 3 marks.
9. The Category 4 curve indicates that, if a student's music ability is located at around 0.5 logits or less, the probability of getting 4 marks is 0 and if their music ability is located at 5.0 logits, the probability of getting 4 marks is 1.0.

dynamics, tempo, rhythm and texture as well as organisation within the group.” To gain 4 marks, they fulfilled these criteria as well as “providing evidence of leadership.” Results indicate that it was more likely for groups to gain 4 marks than 3 marks. This means that, if the group’s composition and performance displayed a well organised structure to show form through a variety of elements, then they usually displayed evidence of leadership (Figure 6.16). Categories four and five were collapsed to attract 3 marks, thus making MP05 a four category item. A re-analysis showed that the thresholds were found to be ordered properly with four categories.

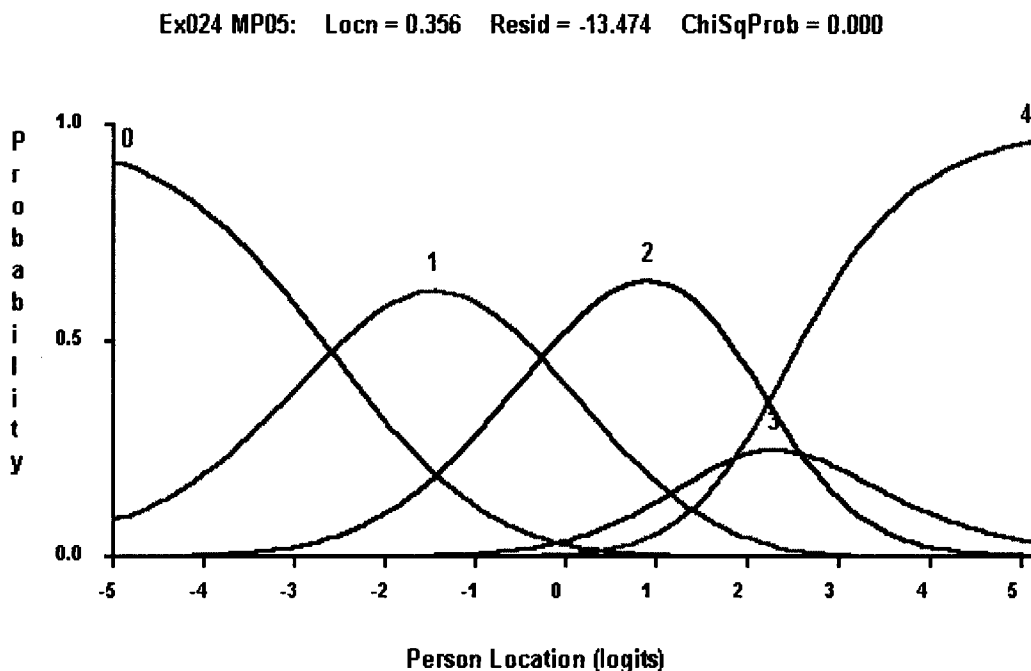


Figure 6.15: Category Probability Curve for Item MP05

NOTE:

1. The Category 0 curve indicates that, if a student’s music ability is located at -5.0 logits, the probability of getting 0 is around 0.9 and if their music ability is located at 1.0 logit, the probability of getting 1 mark is 0.
2. The Category 1 curve indicates that, if a student’s music ability is located at -5.0 , the probability of getting 1 is around 0.1 and if their music ability is located at -1.8 logits, the probability of getting 1 mark is around 0.6.
3. The Category 2 curve indicates that, if a student’s music ability is located at -4.0 logits, the probability of getting 2 marks is 0 and if their music ability is located at 1.0 logits, the probability of getting 2 marks is around 0.7.
4. The Category 3 curve indicates that, for no student ability group is there a probability of getting 3 marks.
5. The Category 4 curve indicates that, if a student’s music ability is around 0 logits, the probability of getting 4 marks is 0 and if their music ability is located at 5.0 logits or better, the probability of getting 4 marks is 1.0.

The fourth Process item that had reversed thresholds was Item MP08, which related to the extent to which the group used the musical element of harmony in their composition and performance. The Item was developed with five categories, attracting 0, 1, 2, 3 and 4 marks. To gain 1 mark, the group created sounds that were simultaneous and linked, and there was evidence of slight contrast to reflect the stimulus. To gain 2 marks, there was evidence of an attempt to create layers of sound and contrast. To gain 3 marks, the group displayed evidence of the ability to “organise sounds to provide contrast of harmony with two or more melodies and clear evidence of texture.” To gain 4 marks for this item, the performance displayed a “pleasing sense of harmony, obviously planned and organised, through use of melodies, texture, contrast and a sense of complementary sounds” in the construction of form. Further examination revealed that it was difficult for markers to differentiate between the ability to “organise sounds to provide contrast of harmony with two or more melodies and clear evidence of texture” and “a pleasing sense of harmony, obviously planned and organised, through use of melodies, texture, contrast”. It appeared easier for students to achieve Category five than Category four (see Figure 6.17). Categories four and five were collapsed so that students received 3 marks if they “organised sounds to provide contrast of harmony with two or more melodies and clear evidence of texture,” or if their composition “displayed a pleasing sense of harmony,

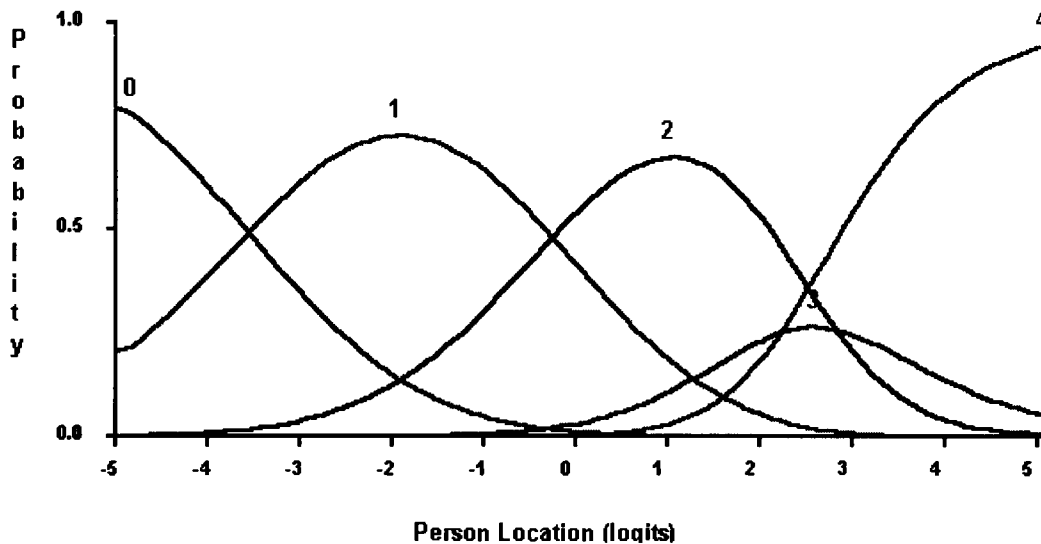


Figure 6.16: Category Probability Curve for Item MP08

NOTE:

1. The Category 0 curve indicates that, if a student's music ability is located at -5.0 logits, the probability of getting 0 is around 0.8 and if their music ability is located at 0.0 logits, the probability of getting 0 marks is 0.
2. The Category 1 curve indicates that, if a student's music ability is located at -5.0 , the probability of getting 1 is around 0.2 and if their music ability is located at -2.0 logits, the probability of getting 1 mark is around 0.7.
3. The Category 2 curve indicates that, if a student's music ability is located at -4.0 logits, the probability of getting 2 marks is 0 and if their music ability is located at 1.0 logits, the probability of getting 2 marks is around 0.6.
4. The Category 3 curve indicates that, for no student ability group is there a probability of getting 3 marks.
5. The Category 4 curve indicates that, if a student's music ability is around 0.5 logits, the probability of getting 4 marks is 0 and if their music ability is located at 5.0 logits or better, the probability of getting 4 marks is 1.0.

obviously planned and organised, through use of melodies, texture, contrast and a sense of complementary sounds." A re-analysis showed that the thresholds were found to be ordered properly after the collapsing of these categories.

Scale values

The second parameter in the hierarchical order is scale (θ). While the first order item parameter; location (δ), specifies the average location of the item on the continuum, the second order item parameter; scale (θ), is associated with the dispersion of the item categories. It provides information about the average spread of the thresholds for each item and can be estimated if three or more categories per item are present. θ = average half-

threshold distance between categories for each item. Ideally, θ should be equal for each item. If scale values are not reasonable, it may be necessary to revisit the wording of items, or look at student responses to determine a reason why categories are not spread equally. While the Rasch model indicates an anomaly, there is no specific theory on how to correct it. The following table provides examples of items in which scale (θ) is reasonable (Table 6.1)

Table 6.1: Item Parameters for Location and Scale

Item	Location		Scale	
	Estm	SE	Estm	SE
MA19	.881	0.039	1.686	0.028
MA25	.435	0.069	1.453	0.069
MP13	-1.313	0.037	1.678	0.037
MP14	-.428	0.039	1.398	0.034
MP26	-.546	0.032	1.158	0.024

The item parameters for the combined Analysis and Process tests are shown as Appendix xvi. In this study, attention is focused on the location and scale parameters.

Power of tests-of-fit

After establishing the item and student estimates and ensuring ordered thresholds, the degree to which the estimates fitted with the conceptual framework was examined. This was done first, by examining the degree to which the expected values differed from the observed data values using item-trait test-of-fit statistics. A chi-square statistic was derived to assess the probability of the degree of divergence between observed and expected

values occurring by chance alone. As well as individual statistics being produced for each item, the data were used across all items to obtain an overall test-of-fit statistic. Second, an item-student interaction test-of-fit was used to examine the degree to which students responded to items of differing levels of difficulty in a logical and consistent manner. These fit statistics approximate a t-distribution when the data fit the model. After rescoring of items, the summary statistics for the test-of-fit are as follows (Table 6.2):

Table 6.2. Test-of-Fit summary statistics

Items	Item-Student Interaction		Persons	
	Location	Standard error	Location	Standard er
Mean	0.000	-1.377	-0.899	-0.248
SD	1.214	6.132	1.109	1.198

The total item Chi Square, after the rescoring of items, is 3456.156 and the Person separation index is 0.900 with total degrees of freedom of 450.000. The total Chi Square Probability is 0.000 and the Test of Fit Power, according to the RUMM (1996) scale is “Excellent” with a reliability of 0.928.

The analysis was undertaken in three different ways in an effort to maximise the reliability and the power of the test of fit. First, the data were analysed using ‘Analysis’ test items only; second, the data were analysed using ‘Process’ test items only and; third, the data were analysed as a combined test of ‘Analysis’ and ‘Process’ items.

Examination of the items when analysing the Process test separately, revealed that the twelve planning and critique items in the Process test showed some misfit to the model. The reason for this was that these items were derived from the '*Appreciating*' strands of the outcome statements and were more compatible with the 'Analysis' test items, whereas the 'true' Process items that involved exploration and performance were derived from the '*Creating*' strands. In addition to this, the 'true' process items showed statistical fit with the 'Analysis' items. For these reasons, the combined sets of Process and Analysis items appeared to have advantages both from the overall perspective and from the perspective of individual item fit. It allowed all items to be located on a single scale and, hence, provided a single estimate of music achievement for each student. The combination of both tests produced the Excellent power of test of fit as rated by RUMM (1996).

Increasing the Probability

The relationship between student ability and item difficulty that had been established using the RUMM (1996) program, reflected a probability of 0.5 that a student would correctly respond to an item of equal difficulty. This is usual for Rasch models. However, to increase the reliability of a student's success in achieving at an outcome level, the probability of success upon which to base an ability estimate was converted, at this stage, from 0.5 to 0.7. The consequence of this was that a student now had at least a 70% chance of correctly responding to an item with a difficulty of equal to or less than the ability of the student. This meant that, when outcome levels were allocated, it

could be said, with more confidence, that a student was performing at a particular outcome level.

Establishment of levels

The simultaneous scaling of item difficulties and student abilities on the same scale derived from a Rasch model of analysis allows for the linking of student performance to *The Arts Student Outcome Statement* (Education Department of Western Australia, 1996) levels. After item difficulties were established, the distribution of items was plotted on an item map. The outcome level that was intended to be operationalised by the item was then entered next to each item.

A study of the item clusters on the item map (Appendix xvii) was then undertaken to establish outcome level boundaries. It was expected that samples of items operationalising an outcome level would be normally distributed with respect to difficulty and that items operationalising a high level outcome would be more difficult than those operationalising a lower level outcome. This was the case in most instances, apart from a few items which did not perform exactly as expected. In some instances, there are some outlying items that do not fit into the intended level (see Appendix xvii). There could be a number of reasons for this. For instance, the format of a question may have created an anomaly that made a question easier or more difficult than anticipated. An item which was intended to operationalise a Level three outcome may have been easier for students because it was presented as a multiple choice format. On the other hand, an item that was designed to

operationalise Level two may have been more difficult because it was an open-ended format. Unfamiliar or unusual wording may have had a similar effect.

The ability levels of students were then plotted on the item map and a study was made of the test scripts of students whose ability levels were close to the level boundaries. This study was carried out by a music expert who had played a significant role in the development of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), the researcher, who was involved in the same project, and two measurement officers with wide experience in system level testing. The overall scores of student test performances were compared to the study of item clusters to ensure that students whose ability estimates appeared within the boundary of a particular outcome level had a high probability of being at that level. Level cut-off points were established and Student performances could now be viewed in relation to *The Arts Student Outcome Statement* (Education Department of Western Australia, 1996). It should be noted that there is an overlapping of levels. This is to be expected, given that progression from one level to another is a developmental process. As student ability estimates approach the upper or lower boundary limits of a level, the probability of their being in that level decreases.

Transforming the logit values

For the purposes of reporting, and to eliminate the use of negative values for student ability, the logit scale was converted to an arbitrary scale

from 0 to 800. The scale of 800 was selected to reflect the eight levels of outcomes contained in *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). After being adjusted to 0.7 probability, the minimum logit value of the sampled students was -3.75 . This minimum logit value was transformed to the arbitrary scale score of 0. The maximum logit value, after being adjusted to 0.7 probability was 4.56. This maximum logit value was transformed to the arbitrary scale score 800. The equation used to perform this conversion is $800/[\text{logit}_{\text{max}} - \text{logit}_{\text{min}}]$.

Table 6.3: Conversion of logit values to scale values

Logit values 3.75 – 4.56	Scale values 0 - 800
-3.75	0
-3.0	72.2
-2.0	168.5
-1.0	164.7
0	361
1.0	457.3
2.0	553.5
3.0	649.8
4.0	746.1
4.56	800

Student performance levels

The mean level for each Year group shows a clear pattern of development from Year 3 through Year 7 to Year 10, although there is considerable overlap in performance between the Year groups. For example,

the highest achieving 10 per cent of Year 7 students performed above the level demonstrated by approximately 25 per cent of Year 10 students.

Over 80 per cent of Year 3 students demonstrated skills associated with Level 2 outcomes in music. This means they can work in a group to plan and create a simple sound piece to interpret a given stimulus, including the creation of a simple score, notating their own rhythms, melodies and accompaniment patterns using simple known methods. They reflect upon music works, noting particular features including melody, instruments used, form and expression. They identify the purpose of a work and how it affects the way it should be performed. They apply simple critical reflections on their preferences and describe sounds using basic musical terms.

Over 55 per cent of Year 7 students demonstrated skills associated with Level 3 outcomes in music. This means they can compose short, simple, structured musical works using tuned or untuned percussion instruments, recorder, sounds from the environment, voice and body percussion. They are able to aurally recognise and describe musical features such as simple rhythmic and melodic patterns, tempo, instrumentation, timbre, dynamics and structure and use and interpret signs and symbols representing pitch, duration of sound and dynamics. They can describe obvious features such as repetition, form, changes in dynamics and texture, as well as identifying music from another culture and associating characteristics of the music with the style.

Over 80 per cent of Year 10 students demonstrated skills associated with Level 4 outcomes in music. This means they can create musical works that capture characteristics of a given stimulus and interpret elements of pitch, rhythm, dynamics and phrasing in composition. They explore major and minor tonalities, textures, forms, media, and invent a soundscape score related to the theme. They explore combinations of sounds from the environment, chords, ostinati, and incorporate known structures such as ternary or binary form. They are able to give reasons why a musical element used in a piece is important and how it was used to create the perceived mood, tension and purpose. They can compare music from different times, places or cultures, identifying notable differences in musical characteristics. Below is a summary of the overall performance of Year 3, 7 and 10 students in music (Table 6.4) and the scale of student performance and outcomes achieved (Figure 6.17).

Table 6.4: Summary of student performance in music

	Number of students	Mean	Standard deviation	Level of mean
Year 3	946	294	85	2
Year 7	921	359	82	3
Year 10	324	525	105	4/5

Student Performance in Music

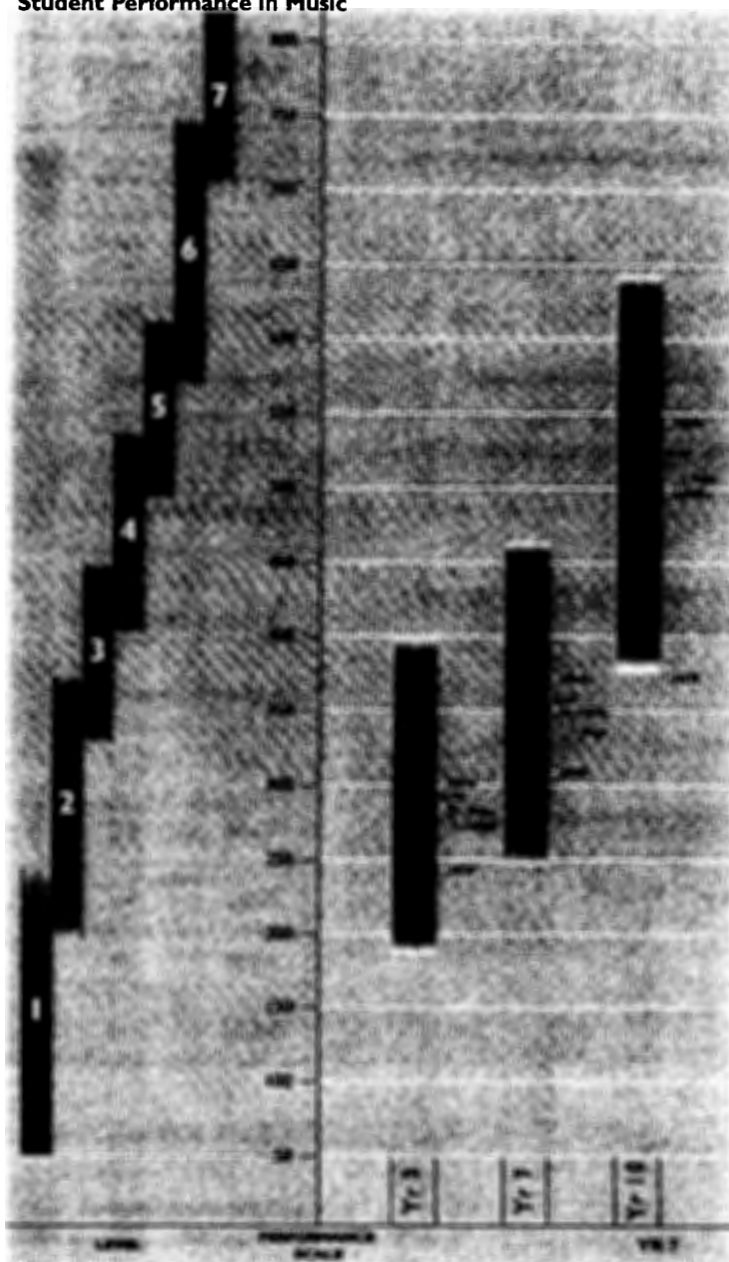


Figure 6.17: Student performance in music by level and mean score
For girls, boys, NESB, ATSI and all students

- NOTE:** 1. ATSI = Mean scale score and level for Aboriginal Torres Strait Islander students (Yr 3 N=59, Yr 7 N=44, Yr 10 N=17)
 2. NESB = Mean scale score and level for non English speaking background students (Yr 3 N=122, Yr 7 N=114, Yr 10 N=41)
 3. Girls = Mean scale score and level for girls (Yr 3 N=426, Yr 7 N=397, Yr 10 N=172)
 4. Boys = Mean scale score and level for boys (Yr 3 N=486, Yr 7 N=487, Yr 10 N=139)
 5. All = Mean scale score and level for the total number of students tested in the relevant Year group (Yr 2 N=946, Yr 7 N=921, Yr 10 N=324)

Implications

The tests have been administered by both generalist and specialist music teachers in schools and are suitable for use by either group. Teachers will be able to use students' raw scores to compare their results with the data gathered across the State for this testing program. Outcomes which relate to aesthetics, critical analysis, interpretation of meaning and music concepts have not been measured with any level of reliability in Western Australian classrooms before, nor has there been any opportunity for teachers to make comparisons using a common framework. These tests will provide these opportunities, as well as providing a model of good classroom practice based on *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) framework.

Marking keys and item descriptions have been worded to provide descriptions that can be understood by generalist as well as specialist teachers at both primary and secondary levels. Although the tests were designed for trialling at Years 3, 7 and 10, they have been developed to reflect a developmental continuum and so can be administered by teachers in other levels as well. This means that, although comparisons with State means at particular Year levels are not possible, the tests can be used at any Year level as a valuable tool for gathering classroom or whole school data in relation to *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996).

Conclusions

The Music Achievement Scale, to measure student outcomes in classroom music across both the *Appreciating* and the *Creating* strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), for Years 3, 7 and 10, has been successfully developed. Validity of the measure has been established by trialling the materials with a sample of 2192 students in Western Australian primary and secondary schools and conducting a Rasch model of analysis using the RUMM (1996) program. Item difficulties have been calibrated on the same scale as the student measures. Overall fit, as well as individual fit, of items to the model has been established. Thresholds have been adjusted where necessary, so that they are properly ordered. Reliability of the scale, as shown by its Separability Index, is high and the power of fit to the model is excellent. Targeting of the items against the student measures is satisfactory. This evidence leads to the conclusion that a valid and reliable scale of music achievement has been created.

The matching of the music achievement scale to outcome level statements indicates that more than 80 per cent of Year 3 students demonstrated skills consistent with Level 2 outcomes. In a group situation, they can plan, notate and perform a simple piece using simple known methods. They are able to listen to and reflect upon excerpts from simple music works and apply simple critical reflections.

Year 7 students' skill levels were spread mainly between Level 2 and Level 4, with over 55 percent demonstrating skills associated with Level 3 outcomes. At Level 3 they can compose short, simple structured musical works using tuned or untuned percussion instruments, recorder or 'found' sound sources. They can aurally recognise and describe simple musical elements as well as obvious features such as repetition, form, dynamics and texture. They can also identify music from a familiar culture, associating characteristics of the music with the style. At Level 4, they creatively solve problems within given structures to create music works that capture the characteristics of a given stimulus and interpret elements of pitch, rhythm, dynamics and phrasing in composition. They explore more sophisticated features of major and minor tonalities, textures, form and media.

Over 80 per cent of Year 10 students also demonstrated skills associated with Level 4 outcomes or higher. Those who demonstrated Level five outcomes are able to explore and develop personal ideas and draw on links to other societies, cultures and times, combining appropriate skills to plan, shape, create and share musical works. They are able to analyse musical works, identifying and discussing key features which locate them in particular societies, cultures or times.

CHAPTER SEVEN
DATA ANALYSIS (PART B):
SUB-GROUPS OF THE SCALE

Introduction

Reliable measures of the performances of sub-groups in all learning areas, including music, would be useful to school administrators in developing the Management Information Systems necessary for school accountability. It would also assist them in future planning and resource allocation, when they would need reliable data about the performances of sub-groups in all learning areas, including music. An awareness of the differential performances in music of boys and girls, students from English speaking backgrounds (ESB) and non-English speaking backgrounds (NESB), Aboriginal Torres Strait Islander students (ATSI) and non-Aboriginal Torres Strait Islander students, would be valuable to both music teachers and principals in providing equal opportunities for these groups to learn.

The collection of data on the sub-groups identified above was a significant part of this study. In an attempt to ensure that tests and stimuli contained no material that was biased towards or against any particular group, advice was sought from personnel with expertise in gender equity, social justice and Aboriginal education throughout the item-development process. Panels examined stimulus materials to ensure there was a balance in style suited to both males and females, as well as cultural minority groups.

They also examined the language used to frame questions to ensure it was suited to the age groups of students, as well as to reduce gender or cultural bias. It was difficult, however, to predict the level of students' literacy skills. For this reason, teachers' guidelines for administering the tests specified that, in Year 3, questions should be read aloud to students, and in Years 7 and 10, teachers should assist students to read questions where necessary. At all Year levels, teachers were instructed to answer any questions related to the comprehension of the question, without actually providing answers. In extreme circumstances, teachers could scribe for students who could not write. In cases where extended responses were required, however, it is still possible that the poor writing skills of some students had an effect on their results. Student background questionnaires were used, during the testing process, to identify sub-groups.

The results of sub-group performances should be treated with caution as the numbers of some sub-groups tested were very low compared with the total cohorts tested. The numbers of sub-groups are indicated in Figure 6.17, "Student performance in music" on page 170.

The student sample for the study exposed some significant access and equity issues at the Year 10 level. It was revealed that there are limitations to the range of arts disciplines, including music, to which Year 10 students have access in Western Australian government schools. Despite the fact that the Arts is now one of the eight compulsory learning areas, of the 171 government secondary and district high schools in Western Australia, there were still 58

schools where access to learning in the Arts was not available at the time the sample was drawn (Pascoe, 1997a, 1997b; see table 7.1).

Table 7.1: Number of schools offering arts disciplines at Year 10 in semester one 1996

No: of arts subjects offered in Year 10	No: of schools with Year 10 in Western Australia
5	32
4	25
3	19
2	20
1	17
0	58

The breakdown of student population enrolled in each of the Arts disciplines varies significantly. The percentage of the total population enrolled in music at Year 10 in government schools is lower than the percentage enrolled for any other Arts discipline (Pascoe, 1997b; see table 7.2).

Table 7.2. Arts enrolments for Year 10, Semester one, 1996

DISCIPLINE		TOTAL ENROLLED	TOTAL POPULATION
Dance	n	1974	17140
	% of population	11.5	
Drama	n	3519	17140
	% of population	20.5	
Media	n	1938	17140
	% of population	11.3	
Music	n	1374	17140
	% of population	8.0	
Visual Arts	n	5812	17140
	% of population	33.9	

NOTE: n = the number of Year 10 students enrolled in the discipline in Western Australian government schools
 % of population = the percentage of the total Year 10 population enrolled in the discipline
 Total population = all students enrolled in Year 10 in Western Australian Government schools

In view of the Education Department of Western Australia's policy to ensure an inclusive curriculum for all students, the low proportion of male enrolments to female enrolments in Arts programs (Pascoe, 1997b), should be cause for some concern to administrators and teachers. Available data indicates that there is a dramatic discord between student participation in the Arts at secondary school, and employment opportunities in the Arts (Davis &

Nunan, 1992). While the number of boys studying the Arts in Western Australian government secondary schools varies from 1.5 per cent of the total enrolment (for dance), to 44.1 per cent of the total enrolment (for media), the proportion of males to females currently in paid employment in the Arts in Australia is around 67 per cent for males and 33 per cent for females (Davis & Nunan, 1992). Girls participate in a variety of Arts disciplines in government schools in Western Australia to a greater extent than Year 10 boys (see table 7.3).

Table 7.3: Year 10 Arts enrolments by gender, Semester one, 1996

DISCIPLINE		FEMALES ENROLLED	MALES ENROLLED
Dance	n	1944	30
	% of enrolled	98.5	1.5
	% of population	11.3	0.2
Drama	n	2516	1003
	% of enrolled	71.5	28.5
	% of population	14.7	5.9
Media	n	1084	854
	% of enrolled	55.9	44.1
	% of population	6.3	5.0
Music	n	791	583
	% of enrolled	57.6	42.4
	% of population	4.6	3.4
Visual Arts	n	3759	2053
	% of enrolled	64.7	35.3
	% of population	21.9	12.0

NOTE: n = the number of Year 10 females or males enrolled in the discipline in Western Australian government schools
 % of enrolled = the percentage of Year 10 females or males of the total number enrolled in the discipline
 % of population = the percentage of Year 10 females or males of the total number of students in Year 10 in Western Australian Government schools

The specialised nature of music teaching compounds the problem of access to music education in remote areas of Western Australia where it is not feasible to install a specialist music teacher to serve a small number of students. This means that most of the 1 600 Aboriginal Torres Strait Islander students in the 29 Remote Community Schools do not have access to specialised classroom music.

At Years 3 and 7, there is no apparent difference in access to music between boys and girls or English speaking background students and non-English Speaking background students. Access to the Arts, where they are available in primary schools, is the same for all students. The remoteness of some schools attended by Aboriginal Torres Strait Islander students in Year 3 and 7, however, has an effect on access to music for these students.

Sub-group analyses

The RUMM (1996) program produced separate analyses for sub-groups, with a differential item function providing the opportunity to compare the results of sub-groups. Comparisons were made between girls and boys, Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students, and English speaking background and non-English speaking background students, at each Year level, on individual items as well as on the overall results of the Analysis and Process tests combined.

Differential item performance

An examination of individual items was made for differential item performances among sub-groups. In the case of all sub-groups, there is a quantitative difference in overall test scores and there is a trend that favours one group. Where a group performed against the trend on a particular item, or where differences were extreme, the item was examined qualitatively for evidence of bias against the group, or for features of the item that appeared to be favourable to the group's performance. Items on which sub-groups

performed against the trend, or items in which differences were extreme, will be discussed.

Year 3 sub-group analysis – males/females

At the Year 3 level, girls significantly outperformed boys in performance on the total results of the Analysis test and Process test combined. The mean score of achievement, on the arbitrary scale of 0-800 (described in Chapter six), for Year 3 girls was 302.7, with a standard deviation of 82.6, and the mean score for boys was 286.2, with a standard deviation of 82.5. A two tail t-test (probability of 0.003) indicates that the difference in performances between the two groups is significant. The results indicated that 7.5 per cent of girls performed at Level 4 or above, compared with 4.5 per cent of boys. There were 26.1 per cent of girls who performed at Level 3 or above, compared with 19.3 per cent of boys. Among the lower achievers, 22.6 per cent of boys performed at below Level 2, whereas only 15.5 per cent of girls performed below Level 2 (see table 7.4). A graphical display of the frequencies of score ranges for Year 3 girls and boys is given in Appendix xviii.

Table 7.4: Percentages of Year 3 girls and boys attaining SOS levels

SOS level	Score Ranges	Percentage	
		Girls	Boys
Level 1	0 - 220	15.5%	22.6%
Level 2	221 - 350	58.2%	57.9%
Level 3	351 - 425	18.6%	13.9%
Level 4 & above	426 +	7.7%	5.6%

Differential item performances – Year 3 males and females

The scores of Year 3 girls were higher than those of boys on all items of both the Analysis and Process tests. This was a general trend and no particular item indicates a significant deviation from the trend. The item that displayed the greatest difference in performance between girls and boys was item MA07. The mean score for girls on this item was 0.944 logits, with a standard deviation of 0.231, and the mean score for boys was 0.907 logits, with a standard deviation of 0.291. The difference in means between the two groups was 0.037 logits which, allowing for a 0.08 error of measurement, is not a big difference. Item MA07 was the easiest of the Analysis items with a difficulty of -3.373 logits. Students were required to indicate whether the beat “stays the same or changes in this part of the music” after listening to an excerpt from the piece *Ballet for Children* (Bliss, 1995). Although the difference between boys and girls was the greatest for this item, it follows the overall trend where Year 3 girls do slightly better than Year 3 boys.

Year 3 sub-group analysis – Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander Students

At the Year 3 level, non-Aboriginal Torres Strait Islander students significantly outperformed Aboriginal Torres Strait Islander students in performance on the Analysis test and Process test combined. The mean score, on the arbitrary scale of 0-800, for Year 3 non-Aboriginal Torres Strait Islander students was 297.4, with a standard deviation of 82.2, and the mean score for Aboriginal Torres Strait Islander students was 243.1, with a standard deviation of 80.6. A two tail t-test (probability of 0.000) indicates that the performance between the two groups is significantly different. The results indicate that 6.8 per cent of non-Aboriginal Torres Strait Islander students

performed at Level 4 or above, whereas no Aboriginal Torres Strait Islander student performed above Level 3. There were 23.6 per cent of non-Aboriginal Torres Strait Islander students who performed at Level 3 or above, compared with 6.8 per cent of Aboriginal Torres Strait students. Among the lower achievers, there were 37.2 per cent of Aboriginal Torres Strait Islander students who performed at below Level 2, whereas only 18.3 per cent of non-Aboriginal Torres Strait Islander students performed below Level 2 (see table 7.5). A graphical display of the frequencies of score ranges for Year 3 ATSI students and non-ATSI students is given in Appendix xix.

Table 7.5: Percentages of Year 3 ATSI and non-ATSI students attaining SOS levels

SOS level	Score Ranges	Percentage	
		ATSI students	non-ATSI students
Level 1	0 - 220	37.3%	18.3%
Level 2	221 - 350	54.2%	57.9%
Level 3	351 - 425	8.5%	16.8%
Level 4 & above	426 - 800	nil	7.0%

Differential item performances – Year 3 Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students

In many cases, the difference in mean logit scores between non-Aboriginal students and Aboriginal Torres Strait Islander students is less than the error of measurement, of 0.08 logits. The test item in which Year 3 non-Aboriginal Torres Strait Islander students most significantly outperformed Aboriginal Torres Strait Islander students was item MA02. This was a relatively easy item, with a difficulty level of -0.641 logits. The logit mean score for non-Aboriginal students on this item was 0.302, with a standard

deviation of 0.358 logits, whereas the mean score for Aboriginal Torres Strait Islander students was 0.131 logits, with a standard deviation of 0.283 logits. The difference in mean between the two groups was 0.172 logits. Even allowing for 0.08 error of measurement, this is still a significant difference. The item required students to justify their choice of answer in the previous item in which they were asked where they would be most likely to hear the piece of music, *Ballet for Children* (Bliss, 1995) which was played to them. The choices were, birthday party, orchestral concert, street parade or rock concert. Students were required to explain what they heard in the music that made them pick their answer. The item was examined for evidence to explain the relatively high difference in performance between the two groups. The item was structured in an extended answer format which means it required a higher level of writing skills, than a multiple choice or one word response item construct, and this may have had an adverse affect on the results of Aboriginal Torres Strait Islander students whose first language was not English. Another explanation for the large difference, may be that, for students in remote areas, the four choices, "birthday party, orchestral concert, street parade and rock concert" were unfamiliar to them. This would have made it difficult for them to explain why they had selected their answer.

There was no deviation from the general trend in the Process tests, that is, non-Aboriginal Torres Strait Islander students outperformed Aboriginal Torres Strait Islander students on all items. A possible explanation for this is that the lack of specialist music teachers in remote areas could mean that

these students have had less experience in performance than some other students in larger metropolitan schools.

Year 3 sub-group analysis – English Speaking Background students and non-English Speaking Background students

English speaking background students significantly outperformed non-English speaking background students at the Year 3 level in performance on the Analysis test and the Process test combined. The mean score, on the arbitrary scale of 0-800, for Year 3 English speaking background students was 298.7, with a standard deviation of 82.8, and the mean score for non-English speaking background students was 282.5, with a standard deviation of 84.3. A two tail t-test (probability of 0.046) indicates that the difference in performances between the two groups is significant. The results indicate that 7.1 per cent of English speaking background students performed at Level 4 or above, compared with 4.1 per cent of non-English speaking background students. There were 24.1 per cent of English speaking background students who performed at Level 3 or above, compared with 14.8 per cent of non-English speaking background students. Among the lower achievers, there were 18.0 per cent of English speaking background students who performed at below Level 2, whereas 20.5 per cent of non-English speaking background students performed below Level 2 (see table 7.6). A graphical display of the frequencies of score ranges for Year 3 English background students and non-English speaking background students is given in Appendix xx.

Table 7.6: Percentages of Year 3 ESB and NESB students attaining SOS levels

SOS level	Score Ranges	Percentage	
		ESB students	NESB students
Level 1	0 - 220	18.0%	20.5%
Level 2	221 - 350	57.6%	60.6%
Level 3	351 - 425	16.8%	14.0%
Level 4 & above	426 +	7.1%	4.9%

Differential performances of Year 3 English speaking background students and non-English speaking background students

The data indicates that, although Year 3 English speaking background students outperformed non-English speaking background students on most items, there were two items that went against this trend. The first of these was item MA08 in which students were asked to explain what they had heard in the excerpt from the music *Ballet for Children* (Bliss, 1995) to justify whether they thought the beat had changed or stayed the same. It was an open-ended response item and it was the most difficult of the Year 3 Analysis items with a difficulty level of 1.214 logits. The mean score for English speaking students on this item was 0.135 logits, with a standard deviation of 0.240 logits and the mean score for non-English speaking students was 0.169 logits, with a standard deviation of 0.264 logits. Although the difference in performance was only 0.034 logits, the results did go against the general trend and so the item was examined. There is no apparent explanation for a reverse in trend in the performance of English speaking and non-English speaking students on this item. The difference may be attributable to measurement error.

The second item that went against the general trend was Process item MP07 on which markers were asked to indicate to what extent the group had used expression in their performance. The item had a difficulty level of 0.617 logits. The mean score for English speaking background students was 0.188 logits, with a standard deviation of 0.203 logits and the mean score for non-English speaking background students was 0.196 logits, with a standard deviation of 0.202 logits. Although the difference in performance was only 0.008 logits, the item was examined for an apparent reason for a reversal in trend. No explanation for the reversal could be found and, given that the difference between the two groups was so small, it is possibly due to error of measurement.

The following table provides a summary of sub-group performances in music in Year 3 (Pascoe, 1998, p.42).

Table 7.7 Summary of sub-group performances in music in Year 3

Sub-group	N	Mean	Standard deviation
All	946	294	85
Females	426	303	83
Males	486	286	83
ATSI students	59	243	81
NESB students	122	283	84

NOTE: All = the total Year 3 population tested
 Females = the total number of girls tested
 Males = the total number of boys tested
 ATSI = Aboriginal Torres Strait Islander students tested
 NESB = non-English speaking background students tested
 Mean = shown on the scale of 0 - 800

Year 7 sub-group analysis of males and females

At the Year 7 level, girls significantly outperformed boys in overall performance. The mean score, on the arbitrary scale of 0-800, for Year 7 girls was 372.9, with a standard deviation of 78.3, and the mean score for boys was 349.7, with a standard deviation of 79.5. A two tail t-test (probability of 0.000) indicates that the difference in performances between the two groups is significant. Results indicate that 20.7 per cent of girls performed at Level 4 or above, compared with 15.4 per cent of boys. There were 64.2 per cent of girls who performed at Level 3 or above, compared with 50.9 per cent of boys. Among the lower achievers, 49.1 per cent of boys performed at below Level 3, whereas 35.5 per cent of girls performed below Level 3 (see table 7.8). A graphical display of the frequencies of score ranges for Year 7 girls and boys is given in Appendix xxi.

Table 7.8: Percentages of Year 7 girls and boys attaining SOS levels

SOS level	Score Ranges	Percentage	Percentage
		Girls	Boys
Level 1	0 - 220	3.5%	4.5%
Level 2	221 - 350	32.0%	44.6%
Level 3	351 - 425	41.1%	33.4%
Level 4	425 - 515	20.4%	16.5%
Level 5 & above	516 - 800	3.0%	1.0%

Differential item performances of Year 7 males and females

The trend for both the Analysis test and the Process test was that girls outperformed boys on all items. There was no deviation from this trend,

although the differences in performance were not great and some may be attributable to measurement error. The Analysis test item in which girls most significantly outperformed boys in Year 7 was item MA19. The mean score for girls on this item was 0.325 logits, with a standard deviation of 0.198 logits, whereas the mean score for boys was 0.268 logits with a standard deviation of 0.181 logits. The difference in mean between the two groups was 0.057 logits. The item related to the piece *Dharpa* (Kellaway & Yunupingu, 1992) and was dependent on the previous item (MA18) which required students to indicate which styles of music had been combined by the composer to create the piece. Item MA19 then asked students, "What is it in the music that helped you to decide?" If students did not correctly answer the previous item, it would not have been possible to answer MA19 correctly. It is possible this item's dependence on the previous one has affected results. The item is also in an extended response form, which makes it reliant on students' literacy skills. This could also have affected the performance of boys. However, given that the difference in performance between the two groups was only 0.057 logits, the difference may be due to error of measurement.

Year 7 sub-group analysis of Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students

Non-Aboriginal students significantly outperformed Aboriginal Torres Strait Islander students at the Year 7 level, in overall scores on the Analysis and Process tests combined. The mean score, on the arbitrary scale of 0-800, for Year 7 non-Aboriginal students was 364.0, with a standard deviation of 78.2, and the mean score for Aboriginal Torres Strait Islander students was

308.8, with a standard deviation of 77.2. A two tail t-test (probability of 0.000) indicates that the difference in performances between the two groups is significant. Results indicate that 21.3 per cent of non-Aboriginal students performed at Level 4 or above, whereas 4.5 per cent of Aboriginal Torres Strait Islander students performed at Level 4 or above. There were 59.1 per cent of non-Aboriginal Torres Strait Islander students who performed at Level 3 or above, compared with 29.5 per cent of Aboriginal Torres Strait Islander students, and 21.3 per cent of non-Aboriginal students who performed at Level 4 or above, compared with only 4.5 per cent of Aboriginal Torres Strait Islander students. Among the lower achievers, 70.5 per cent of Aboriginal Torres Strait Islander students performed at below Level 3, and 40.9 per cent of non-Aboriginal Torres Strait Islander students performed below Level 3 (see table 7.9). A graphical display of the frequencies of score ranges for Year 7 ATSI and non-ATSI students is given as Appendix xxii.

Table 7.9: Percentages of Year 7 ATSI and non-ATSI students attaining SOS levels

SOS level	Score Ranges	Percentage	
		Girls	Boys
Level 1	0 - 220	11.4%	3.3%
Level 2	221 - 350	59.1%	37.6%
Level 3	351 - 425	25.0%	37.8%
Level 4	425 - 515	4.5%	19.2%
Level 5 & above	516 - 800	nil	2.1%

Differential item performances of Year 7 Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students

The individual item analysis indicates that most items follow the general trend, with Year 7 non-Aboriginal students outperforming Aboriginal Torres Strait Islander students. There was one item, however, on which there was no difference in performance between the two groups. This was Process item MP22 in which markers indicated the extent to which the group expressed form (i.e. changes in expression, instrumentation, dynamics etc to denote beginning, middle, end). The mean score for both non-Aboriginal and Aboriginal Torres Strait Islander students on this item was 0.417 logits, with a standard deviation for non-Aboriginal students of 0.210 logits and a standard deviation for Aboriginal Torres Strait Islander students of 0.189 logits. There is no apparent reason for a deviation from the general trend on this item. It may be that the Aboriginal Torres Strait Islander students preferred using their instruments or sound sources to completing written responses. There was no writing related to this item, so they did not need to display literacy skills in their responses.

The item in which Year 7 non-Aboriginal students most significantly outperformed Aboriginal Torres Strait Islander students was item MA17. The mean score for non-Aboriginal students on this item was 0.287 logits, with a standard deviation of 0.235 logits, and the mean score for Aboriginal Torres Strait Islander students was 0.149 logits with a standard deviation of 0.181 logits. The difference in mean between the two groups was 0.138 logits, which, even allowing for error of measurement, is significant. This item

required an extended response and students were asked to explain what they heard in the piece *Ballet for Children* (Bliss, 1995) to justify their interpretation of the mood in the previous item. In the previous item, which was not marked right or wrong, they chose a mood from the selection: happiness, sadness, anger, and excitement. Among the possible reasons for Aboriginal Torres Strait Islander students' poor performance on this item could be the fact that this piece of music generated none of these moods for them. Another possible reason is the fact that it was an extended answer format and poor literacy skills may have affected results.

Year 7 sub-group analysis – English Speaking Background students and non-English Speaking Background students

At the Year 7 level, English speaking background students outperformed non-English speaking background students by a small margin in performance on the Analysis test and Process test combined. The mean score, on the arbitrary scale of 0-800, for Year 7 English speaking background students was 363.2, with a standard deviation of 78.4, and the mean score for non-English speaking background students was 350.4, with a standard deviation of 83.9. However, a two tail t-test (probability of 0.109) indicates that the difference in performances between the two groups is not statistically significant. Results indicate that 20.1 per cent of English speaking background students performed at Level 4 or above, compared with 22.8 per cent of non-English speaking background students. There were 58.8 per cent of English speaking background students who performed at Level 3 or above, compared with 50.9 per cent of non-English speaking background students.

Among the lower achievers, 41.2 per cent of English speaking background students performed at below Level 3, whereas 49.1 per cent of non-English speaking background students performed below Level 3 (see table 7.10). A graphical display of the frequencies of score ranges for Year 7 English speaking background and non-English speaking background students is given in Appendix xxiii.

Table 7.10: Percentages of Year 7 ESB and NESB students attaining SOS levels

SOS level	Score Ranges	Percentage	
		ESB students	NESB students
Level 1	0 - 220	3.4%	5.3%
Level 2	221 - 350	37.8%	43.8%
Level 3	351 - 420	38.7%	28.1%
Level 4	421 - 515	17.8%	22.8%
Level 5 & above	516 +	2.3%	nil

Differential item performances of Year 7 English speaking background and non-English speaking background students

Although Year 7 English speaking background students outperformed non-English background students on most items, the differences were small and could be due to measurement error. There were two items, however, that went against this trend, with non-English speaking background students outperforming English speaking background students. It is interesting that the first of these was item Analysis item MA17 which is the same item on which Year 7 Aboriginal Torres Strait Islander students scored significantly worse than the rest of the Year 7 population. English speaking background students scored a mean of 0.280 logits on this item, with a standard deviation of 0.235 logits, and non-English speaking background

students scored a mean of 0.285 logits with a standard deviation of 0.231. Students were asked to explain what they heard in the piece *Ballet for Children* (Bliss, 1995) to justify their interpretation of the mood they had selected in the previous item. It is possible that, unlike Aboriginal Torres Strait Islander students, non-English speaking background students were better at interpreting and describing mood than other students. Although the item was presented as an extended answer response format, the expected weaker literacy skills of non-English speaking background students did not adversely affect their performance on this item. It should be noted that the mean score for non-English speaking background students was only 0.005 logits better than English speaking background students, and error of measurement could account for the difference.

The second item on which non-English speaking background students outperformed English speaking background students was Process test item MP15. This item required students to critically analyse their group's composition and performance and make suggestions, in terms of musical elements, as to what improvements could be made. English speaking background students scored a mean of 0.338 logits, with a standard deviation of 0.179 logits and non-English speaking background students scored a mean of 0.366 logits, with a standard deviation of 0.169 logits. The difference in mean between the two groups was 0.028 logits. Examination of the item did not reveal an apparent reason for the better performance of non-English speaking background students, especially in view of the fact that an extended written response was required.

The following table provides a summary of sub-group performances in music in Year 7 (Pascoe, 1998, p.42).

Table 7.11: Summary of sub-group performances in music in Year 7

Sub-group	N	Mean	Standard deviation
All	921	359	82
Females	397	373	78
Males	487	350	80
ATSI students	44	309	77
NESB students	114	350	84

NOTE: All = the total Year 7 population tested
 Females = the total number of Year 7 girls tested
 Males = the total number of Year 7 boys tested
 ATSI = The total number of Year 7 Aboriginal Torres Strait Islander students tested
 NESB = The total number of Year 7 non-English speaking background students tested
 Mean = shown on the scale of 0 - 800

Year 10 sub-group analysis – males/females

At the Year 10 level, girls significantly outperformed boys in overall performance. The mean score, on the arbitrary scale of 0-800, for Year 10 girls was 544.4, with a standard deviation of 96.6, and the mean score for boys was 500.2, with a standard deviation of 102.7. A two tail t-test (probability of 0.000) indicates that the difference in performances between the two groups is significant. Results indicate that 34.9 per cent of girls performed at Level 6 or above, compared with 20.9 per cent of boys. There were 63.5 per cent of girls who performed at Level 5 or above, compared with 52.5 per cent of boys. Among the lower achievers, there were 25.9 per cent of boys who performed at below Level 4, whereas 10.5 per cent of girls performed below Level 4 (see table 7.12). A graphical display of the frequencies of score ranges for Year 10 girls and boys is given in Appendix xxiv.

Table 7.12: Percentages of Year 10 girls and boys attaining SOS levels

SOS level	Score Ranges	Percentage	Percentage
		Girls	Boys
Level 1	0 - 220	nil	nil
Level 2	221 - 350	3.5%	7.9%
Level 3	351 - 425	7.0%	18.0%
Level 4	425 - 515	26.1%	21.6%
Level 5	516 - 590	28.5%	31.6%
Level 6 & above	591 +	34.9%	20.9%

Differential item performances of Year 10 males and females

Although the differences did not exceed 0.197 logits, girls outperformed boys on every item at the Year 10 level and there were no items that did not fit this trend. There was one item, however, on which the difference in mean score was only 0.001 logits. This was Analysis item MA18 which had a relatively low level of difficulty of -2.892 logits, and which related to the piece *Dharpa* (Kellaway & Yunupingu, 1992). Students were required to identify the two different styles of music that had been combined for this piece. Although an open ended response was required, the expected poorer literacy skills of boys did not appear to adversely affect their performance on this item. A possible reason for their better performance may be that they enjoyed this 'rock' piece more than the classical and contemporary pieces presented, and so engaged more with this task.

Year 10 sub-group analysis – Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students

At the Year 10 level, non-Aboriginal Torres Strait Islander students significantly outperformed Aboriginal Torres Strait Islander students in performance on the Analysis and Process tests combined. The mean score, on the arbitrary scale of 0-800, for Year 10 non-Aboriginal Torres Strait Islander students was 533.3, with a standard deviation of 94.7, and the mean score for Aboriginal Torres Strait Islander students was 372.4, with a standard deviation of 87.5. A two tail t-test (probability of 0.000) indicates that the difference in performances between the two groups is significant. It should be noted at this point, however, that the number of Aboriginal Torres Strait Islander students who undertook the tests at Year 10 was only about 6.0 per cent of the total number of students tested. For this reason, these results should be treated with caution. The data indicates that 30.1 per cent of non-Aboriginal Torres Strait Islander students performed at Level 6 or above, whereas no Aboriginal Torres Strait Islander student performed above Level 4. There were 85.6 per cent of non-Aboriginal Torres Strait Islander students who performed at Level 4 or above, compared with 29.4 per cent of Aboriginal Torres Strait Islander students. Among the lower achievers, there were 29.4 per cent of Aboriginal Torres Strait Islander students who performed at below Level 3, whereas only 4.1 per cent of non-Aboriginal Torres Strait Islander students performed below Level 3 (see table 7.13). A graphical display of the frequencies of score ranges for Year 10 ATSI and non-ATSI students is given in Appendix xxv)

Table 7.13: Percentages of Year 10 ATSI and non-ATSI students attaining SOS levels

SOS level	Score Ranges	Percentage	
		ATSI students	non-ATSI students
Level 1	0 - 220	nil	nil
Level 2	221 - 350	29.4%	4.1%
Level 3	351 - 425	35.3%	10.3%
Level 4	426 - 515	35.3%	23.6%
Level 5	516 - 590	nil	31.9%
Level 6 & above	591 +	nil	30.1%

Differential item performances of Year 10 Aboriginal Torres Strait Islander students and non-Aboriginal Torres Strait Islander students

There are significant differences in performance on all items between non-Aboriginal Torres Strait Islander students and Aboriginal Torres Strait Islander students. The non-Aboriginal Torres Strait Islander students have outperformed the Aboriginal Torres Strait Islander students with differences in mean scores ranging up to 0.300 logits. There was only one item in which the difference in mean score was below 0.120 logits. This was Analysis item MA25, on which the difference in mean score was 0.086 logits. In this item, students were required to describe the mood of the landscape the composer had created in the excerpt from the contemporary piece, *Earth Cry Kakadu* (Sculthorpe, 1989). The piece provides a haunting, dramatic interpretation of the Australian landscape and it is possible the Aboriginal students engaged with this, and so performed better on this item. The item had a relatively high difficulty level of 0.948 logits.

Year 10 sub-group analysis – English speaking background students and non-English speaking background students

At the Year 10 level, English speaking background students outperformed non-English speaking background students by a small margin in overall performance. The mean score, on the arbitrary scale of 0-800, for Year 10 English speaking background students was 531.5, with a standard deviation of 94.3, and the mean score for non-English speaking background students was 498.9, with a standard deviation of 127.2. However, a two tail t-test (probability of 0.123) indicates that the difference in performances between the two groups is not statistically significant. Results indicate that 29.1 per cent of English speaking background students performed at Level 6 or above, compared with 26.8 per cent of non-English speaking background students. There were 62.1 per cent of English speaking background students who performed at Level 5 or above, compared with 46.3 per cent of non-English speaking background students. Among the lower achievers, 14.2 per cent of English speaking background students performed at below Level 4, whereas 31.7 per cent of non-English speaking background students performed below Level 4 (see table 7.14). A graphical display of the frequencies of score ranges for Year 10 ESB and NESB students is given in Appendix xxvi.

Table 7.14: Percentages of Year 10 ESB and NESB students attaining SOS levels

SOS level	Score Ranges	Percentage	
		ESB students	NESB students
Level 1	0 - 220	nil	nil
Level 2	221 - 350	3.4%	14.6%
Level 3	351 - 420	10.8%	17.1%
Level 4	421 - 515	23.7%	22.0%
Level 5	516 - 590	32.6%	19.5%
Level 6 & above	591 +	29.5%	26.8%

Differential item performances of Year 10 English speaking background students and non-English speaking background students

There were no extreme differences in performance between Year 10 English speaking and non-English speaking background students on any one item and there was no definite trend. The greatest difference in the mean score between English speaking and non-English speaking background students was for item MA18 in which English speaking background students outperformed non-English speaking background students by 0.128 logits. For most items the differences were less than 0.100 logits. There were six items on which non-English speaking background students outperformed English speaking background students. The differences between performance, however, were all less than 0.070 logits and it is possible they are attributable to error of measurement.

There were nineteen items on which English speaking background students outperformed non-English speaking background students. The greatest difference in performance was for Analysis item MA18. The mean

score for English speaking background students on this item was 0.900 logits, with a standard deviation of 0.249 logits, and the mean score for non-English speaking background students was 0.773 logits with a standard deviation of 0.349 logits. The difference in mean between the two groups was 0.128 logits. The item, which related to the piece *Dharpa* (Kellaway & Yunupingu, 1992), had a very low difficulty level of -2.892 . Students were asked to describe the different styles of music combined by the composer to create the piece. The two styles of music were 'rock' and 'traditional Aboriginal' music which English speaking background students may have recognised more readily than non-English speaking background students. The difference of 0.128 logits is not exceptionally high, however, and error of measurement may be a factor.

The following table provides a summary of mean sub-group performances in music in Year 10 (Pascoe, 1998, p.42). On average, females outperform males, males and females outperform Aboriginal Torres Strait Islander students, and non-English speaking background students outperform Aboriginal Torres Strait Islander students at Year 10 level.

Table 7.15: Summary of sub-group performances in music in Year 10

Subgroup	N	Mean	Standard deviation
All	324	525	105
Females	172	544	97
Males	139	500	103
ATSI	17	372	87
NESB	41	499	127

NOTE: All = the total Year 10 population tested
 Females = the total number of Year 10 girls tested
 Males = the total number of Year 10 boys tested
 ATSI = The total number of Year 10 Aboriginal Torres Strait Islander students tested
 NESB = The total number of Year 10 non-English speaking background students tested
 Mean = shown on the scale of 0 - 800

Summary

At Year 3, Year 7 and Year 10, girls scored significantly better than boys on the total performance of the Analysis and Process tests combined. Although the difference in total scores was statistically significant, differences in performance on individual items were not extreme, and in many cases, do not exceed the 0.08 logit error of measurement.

At the three Year levels, non-Aboriginal students significantly outperformed Aboriginal Torres Strait Islander students on the total scores of the Analysis and Process tests combined. The performance of non-Aboriginal students on individual items was also better than Aboriginal Torres Strait Islander students. In most cases, even allowing for error of measurement, these differences are statistically significant.

The differential performance of English speaking background and non-English speaking background students varied among the Year levels and did not appear to follow the predictable trends apparent in the male/female subgroup or the Aboriginal/non-Aboriginal Torres Strait Islander subgroup. At the Year 3 level, English speaking background students scored better than non-English speaking background students on the Analysis and Process tests combined. Although the difference was statistically significant, the differences on individual items were minimal and some could have been attributed to error of measurement. There were two items on which non-English speaking background students scored better than English speaking background

students and seventeen items on which English speaking background students performed better.

At Year 7, there was no statistically significant difference between English speaking and non-English speaking background students in the overall performance on both tests. Differences on individual items were minimal and, in most cases, did not exceed the 0.08 error of measurement. There were five items on which non-English speaking background students outperformed English speaking background students and nineteen items on which English speaking background students outperformed non-English speaking background students.

At Year 10 there was no statistically significant difference between English speaking and non-English speaking background students in their total performance on the Analysis and Process tests. There was no definite trend in performance on individual items and any differences were minimal, with most being less than the error of measurement. There were six items on which non-English speaking background students outperformed English speaking background students and nineteen items on which English speaking background students outperformed non-English speaking background students.

CHAPTER EIGHT

SUMMARY, CONCLUSIONS AND IMPLICATIONS

Summary of the study

The recognition of *The Arts* as one of the eight compulsory learning areas in the Western Australian K-10 curriculum, together with the demand for accountability in education, has revealed a lack of reliable and systematic methodology for evaluating progress in this learning area. Music is one of the five disciplines included in *The Arts* learning area in Western Australian school and the aims of this study were to address the problem of evaluating progress in music, by developing music assessment instruments that would show patterns of development in student achievement in music. This included the administration of the tests to a sample of 2 191 students in Western Australia and analysis of the data, so that teachers could be provided with a useful profile of achievement, upon which students' progress could be measured. It would also provide them with state means of student achievement that could be used as benchmarks for comparisons, as well as information related to the differential performance of sub-groups.

The study has addressed the problem of evaluating progress in music by developing an innovative range of assessment tasks appropriate for use at system, school or classroom level so that meaningful reporting of student outcomes in music can occur. A Rasch measurement model was used to

transform student raw scores into ability estimates. A scale was developed to allow teachers the opportunity to map students' skills and understandings onto a profile of achievement matched to a standards framework based on *The Arts Student Outcome Statements* (Education Department of Western Australia 1996). As well as measuring their students' achievements against an outcomes framework, teachers can compare them with the means of other students in the same Year levels across Western Australia.

Tasks were developed for students in Years 3, 7 and 10. The reason for selecting these three levels is that they represent three significant stages of students' compulsory schooling, that is, the conclusion of early childhood education, the conclusion of primary school and the conclusion of the compulsory years of schooling. Teachers can, however, use the assessment materials at other Year levels, thus providing valuable information on student progression through outcome levels and diagnostic information about strengths or weaknesses. This was made possible by the linking of items across the three tests through the use of common tasks and common stimulus materials. Obviously, if using the tests at Year levels other than 3, 7 and 10, it will not be possible for teachers to make comparisons with state means.

In order to evaluate student achievement in the range of skills and abilities contained in *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), two assessment instruments were developed. The first, the Analysis test, assesses student ability to understand and appreciate music. It consists of a set of stimulus material to which

students respond, primarily in relation to the strands, *Responding, Reflecting and Evaluating* and *Understanding the role of the Arts in Society*. These are known as the *Appreciating* strands. Students produce responses in relation to aesthetics, critical analysis, interpretation of meaning and music concepts, such as beat, rhythm, melody, dynamics, shape, mood and tension. Developmental processes involve comparisons and contrasts and the exploration of critical and contextual understanding focusing on particular periods of music history.

The second assessment instrument, the Process test, addresses student ability to 'make' music. It offers a broad view of student abilities through their documentation of the steps in the learning process, which lead to the performance of their final products. The process test addresses *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996) strands of *Creating, exploring and developing ideas* and *Using skills, techniques, technologies and processes*. These are known as the *Expressing* strands. This test provides evidence of students' planning processes towards a simple composition and performance. The activities in which students engage provide opportunity for inquiry and the use of music language, which are fundamental elements in the creative process leading to the development of a composition and its performance.

Conclusions from the study

Some arts educators were of the belief that, although students could be 'examined' on their ability to read music, remember pieces by rote, and play instruments, assessment of music in such things as critical thinking skills and aesthetics would not be possible. Because it involved the development of a marking key for the quantitative assessment of such things as imagination and creativity and the scoring of musical compositions and performance, they believed that achieving reliability in assessment would not be possible. This study demonstrates that, provided there is a clear developmental framework of achievement, together with marking keys that define and describe precisely what evidence is sought to demonstrate that achievement, then reliable assessment of music achievement at primary and secondary schools is possible. Double marking of tests would have provided more reliability in relation to the markers' use of the marking keys and the elimination of possible discrepancies in the scoring of items by individual markers. However, although significant marker-training was carried out, budgeting constraints did not allow for double marking.

Unlike tests in which a raw score is allocated, the use of a Rasch analysis provides scale estimates that place items in order of difficulty and students in order of ability on the same scale. This means, for instance, that a student who was able to respond to five difficult items would be placed higher on the scale than a student who responded to five easy items. Unlike a conventional classroom test (in which a raw score is allocated by adding up

the scores on individual items), a zero score in this testing process does not imply a complete lack of knowledge. It may simply mean that the test did not ask a question related to the student's knowledge. On the other hand, a perfect score does not imply that the student knows everything there is to know about music. It would be impossible to include everything there is to know about music in a test.

A reliable music achievement scale to measure student outcomes across a wide range of skills in classroom music has been successfully developed in this study. Validity of the measure has been established through the trialling of the materials with a large sample of Western Australian students and the analysis of data using a Rasch model of analysis. Overall fit to the model has been established and thresholds are properly ordered. Reliability of the scale, as shown by its Separability Index, is high, targeting of the items against student measures is satisfactory, and the power of the tests-of-fit to the model is excellent. This means a reliable scale of music achievement, that will assist in the measurement of student outcomes in music in Western Australian schools, has been achieved.

Results of the data analysis indicate that the mean level for each Year group shows a clear pattern of development from Year 3 through Year 7 to Year 10, although there is considerable overlap in performance between the Year groups. For example, the highest achieving 10 per cent of Year 7 students performed above the level demonstrated by approximately 25 per cent of Year 10 students.

Over 80 per cent of Year 3 students demonstrated skills associated with Level 2 outcomes in music. This means they can work in a group to plan and create a simple sound piece to interpret a given stimulus, including the creation of a simple score, notating their own rhythms, melodies and accompaniment patterns using simple known methods. They reflect upon music works, noting particular features including melody, instruments used, form and expression. They identify the purpose of a work and how it affects the way it should be performed. They apply simple critical reflections on their preferences and describe sounds using basic musical terms.

Over 55 per cent of Year 7 students demonstrated skills associated with Level 3 outcomes in music. This means they can compose short, simple, structured musical works using tuned or untuned percussion instruments, recorder, sounds from the environment, voice and body percussion. They are able to aurally recognise and describe musical features such as simple rhythmic and melodic patterns, tempo, instrumentation, timbre, dynamics and structure and use and interpret signs and symbols representing pitch, duration of sound and dynamics. They can describe obvious features such as repetition, form, changes in dynamics and texture, as well as identifying music from another culture and associating characteristics of the music with the style.

Over 80 per cent of Year 10 students demonstrated skills associated with Level 4 outcomes in music. This means they can create musical works

that capture characteristics of a given stimulus and interpret elements of pitch, rhythm, dynamics and phrasing in composition. They explore major and minor tonalities, textures, forms, media, and invent a soundscape score related to the theme. They explore combinations of sounds from the environment, chords, ostinati, and incorporate known structures such as ternary or binary form. They are able to give reasons why a musical element used in a piece is important and how it was used to create the perceived mood, tension and purpose. They can compare music from different times, places or cultures, identifying notable differences in musical characteristics.

These figures have provided baselines of achievement that can be used by teachers and administrators to make comparisons between the achievements of their students and those of the Western Australian population. The Rasch model of analysis used in this study also produced separate analyses for sub-groups. This provided the opportunity for comparisons to be made between the results of girls and boys, non-Aboriginal students and Aboriginal Torres Strait Islander students, and English speaking background and non-English speaking background students. At the three Year levels, non-Aboriginal students significantly outperformed Aboriginal Torres Strait Islander students on the total scores of the Analysis and Process tests combined. The performance of non-Aboriginal students on individual items was also better than Aboriginal Torres Strait Islander students. In most cases, even allowing for error of measurement, these differences are statistically significant.

The differential performance of English speaking background and non-English speaking background students varied among the Year levels and did not appear to follow the predictable trends apparent in the male/female subgroup or the Aboriginal/non-Aboriginal Torres Strait Islander subgroup. At the Year 3 level, English speaking background students scored better than non-English speaking background students on the Analysis and Process tests combined. Although the difference was statistically significant, the differences on individual items were minimal and some could have been attributed to error of measurement. There were two items on which non-English speaking background students scored better than English speaking background students and seventeen items on which English speaking background students performed better.

At Year 7, there was no statistically significant difference between English speaking and non-English speaking background students in the overall performance on both tests. Differences on individual items were minimal and, in most cases, did not exceed the 0.08 error of measurement. There were five items on which non-English speaking background students outperformed English speaking background students and nineteen items on which English speaking background students outperformed non-English speaking background students.

At Year 10 there was no statistically significant difference between English speaking and non-English speaking background students in their total performance on the Analysis and Process tests. There was no definite trend

in performance on individual items and any differences were minimal, with most being less than the error of measurement. There were six items on which non-English speaking background students outperformed English speaking background students and nineteen items on which English speaking background students outperformed non-English speaking background students.

Implications for teachers

This study is of importance to Western Australian teachers and schools, because, for the first time, specialist and generalist teachers have access to reliable, authentic assessment materials that reflect exemplary classroom practice. The tests will not only provide them with a useful set of instruments with which to measure student progress in music, but will also provide them with authentic models on which to base future classroom activities that can be assessed using an outcomes framework. The use of outcome levels to measure student progress is now compulsory in Western Australia and this study will contribute significantly to teacher knowledge in music education and in the use of an outcomes framework to measure student achievement.

The marking keys and student profiles have been developed so that teachers can administer the tests, mark them and use students' scores to establish levels of achievement and to compare their results with state means. They will not need to undertake any complicated analysis of the data to get this information. Outcomes that relate to aesthetics, critical analysis,

interpretation of meaning and music concepts can, for the first time, be reliably measured in Western Australian classrooms. This information can now be used to provide individual, class and school information for parents and school administrators as well as for school Management Information Systems necessary for system accountability.

Marking keys and item descriptions are worded in terms that can be understood by generalist as well as specialist teachers at both primary and secondary levels. This will provide access to this material to students in all schools, regardless of whether there is a music specialist or not. Although the tests were developed and trialled for use at Years 3, 7 and 10, they are not confined to these Year levels. The use of 'link' items and the developmental nature of the instruments allows for the mapping of student progress on a continuum of achievement so that teachers can use them to measure student performance against the outcomes framework at any Year level.

In order to evaluate student achievement in all strands of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996), it is necessary to assess students in the exploration and creation of music, as well as their appreciation of it. For this to occur, the data from both the Analysis and Process tests need to be combined to attain an overall profile of student performance.

The use of open-ended questions and partial credit in marking keys eliminates the notion that answers are 'right' or 'wrong' and allows the

opportunity for students to demonstrate their skills and knowledge to the maximum of their abilities. However, the use of open-ended questions that require extended responses needs to be approached with caution. Teachers need to be aware that they are testing music abilities. They are not testing student literacy skills. In order to obviate weaknesses in students' reading skills, teachers need to read questions aloud at lower levels and answer all questions related to the comprehension of questions at other levels. Spelling, grammar and sentence construction should not be factors in the marking of the tests.

Implications for school administrators

The gathering of data in music for whole school accountability purposes will be much improved using this method of assessment. The use of a quantitative measure in a learning area such as music has, in the past, been difficult, particularly for generalist teachers, and data collated for the school's Management Information System was based mainly upon the attitude of students within the classroom. It will now be possible for teachers to map the achievements of students in relation to outcome levels, thus providing a much more accurate measure of progress.

Reporting to parents using the method of assessment developed in this study will provide more information on their children's skills and abilities than in the past. Typical report comments for music in the past related to students' attitudes, rather than what they actually knew. For example, there were

comments such as “participates enthusiastically,” or “enjoys music.” Whilst attitude is an important part of the learning process, it does not provide information on the skills and abilities of students. Using the method of assessment developed in this study, teachers can actually provide the description of skills and abilities appropriate to the score the student has achieved which, for the first time, indicates the actual level of achievement of students. Improved reporting will also assist in emphasising the importance of music in the school curriculum. It is difficult to believe that parents would be satisfied with a comment such as “participates enthusiastically” in relation to maths or reading. The fact that they have accepted it for music in the past, indicates that they have not regarded it as being important. The need for accountability in all learning areas, including *The Arts*, as well as the impending compulsory use of the outcomes framework in schools means that administrators will, in the future, have to report on students’ skills and abilities in *The Arts*, including music.

Implications for centrally administered educational systems

Data indicate that, on these tests, boys perform at a lower level than girls, Aboriginal Torres Strait students perform at a lower level than non-Aboriginal students and non-English speaking background students may perform at a marginally lower level than English speaking background students. Teachers and administrators need to be aware that there are differences in the performances of sub-groups and that there could be a wide range of factors impinging upon their differential performances. It is likely that

the two factors most likely to cause differences in the performance of the sub-groups examined in this study are differences in literacy skills and issues of access and inclusivity. The issue of literacy skills impinges on all testing situations in which students are required to read and interpret questions and write their answers. The issue of access and inclusivity is one that needs to be examined and reviewed by administrators and teachers. In Western Australia where some of the world's most remote schools exist, the issue of accessibility through remoteness is significant. It will, in all probability, never be economically feasible to supply music teachers to remote areas of Western Australia. However, the rapid growth and development of the electronic media, and the use of computers in schools, will provide system-level administrators with an alternative solution to the problem. There is no reason why the learning of music cannot take place through these media, if serious consideration is given to the task.

The limited access of boys to music and the Arts generally, and the imbalance of girls and boys participating in the Arts, which was revealed during the drawing of the sample for this study (Pascoe, 1997a, 1997b, 1998), are issues which need to be addressed at a system level, as well as by Arts educators. Arts teachers need to assist in dispelling the outdated Australian culture's image of masculinity, which views artistic pursuits as 'sissy' (Pascoe, 1997b). If boys are to succeed in *The Arts*, they must feel comfortable in participating in performance and should not fear the ridicule of peers in doing so. The role models of successful males of stage, film and television should be used by teachers to promote *The Arts* to boys. The fact that boys have a

lower level of performance in literacy (Cook et al., 1995, p.63) is likely to have had an effect on the written sections of the music tests. The weaker literacy skills of boys is an ongoing issue for teachers and education systems and will continue to affect the results of written tests in all subject areas until the problem is addressed.

The literacy skills of both Aboriginal Torres Strait Islander students and non-English speaking background students is also likely to be one of the most significant factors contributing to the lower achievement of these groups. Western Australian system level testing data in reading and writing indicates that boys' literacy skills are weaker than those of girls, Aboriginal Torres Strait Islander students' literacy skills are weaker than those of non-Aboriginal students, and students from non-English speaking backgrounds have weaker literacy skills than those from English speaking backgrounds (Cook et al, 1995, p.63). These weaker literacy skills are likely to have an effect on testing in any subject area where students are required to read questions and write answers.

Cultural differences, in the case of both Aboriginal Torres Strait Islander students and non-English speaking background students, may have had an effect on results in the Process tests for three reasons. First, it may be culturally unnatural for some students to perform publicly. Second, a lack of language expertise could affect the confidence of some students to work within a group. Third, the interpretation of the painting which was presented as a stimulus may have been difficult for students from remote schools where

there is no art specialist to provide experience in the interpretation of paintings.

In remote areas of Western Australia where many Aboriginal Torres Strait Islander students attend school, access to specialised teaching is an issue. Where it is not economically feasible to install specialist music teachers in primary or secondary schools to serve a very small proportion of the population, Aboriginal Torres Strait Islander students may have been disadvantaged in music learning during all or part of their schooling. This may have had an effect on their results.

Serious consideration should be given to the implementation of strategies to ensure the delivery of arts learning to students in remote areas through the use of interactive multi-media. Programs that provide students with arts experiences that do not rely heavily on students' literacy skills could be developed for delivery by the Western Australian Education Department's School of Isolated and Distant Education. The use of sound and graphics, with the opportunity for students to interact, would provide a feasible alternative to the limited supply of arts teachers to remote areas. The Aboriginality of students in various regions should be taken into account in the development of these materials, so that the developed materials are suited to students' cultural needs.

Curriculum developers for *The Arts* learning area should consider the development of support materials for schools to assist in encouraging boys to

select arts courses in secondary school. The provision of stimulus materials and programs, that reflect the contemporary, more vigorous and aggressive styles in performing arts, could assist teachers and school administrators to entice boys into these disciplines.

Implications for theory

Future developers of music testing materials for use in Australian classrooms will need to be cognizant of the issues mentioned above. They will need to be taken into account in the development of tests for the assessment of the performance of music, as well as tests that require the written responses of students in the appreciation of music.

Future developers of tests should benefit from the success of the use of an outcomes framework and the use of a common scale as demonstrated in this study. The transforming of students' raw scores onto a scale in which the unit of measure is constant and the estimating of the difficulty of the items on the same scale has provided a reliable measure of student achievement in terms of outcome levels. This is a significant improvement on previous methods of allocating raw scores which, apart from ranking students, produced no information about what students could actually 'do'; that is, what their skills and abilities were in relation to any standard or framework. It was possible for one teacher to set a test on which most students achieved 80% or more, and another teacher, at the same school to set a test on which no student scored over 50%. There was no benchmark or framework to say the

first group necessarily had more knowledge than the second. It could possibly have meant the first test was easier than the second. The method provided in this study gives teachers the opportunity, not only to rank students, but to obtain a descriptive profile of what they know and can actually 'do' at a particular level of *The Arts Student Outcome Statements* (Education Department of Western Australia, 1996). It also provides the opportunity to make comparisons with state means of achievement in music.

In the past, many educators avoided assessment of aesthetics in music, in the knowledge that a student's personal opinion or interpretation of a piece of music should not be considered 'right' or 'wrong'. This issue has been addressed in this study by using questions related to aesthetics as prompts that require further justification by students in terms of musical knowledge. No mark is awarded for the prompt questions, thus allowing teachers, for the first time, to negotiate this previously unexplored territory without dampening students' enthusiasm to voice opinions and produce their own interpretations.

The use of partial credit to mark responses has provided a significantly improved method of assessment for teachers. The notion that a student is 'wrong' because she/he has not provided as much information as another student has been obviated using this method. It is possible to ask the same question of any student from Year 3, 7 or 10 and give them credit for the level of knowledge they display in their response. This also provides the opportunity to map students' progression along the outcomes continuum over time.

The methods of assessment in music developed in this study should be regarded by education systems and test developers as only a beginning. Now that baseline data have been gathered and new methods pioneered, the way has been paved for future, improved methods of assessment in *the Arts*, and music in particular. The confines of testing in this study did not allow for assessment in either group or individual singing, nor did it allow the opportunity for individual music reading or playing of instruments. Accurate measurement in these skills is possible through the use of partial credit marking keys, the outcomes framework and a common scale of measure, as used in this study.

To build on the theoretical model, the next project in music testing could include the testing of a sub-set of the sample in some individual vocal or instrumental activities and analysis using a Rasch model. The model could also be improved with the inclusion of activities to assess students' group singing skills, as group singing plays a significant part in classroom music, particularly at the primary school level. This study did not provide opportunity for assessment in this important area.

Now that a model has been provided for teachers in assessing the *Appreciation* strands of Arts outcomes, further assessment instruments should also be developed to establish and enhance the importance of these skills. The opportunity for group interaction, that is, the opportunity for students to discuss the music of particular societies or different cultures and talk about

stylistic elements that locate it in a time, place or culture would authenticate the assessment procedure in addressing the outcomes framework.

LIST OF REFERENCES

- Adams, R. J. & Khoo, S. T. (1994). *Quest: The interactive test analysis system*. Camberwell, Victoria: ACER.
- Alexander, R. R. (1982). Participant observation, ethnography, and their use in educational evaluation: A review of selected works. *Studies in art education*, 24 (1) 63-69 cited in Wright, S. K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.
- Allan D. & McDonald, J. (1993). *Keeping student performance central, The New York assessment collection, Studies on Exhibitions 14*. New York : Brown University.
- Almy, M. & Genishi, C. (1979). *Ways of studying children. (rev. ed.)*. New York: Teachers College Press cited in Wright, S. K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.
- Andrich, D. (1978a). Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement*, 2 (2) 449-460.
- Andrich, D. (1978b). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Andrich, D. (1982a). An extension of the Rasch model for ratings providing for both location and dispersion parameters, *Psychometrika*, 47, 105-113.
- Andrich, D. (1982b). An index of person separation in latent trait theory, the traditional KR.20 index, and the Guttman scale response pattern. *Education Research and Perspectives*, 9 (1), 95-104.
- Andrich D. (1983). *Diagnosing and accounting for response sets provoked by items of a questionnaire*. (Unpublished paper) presented to the National Council for Measurement in Education Conference, Montreal. Cited in Fairhall, J.H. *Dogmatism and Aesthetic Judgment, a study of response to paintings* (unpublished thesis). University of Western Australia.
- Andrich, D. (1985a). A latent trait model for items with response dependencies: Implications for test construction and analysis. In S. Embretson (Ed.), *Test design: Contributions from psychology, education and psychometrics*. New York: Academic Press.
- Andrich, D. (1985b). An elaboration of Guttman scaling with Rasch models for measurement. In N. Brandon-Tuma (Ed.), *Sociological Methodology*, San Francisco: Jossey-Bass.

- Andrich, D. (1988a). A general form of Rasch's Extended Logistic Model for partial credit scoring. *Applied Measurement in Education*, 1 (4) 363-.
- Andrich, D. (1988b). *Rasch models for measurement*. Sage university paper on quantitative applications in the social sciences, series number 07/068. Newbury Park: Sage Publications.
- Andrich, D. (1989). Distinctions between assumptions and requirements in measurement in the social sciences. In Keats, J. A., Taft, R., Heath, R. A., and Lovibond, S. H. (Eds) *Mathematical and Theoretical Systems*, North Holland: Elsevier Science Publishers.
- Andrich, D. (1991). Rasch models for measurement, *Quantitative applications in the social sciences* 68. Beverley Hills: Sage Publications.
- Andrich, D. (1992). *On the function of fundamental measurement in the social sciences* (unpublished paper) for the International Conference on Social Science Methodology, Trento.
- Andrich, D. (1995). *Advanced Social and Educational Measurement. Course materials*. Perth: Murdoch University.
- Andrich, D. & Godfrey, J. R. (1979). Hierarchies in the skills of Davis' reading comprehension test form D: an emperical investigation using a latent trait model. *Reading Research Quarterly*, 14 (2) 182-200.
- Andrich, D., Luo, G. & Sheridan, B. (1996). *Rumm. A Windows-based item analysis program employing Rasch Unidimensional Measurement Models*. Perth: Murdoch University and Edith Cowan University.
- Andrich, D. & Styles, I. (1994). Psychometric Evidence of Intellectual Growth Spurts in Early Adolescence. *Journal of Early Adolescence*, 14 (3) 328-344.
- Armstrong, C.L. (1994). *Designing assessment in Art*. Reston: National Art Education Association.
- Bailey, M. (1993). *Theoretical models and the assessment of competency : A report for the School of Educational Psychology, Measurement and Technology*. Sydney: University of Sydney.
- Baker, E. L. (1994). Making performance assessment work: The road ahead. *Educational Leadership*, 51 (6) 58-62.
- Bannister, R. (1992). Difficult but sensitive: Participant obversation research in music education. *British Journal of Music Education*, 9, 131-141.
- Bartel, L.R. (1994). Cognitive-affective response test - music: Questions of validity. *Bulletin*, 120, 1-15.

- Bates, D. (1992a). Opinions on the status of evaluation in music education in Canada. *Canadian Music Educator*, 33 (5) 15-17.
- Bates, D. (1992b). The Queen's symposium on evaluation in music education. *Canadian Music Educator*, 33 (5) 3-5.
- Beard, A. D. (1978). The quantification of visual aesthetic judgment using verbal and nonverbal scales (Doctoral dissertation, University of Chicago). *Dissertation Abstracts International*, 39, 3789A. cited in Myford, C.M. (1989). The nature of expertise in aesthetic judgment: Beyond inter-judge agreement. (Doctoral dissertation) University of Chicago.
- Beatty, R. (1992). Evaluation in music education: an Ontario perspective. *Canadian Music Educator*, 33 (5), 35.
- Bentley, A. (1966). *Aural foundations of music reading*. London: Novella & Company.
- Bliss, A. (1995). *Bliss Conducts Bliss. A Colour Symphony Things to Come, suite introduction & allegro*. [CD Rom] London Symphony Orchestra. Middlesex: Dutton Laboratories.
- Board of Studies. (1994). *Curriculum and Standards Framework: Draft for Consultation*. Victoria: Australian Government Publishing Service.
- Bonser, S. & Grundy, S. (1995). *Trialing student outcome statements: An Arts Faculty focus*. Perth: Quality Development and Educational Services.
- Boston, B. O. (1994). Perspectives on Implementation. *Arts Education Standards for America's Students*. Reston: Music Educators National Conference.
- Boughton, D. (1995). *Assessing learning effects in the visual arts: What criteria should be used?* Presented at the Art and Fact Conference, Rotterdam (unpublished paper, University of South Australia).
- Boughton D, Eisner E. W. & Ligtoet J. (1996). Evaluating and assessing the visual arts in education. *International Perspective*. New York: Teachers College Press.
- Bruner, J. S. (1960). *The process of education*. New York: Vintage Books, cited in Interim Curriculum Council of Western Australia (1997). *Draft Curriculum Framework*. Perth: Curriculum Council of Western Australia.
- Cancel, L.R. (1992). Remarks by the President. In Mitchell R. *Measuring up to the challenge*. Atlanta: American Council for the Arts.
- Carlin, J. (1994). Evaluation and assessment tools in the arts: creative dance unit. *Canadian Music Educator*. 36, (2) 29-35.

- Carlton, M. (1987). *Music in education*. London: The Woburn Press.
- Cartwright, P. (1989). Assessment and examination in arts education: Teachers talking. In M Ross (Ed.). *The claims of feeling: Readings in aesthetic education*. London: Falmer. Cited in Wright, S.K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.
- Centre for Educational Studies. (1987). *Task Group on Assessment and Testing, A Report*. (Unpublished report, University of London).
- Cohen, D. (1989). Tests fail to pass their exams. *Counterpoints No.1*. Rozelle: Australian Educational Network.
- Consortium of National Arts Education Associations. (1994). *National Standards for Arts Education*. Reston: Music Educators National Conference.
- Cook, J., Randall, K. & Richards, L. (1995). *Student Achievement in English*. Perth: Education Department of Western Australia.
- Curriculum Corporation. (1994a). *A statement on the arts for Australian Schools*. Melbourne: A.E. Keating (Printing) Pty. Ltd.
- Curriculum Corporation. (1994b). *The Arts - a curriculum profile for Australian schools*. Melbourne: A.E. Keating (Printing) Pty Ltd.
- Curriculum Programmes, Ministry of Education Western Australia. (1980). *Music in Schools*. Perth: Government Printer.
- Cutietta, R. (1985). An analysis of musical hypotheses created by the 11-16 year-old learner. *Bulletin*, 84, 1-11.
- Darling-Hammond, L. (1994). Performance-Based Assessment and Educational Equity. *Harvard Educational Review*, 64 (1) 5-30.
- Davis S. & Nunan, K. (1992, September). Musing on the Arts. *The Gen*, 1-4.
- Department of Education and Science. (1989). *National Curriculum. From policy to practice*. Stanmore: Publications Despatch Centre.
- Down, A. G. (1994). The view from the National Committee. *The vision for Arts education in the 21st century*. Reston: Music Educators National Conference.
- Down, A. G. & Mahlmann, J. J. (1994). Introduction. In *The Vision for Arts Education in the 21st Century*. Reston: Music Educators National Conference.

Duke, L. L. (1994). Potential of the standards for Arts literacy. In *The Vision for Arts Education in the 21st Century*. Reston: Music Educators National Conference.

Dwyer, B. (1989). Testing times produce their own myths. *Counterpoints No.1*. Rozelle: Australian Education Network.

Dyson, A. H. (1988). Symbol makers, symbol weavers: How children link play, pictures and print. *Young Children*, 45 (4) 4-9 cited in Wright, S.K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.

Ebel, R. L. & Frisbie, D. A. (1986). *Essentials of Educational Measurement*. New Jersey: Prentice-Hall Inc.

Education Department of Western Australia. (1993). *SEA Assessment Policy: Art Years 11 and 12*. East Perth: Education Department of Western Australia.

Education Department of Western Australia. (1987a). *Practical and Creative Arts. Class music*. Perth: Curriculum Branch.

Education Department of Western Australia. (1987b). *The Unit Curriculum. Assessment and Grading Procedures*. Perth: Curriculum Branch.

Education Department of Western Australia. (1993). *Tertiary Entrance Examinations*. Perth: Education Department of Western Australia.

Education Department of Western Australia. (1994a). *Profiles of Student Achievement*. Perth: Education Department of Western Australia.

Education Department of Western Australia. (1994b). *Student Achievement in Health and Physical Education in Western Australian Government Schools*. Perth: Education Department of Western Australia.

Education Department of Western Australia. (1994c). *Student Outcome Statements. The Arts. Working edition*. Perth: Education Department of Western Australia.

Education Department of Western Australia. (1995). *Student achievement in English in Western Australian Government schools*. Perth: Monitoring Standards in Education.

Education Department of Western Australia. (1996). *Student Outcome Statements. The Arts. Draft version*. Perth: Education Department of Western Australia.

Education Department of Western Australia. (1998). *Student Outcome Statements. Technology and Enterprise*. Perth: Education Department of Western Australia.

- Education Research and Development Committee. (1980). *National Assessment of Educational Progress*. Canberra: Australian Government Publishing Service.
- Edwin & Gordon. (1989). *Advanced measures of music audiation*. Chicago: G.I.A. Publications.
- Eibell, B. A. (1993). Developing an exhibition of skills: A cooperative venture. *Studies on Exhibitions (No.11)*. New York: Coalition of Essential Schools, Brown University.
- Emery, L. (1994). Room to move in the Arts. *EQ Australia*, 1, 5-7.
- Engelhard, G.J. R. & Wilson, M. (Ed.). (1996). *Objective measurement: Theory into practice Volume 3*. New Jersey: Ablex Publishing Company.
- Fairhall, J. (1989). Dogmatism and aesthetic judgment. A study of response to paintings. (Unpublished PhD thesis) University of Western Australia.
- Farmer, B. (1984). *Springboards*. Melbourne: Thomas Nelson Australia.
- Fehrs-Rampolla, B. (1994). Opportunities in the classroom. *Perspectives on implementation*. Reston: Music Educators National Conference.
- Forster, M. & Masters G. (1996a). Portfolios: Assessment Resource Kit. *Ark*. Melbourne: The Australian Council for Educational Research Ltd.
- Forster, M. & Masters G. (1996b). What is performance assessment? *Ark*. Melbourne: The Australian Council for Educational Research Ltd.
- Fortney, P. M. (1992). The construction and validation of an instrument to measure attitudes of students in high school instrumental music programs. *Contributions to Music Education*, 19, 32-45.
- Frederiksen, J. R. & Collins, A. (1989). A systems approach to educational testing. *Education Researcher*, 18 (9) 27-32.
- Gardner, H. (1996). The Assessment of Student Learning in the Arts. In Boughton D., Eisner, E. W. & Ligtoet, J. (1996). *Evaluating and assessing the visual arts in education: International perspective*. New York: Teachers College Press.
- Gillies, M. (1986). Levels of standardization in tertiary music assessment. *International Journal of Music Education*, 7, 9-15.
- Glaser, R. and Nitko, A.J. (1971). Measurement in learning and instruction, in Thorndike (ed.), *Educational Measurement*. Washington D.C: American Council on Education cited in Griffin, P. & Nix, P. (1991). *Educational Assessment and Reporting, A New Approach*. Sydney: Harcourt, Brace, Jovanovich.

Glass, G. V. & Stanley, J. C. (1970). *Statistical Methods in Education and Psychology*. New Jersey: Prentice Hall.

Gordon, E. E. (1992). Is it only in academics that Americans are lagging? *American Music Teacher*, 42 (3), 24-83.

Griffin, P. (1991). *Resolutions; Monitoring School Achievements*. Geelong: Deakin University.

Griffin, P. & Nix, P. (1991). *Educational Assessment and Reporting, A New Approach*. Sydney: Harcourt, Brace, Jovanovich.

Gronlund, N. E. & Linn, R.L. (1985). *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Company.

Gurin, R. S. (1994). The Arts education standards. In *Perspectives on implementation*. Reston: Music Educators National Conference.

Hanley B. (1992). Assessment and evaluation in music education, reflections on British Columbia initiatives. *Canadian Music Educator*, 33 (5) 7-13.

Hanley, B. (1992). Student assessment in music education. *Canadian Music Educator*, 33 (5) 19-24.

Hanley, B. (1994). Assessment Exchange, Student self evaluation. *Canadian Music Educator*, 35 (5) 7-10.

Hewton, J. (1991). *Performance indicators in education: What can they tell us?* Brisbane: V. R. Ward, Government Printer.

Hewton, J. (1985). *Primary Music Evaluation*. Brisbane: Queensland Dept. of Education.

Holmes, R. (1994). Shaping the Future: Queensland. *Review of the Queensland School Curriculum*, 2, 235-274.

Honeyman, A. (1996). *Monitoring standards in education: a handbook of the operations, policies and procedures of the Education Department of Western Australia* (draft). Perth: Education Department of Western Australia.

Interim Curriculum Council of Western Australia (1996). *The Curriculum Framework for K-12 Education in Western Australia* (draft). Perth: Interim Curriculum Council of Western Australia.

Interim Curriculum Council of Western Australia (1997). *Curriculum Framework* (draft). Perth: Curriculum Council of Western Australia.

International Baccalaureate. (1995). *Art/Design Guidelines and Instructions for Visiting Examiners*. Cardiff: International Baccalaureate Examinations Office.

Jorgensen, E. R. (1994). Justifying music instruction in American public schools. *Bulletin, Council for Research in Music Education*, 120, 15-31.

Kellaway, S. & Yunupingu, M. (1992). *Tribal Voice*. [CD Rom] Sydney: Mushroom Records.

Kemp, A. E. & Freeman, S. W. (1988). New tasks for music in primary schools and teacher training. *International Journal of Music Education*, 11, 21-23.

Knight, S. (1992). Evaluation-Alpha and Omega. *Canadian Music Educator*, 33 (5) 25-32.

Kubiszyn, T & Borich, G. (1987). *Educational Testing and Measurement: Classroom application and practice*. Glenview: Scott, Foresman and Company.

Leask, J. (1989), *Upbeat*: Gosford: Ashton Scholastics.

Lehman, P. (1994a). Issues of Assessment. In *Perspectives on Implementation*, Reston: Music Educators National Conference.

Lehman, P. (1994b). Writing Standards for Music. In *The vision for arts education in the 21st century*. Reston: Music Educators National Conference.

Lehman, P. (1996). *Performance Standards for Music*. Reston: Music Educators National Conference.

LeMahieu, P. G. (1995). Portfolios in Large-scale Assessment: Difficult but not impossible. *Educational Measurement: Issues and Practice*. 14 (3) 11-28.

Lokan, J. & Ford, P. (1994). *Mapping State Testing Programs*. Presented at the National Industry Education Forum, Melbourne. (Unpublished paper) Australian Council for Educational Research.

Leonhard, C. (1990). The National Arts Education Research Center: Background. *Bulletin, Council for Research in Music Education*, 105, 1-21.

Linn, R. L., Baker, E. L. & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Education Researcher*, 20 (8) 15-21.

Long, R. (1992). The future of the Arts in education? Alternative strategies. *British Journal of Music Education*, 9, 3-17.

- McGuire, M. (1983). Creativity and arts education. *Curriculum and Research Bulletin*, 18 (2) 2-14.
- Madsen, C. K. (1990). Teacher intensity in relationship to music education. *Bulletin*, 104, 38-45.
- Marzano, R. J. (1994). Lessons from the field about outcome-based performance assessments. *Educational Leadership*, 51 (6) 44-50.
- Marzano, R. J., Pickering, D. & McTighe, J. (1993). *Assessing student outcomes*. Alexandria: ASCD Publications.
- Masters, G .N. & Forster, M. (1996a). Developmental assessment. *Ark*. Melbourne: Australian Council for Educational Research Ltd.
- Masters, G. N. & Forster, M. (1996b). Portfolios. *Ark*. Melbourne: Australian Council for Educational Research Ltd.
- Masters, G. N. & Forster, M. (1996c). Performances. *Ark*. Melbourne: Australian Council for Educational Research Ltd.
- Masters, G. N. (1991). *Assessing achievement in Australian schools*, commissioned by the Industry Education Forum, Melbourne. (Unpublished paper) Australian Council for Educational Research.
- Masters, G. N. (1992). In *Profiles of student achievement*. Perth: Education Department of Western Australia.
- Masters, G. N. (1993). Rasch measurement and IEA studies. *IEA Bulletin*, 2 (2) 1-2.
- Masters, G. N. (1994a). *Setting and Measuring Performance Standards for Student Achievement* prepared for Public Investment in School Education: Costs and Outcomes Conference. (Unpublished paper) Australian National University.
- Masters, G. N. (1994b). *Nationally Comparable Achievement Measures: Principles and Options*. Melbourne: Australian Council for Educational Research.
- Masters, G. N. & Wright, B. D. (1984). The essential process in a family of measurement models. *Psychometrika*, 49 (4) 529-544. In Engelhard, G.J. R. & Wilson, M. (Ed.) (1996). *Objective measurement: Theory into practice Volume 3*. New Jersey: Ablex Publishing Company.
- McGuire, M. (1983). Creativity and arts education. *Curriculum and Research Bulletin*, 18 (2) 2-14.
- Mercer, A. (1997). *Report on MSE Arts Analyses*. (Unpublished report) Educational Testing Centre, Murdoch University.

- Mercer, A. & Church, T (1998). *Assessment in the Arts* prepared for the Rasch International Measurement Conference. (Unpublished paper) University of Western Australia and Murdoch University.
- Ministry of Education and Training. (1993). *The Common Curriculum Grades 1-9*, Ontario: Ministry of Education and Training.
- Ministry of Education, British Columbia. (1994a). *Assessment Handbooks Series, Performance Assessment*. Victoria: Queens Printer for British Columbia.
- Ministry of Education, British Columbia. (1994b). *Assessment Handbooks Series, Portfolio Assessment*. Victoria: Queens Printer for British Columbia.
- Ministry of Education, British Columbia. (1994c). *Assessment Handbooks Series, Student Self-Assessment*. Victoria: Queens Printer for British Columbia.
- Ministry of Education, British Columbia. (1994d). *Assessment Handbooks Series, Student-Centred Conferences*. Victoria: Queens Printer for British Columbia.
- Ministry of Education. (1991). *School Accountability-Policy and Guidelines*. Perth: Ministry of Education.
- Ministry of Education. (1989a). *Frameworks: Unit curriculum art guide 8-10*. Perth: Programmes Branch, Ministry of Education.
- Ministry of Education (1989b). *Music in schools*. Perth: Curriculum programmes, Ministry of Education.
- Mitchell, R. (1992). *Measuring up to the challenge*. Atlanta: American Council for the Arts.
- Music Educators National Conference. (1994). *National Standards for Arts Education*. Reston: Music Educators National Conference.
- Music Educators National Conference. (1994). *The Vision for Arts Education in the 21st Century*. Reston: Music Educators National Conference.
- Music Educators National Conference Committee on Performance Standards. (1996). *Performance Standards for Music*. Reston: Music Educators National Conference.
- Myford, C. M. (1989). The nature of expertise in aesthetic judgment: Beyond inter-judge agreement. (Unpublished doctoral dissertation) University of Chicago.

- Myford, C. M. & Mislevy, R. J. (1994). *Monitoring and Improving a Portfolio Assessment System*. Princeton: Educational Testing Service.
- National Assessment of Educational Progress (1981). *Art and Young Americans, 1974-79: Results from the second national art assessment*. Denver: U.S. Department of Education.
- National Assessment of Educational Progress (1997). *The NAEP Report Card*. Washington D.C: U.S. Department of Education.
- Nierman, G E. (1985). The Development of a test of verbal descriptors in music. *Bulletin*, 84, 20-33.
- Niguidula, D. (1993). The digital portfolio: A richer picture of student performance. *Studies on Exhibitions* , 13. New York: Brown University, Coalition of Essential Schools.
- Ogilvie, L. (1992). Stage-struck! Assessment and class music making. *British Journal of Music Education*, 9, 201-209.
- Pascoe, B. J. (1997a). Monitoring standards in music education in Western Australian schools. *New sounds for a new century. ASME XI*. Brisbane: Australian Society for Music Education 30th Anniversary National Conference.
- Pascoe, B. J. (1997b). *Arts education: Who is missing out?* Unpublished paper. Women in Education Conference. Perth: Murdoch University.
- Pascoe, B. J. (1998). *Student Achievement in The Arts in Western Australian Schools*. Monitoring Standards in Education. Perth: Education Department of Western Australia.
- Pericles, L. (1978). *Heaven and Earth*, [Painting] Acrylic on board, 900x480mm. Perth.
- Polisi, J. W. (1994). Balancing the elements of arts literacy. *The vision for arts education in the 21st century*. Reston: Music Educators National Conference.
- Porter, B. (1992). Report from the Halifax County-Bedford District School Board, Nova Scotia. *Canadian Music Educator*, 33 (5), 37-39.
- Queensland Department of Education. (1985). *Assessment in music education*. Brisbane: Research Services Branch.
- Rasch G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research [Expanded edition] Chicago: University of Chicago Press.
- Rasch, G. (1980). Probabilistic models for intelligence and attainment tests [expanded edition]. Chicago: University of Chicago Press.

- Ratzlaff, L. A. (1987). *The Education Evaluator's Workbook*. Alexandria, Virginia: Capitol Publications.
- Ratzlaff, L. A. (1994). *The Education Evaluator's Handbook: How to assess education programs*. Virginia: Education Research Group, Capitol Publications, Inc.
- Reid, A. (1992). 'Accountability' and Education: The Great Profile Debate, Point and Counterpoint, *Curriculum Perspectives*, 12 (1) 55-66.
- Rendell, J. (1994). The art room arsenal, *EQ Australia*, 1, 16-17.
- Resnick, L. B. (1994). Performance puzzles. *American Journal of Education*, 102, 511-526.
- Roberts, B. (1994). Assessment in music education. A cross-Canada study. *Canadian Music Educator*, 35 (5) 3-5.
- Romer, R. (1994). In *The vision for arts education in the 21st century*. Reston: Music Educators National Conference.
- Ross, M., Radnor, H., Mitchell, S. & Berton, C. (1993). *Assessing Achievement in The Arts*. Buckingham: Open University Press.
- Salvia, J. & Ysseldyke, J. E. (1977). *Assessment: Assessment in special and remedial education* (4th ed.). Boston: Houghton-Mifflin, cited in Wright, S.K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.
- Sculthorpe, P. (1989). *Earthcry kakadu, mangrove*. ABC Classics Recording. [CD] Sydney Symphony Orchestra. Sydney: Australian Broadcasting Corporation.
- Secondary Education Authority (1996). *Secondary education statistics (Year 11 and Year 12). Course statistics*. Perth: Secondary Education Authority.
- Shepard, L. A. (1995). Using assessment to improve learning. *Educational Leadership*, 52 (5) 38-43.
- Shrubb, G. (1989). The basic skills tests in New South Wales, *Counterpoints No.1*. Rozelle: Australian Education Network.
- Spodek, B., Saracho, O. N. & David, M. D. (1987). *Foundations of early childhood education*. Englewood Cliffs, New Jersey: Prentice-Hall cited in Wright, S. K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.

Strauss, J. (1992). Foreward. In Mitchell, R. *Measuring up to the challenge*, Atlanta: American Council for the Arts.

Styles I. & Andrich D. (1993). Linking the standard and advanced forms of the Raven's progressive matrices in both the pencil-and-paper and computer-adaptive-testing formats. *Educational and Psychological Measurement*, 53 (4) 905-923.

Swanwick, K. (1984). Some observations on research and music education. *British Journal of Music Education*, 1 (3) 195-204 cited in Bannister, R. (1992). Difficult but sensitive: Participant observation research in music education. *British Journal of Music Education*, 9, 131-141.

The Council of Chief State School Officers. (1994). *Arts Education Assessment Framework*. Washington: The National Assessment Governing Board.

The Joint Committee on Standards for Educational Evaluation. (1980). *Standards for evaluation of educational programs, projects and materials*. New York: McGraw-Hill Book Company.

Vale, C. D. (1986). Linking item parameters onto a common scale. *Applied Psychological Measurement*, 10 (4) 333-344.

Warhurst, J. (1994). *Understanding cross curricular practice in schools: Occasional paper no. 5*. Belconnen: Australian Curriculum Studies Association Inc.

Washor, E. (1993). Show, don't tell: Video and accountability. *Studies on Exhibitions (No.10)*. New York: Brown University, Coalition of Essential Schools.

Waugh, R. F., Godfrey, J. R., Evans, E. D. & Craig, D. (1995). Measuring students' perceptions about cheating in six countries. *Australian Journal of Psychology* 47 (2) 73-80.

Waugh, R. F. (1998). The Course Experience Questionnaire: A Rasch Measurement Model Analysis. *Higher Education Research and Development*, 17(1) 45-64.

Webster, P. R. (1995). *Creativity, Composition and Computers: Connections for the New Century*. Hobart: Australian Society for Music Education 10th National Conference.

West Australian Newspapers. (1996). *The West Australian*. Perth: West Australian Newspapers Ltd.

Wiggins, G. (1993). Assessment: Authenticity, context, and validity. *Phi Delta Kappan*, 200-214.

- Willingham, L. (1992). Musical Growth: Need we evaluate it? *Canadian Music Educator*, 33 (5) 41-43.
- Willis S. (1994). *Student outcome statements: Fitting the pieces together* presented at Joondalup conference. (Unpublished paper) Murdoch University
- Willis, S. & Kissane, B. (1995). *Outcome-based education: A review of the literature* for Education Department of W.A. (Unpublished paper) Murdoch University.
- Withers, G. (1991). *From marks to profiles and 'records of achievement'*. Geelong: Deakin University Press.
- Worby, G. (1994). Multi-arts: a nest of heresies. *EQ Australia*, 1, 12-14.
- Worthen, B. R. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappan*, 74 (6) 444-454.
- Wright, B. D. (1985). Additivity in psychological measurement. In E. E. Roskam (Ed.), *Measurement and personality assessment*. Amsterdam: Elsevier Science Publishers B.V.
- Wright, B. D. (1985). Additivity in psychological measurement cited in Waugh, R. F., Godfrey, J. R., Evans, E. D. & Craig, D. (1995). Measuring students' perceptions about cheating in six countries. *Australian Journal of Psychology* 47 (2) 73-80.
- Wright, B. D. & Masters, G. N. (1982). *Rating Scale Analysis: Rasch Measurement*. Chicago: Mesa Press.
- Wright, B. D. & Stone, M. H. (1979). *Best Test Design*. Chicago: Mesa Press.
- Wright, B. & Masters, G. (1981). *The measurement of knowledge and attitude* (Research memorandum no. 30). Chicago: Statistical Laboratory, Department of Education, University of Chicago, cited in Waugh, R. F. *A Rasch measurement model analysis of the course experience questionnaire*. (unpublished paper) Edith Cowan University.
- Wright, S. (1987). *In search of creative art in young children*. Canberra: Australian Early Childhood Resource Booklet No. 4, 1-17 cited in Wright, S. K. (1994). Assessment in the Arts: Is it appropriate in the early childhood years? *Studies in Art Education. A Journal of Issues and Research*. 36 (1) 28-43.

Appreciating

Responding, reflecting and evaluating	Understanding the role of the Arts in society
<i>The students</i>	<i>The students</i>
APP Responds to sensory experiences: auditory, visual, tactile and/or kinaesthetic stimuli 1.F	APP Shows an awareness of auditory, visual, tactile and/or kinaesthetic stimuli in their immediate environments 2.F
APP Responds to arts works and activities in personal ways showing interest in the responses of others 1.1	APP Identifies arts experiences in their own lives 2.1
APP Outlines features of their own and others' arts works and activities using simple arts terminology relating their responses to these features 1.2	APP Outlines how the Arts are used for a range of different purposes in their everyday lives and familiar other cultures 2.2
APP Outlines key features of arts works and activities, giving reasons for their responses using appropriate arts terminology and critical processes 1.3	APP <ul style="list-style-type: none"> a Uses their understandings of the arts in their community and other cultures and times in making and sharing their own arts activities and arts works b Identifies the contribution of the arts and artists in their immediate community 2.3
APP <ul style="list-style-type: none"> a Makes critical observations about arts works and activities using given criteria b Recognises and accepts that different people have different points of view and personal responses 1.4	APP <ul style="list-style-type: none"> a Recognises similarities and differences and makes links between the Arts from different times and places b Recognises and understands the contributions the Arts and artists make to Australian society 2.4
APP Uses arts terminology and critical frameworks to analyse and express informed opinions about arts works and activities 1.5	APP <ul style="list-style-type: none"> a Identifies and discusses distinguishing features of arts works which locate them in a particular time, place or culture b Identifies and discusses the distinguishing features of arts works and activities in contemporary Australian society 2.5
APP Identifies, analyses and interprets features of arts works and activities expressing and discussing responses to them 1.6	APP <ul style="list-style-type: none"> a Shows an understanding of how the arts are shaped by particular historical, social, economic and political contexts and values and how these change over time b Identifies career opportunities in and related to the Arts 2.6
APP Uses processes of critical analysis to support interpretations and personal judgements about arts works and activities 1.7	APP Discusses the effect of continuity and change in local, national and international Arts 2.7
APP Critically reflects on meanings and values associated with particular arts works and activities 1.8	APP <ul style="list-style-type: none"> a Researches arts works from a variety of contexts understanding how histories are constructed in the arts and how their own expression and appreciation of the arts is shaped by them b Critically examines the ways the arts challenge and shape values and are influenced by prevailing values 2.8

Expressing

Creating, exploring and developing ideas

The student

EXP 1.1 Explores a variety of sensory experiences to make meaning

Using skills, techniques, technologies and processes

The student

EXP 2.1 Uses basic elements of auditory, visual, tactile and/or kinaesthetic experiences to communicate ideas

EXP 1.2 Uses play, sensory experiences and imagination as starting points for arts activities

EXP 2.2 Uses simple arts skills and processes in sharing their arts activities informally with others

EXP 1.3 Builds on their experiences in exploring, shaping and communicating ideas for personal creative expression

EXP 2.3 Uses the skills, technologies and simple arts processes that they have learned in making and sharing arts activities informally with others

EXP 1.4 Explores and develops ideas and feelings through arts activities using given contexts

EXP 2.4 Uses a range of skills, techniques, technologies and processes in communicating for an audience or purpose

EXP 1.5 Uses creative problem solving to explore and develop ideas for individual and group solutions to given tasks

EXP 2.5 Selects and manipulates a range of skills, techniques, technologies and processes to demonstrate and share solutions to given tasks

EXP 1.6 Uses creativity and originality to explore and develop ideas for individual and group solutions to given tasks drawing on links to the arts of different times and places

EXP 2.6 Chooses appropriate arts skills, techniques, technologies and processes to shape and share meaning through arts activities

EXP 1.7 Draws on a wide range of ideas, contexts and past and contemporary practice in arts activities to create and re-create arts works

EXP 2.7 Structures arts works by applying skills, techniques, technologies and processes to specific styles and forms

EXP 1.8 Aligns ideas with artistic purpose and articulates why and how their choices were made

EXP 2.8 Manages and refines skills, techniques, technologies and processes in the Arts to communicate for a specific audience or purpose

EXP 1.9 Selects and manipulates ideas, arts theory and practice to make arts works that show personal commitment and control of the art form

EXP 2.9 Integrates technical and structural elements to control the chosen medium using a range of skills, techniques technologies and processes

Major Outcomes

Students respond to a selection of local and international artworks, exhibitions and events using their own skills of understanding.

Students respond to the Arts in their own worlds and in the worlds of others. They recognise and discuss the ways the Arts both challenge and confirm prevailing values.

APP 1 Students at Foundation Level attend to Arts presentations, particularly those with a personal and familiar content. They respond to presentations by watching, participating, and applauding. They are able to express their likes and dislikes in simple terms. Students can distinguish between the five art forms: Dance, Drama, Media, Music, and Visual Art. They show interest in participating in arts activities. They develop preferences and can make choices about which art form they would prefer to participate in. They communicate their ideas through verbal and non-verbal communication sometimes using aids.

APP 2 Students make simple personal responses to their own arts works and activities and those of others showing engagement or enjoyment. They express their opinions directly but also show acceptance of the arts of others. Students recognise the arts in their own worlds and in their immediate communities.

APP 3 Students make critical responses by making subjective observations about features and elements of their own arts works and those of others. They use a variety of simple given frameworks in making their judgements. Students recognise a range of different purposes for The Arts such as entertainment, ceremonial and advertising.

APP 4 Students recognise and identify the important features of arts works and activities. Reflection is facilitated by their use of appropriate terminology and critical processes which enables them to articulate their reasons for their personal responses. Students identify the contributions of The Arts in communities. They consciously explore using distinctive features of known arts works from other cultures and times in their expressive activities.

APP 5 Students respond to and reflect on their own and others arts works using a given set of criteria such as structured questioning. They recognise and accept others' views and opinions. They recognise similarities and differences in the arts forms and arts works of different times, cultures and places. They show awareness and understanding of the role of the Arts and artists in Australian society.

APP 6 Students use formal critical processes such as analysis to express informed opinions about arts works. They respect differences of opinions. They recognise, discuss, and use distinguishing features of arts works or experiences that come from a particular place, time or culture as well as those that are identifiably Australian.

APP 7 Students use critical processes to describe, analyse, and interpret arts works. They give personal points of view and interpretations. They understand and discuss how arts works communicate ideas and both reinforce and challenge social, cultural and artistic values. They show an understanding of the importance of historical, social, economic and political contexts and analyse, describe and interpret arts works from these perspectives.

APP 8 Students use formal processes of analysis and interpretation to make judgments about their own and others' arts works and experiences. They discuss how the arts stay the same and how they change over time. They focus their understanding from three perspectives: local (WA), national (Australia), and international.

APP 9 In responding to the arts, students use a full range of critical processes to show their understanding of the meaning and value of arts works and experiences. In making and appraising the Arts, students use a range of past and contemporary contexts showing how their own understanding is shaped and influenced by them. In understanding the arts in a contemporary society, they recognise and discuss the ways the Arts both challenge and confirm prevailing values. They extrapolate criteria for evaluating a performance from a variety of professional critical reviews and apply these when reviewing a performance.

Major Outcomes

Students explore, develop, create, and communicate ideas through a range of arts forms as well as three domains: the music, visual arts and performing arts.

Students use skills, techniques and technologies to create, develop and communicate their arts works and activities in informal settings. They use critical processes to develop personal solutions in their arts work. They set goals and confidently work towards them. The processes involved are more personal as students work purposefully to the intended result.

EXP 1 Students at Foundation Level begin to explore elements of The Arts. They attend to and respond to stimuli in their environments and express themselves in a variety of ways through play. Students experiment and develop basic techniques in and across the arts forms. They learn simple routines and imitate actions modelled for them through structured activities.

EXP 2 Students value and use ideas and imagination as well as those of their play and sensory experiences as the basis for making and sharing arts activities. They work from the familiar and make connections to their arts activities. In doing so, students use simple skills to explore, create, develop and communicate their arts works and activities in informal settings. They use simple processes involving planning, expressing and reflecting.

EXP 3 Students explore their experiences and use skills and understandings they have developed about technologies and problem solving processes to shape arts ideas in the arts. Students make and share arts works informally.

EXP 4 Students have acquired a degree of technical skill and facility (technical control) which enables them to successfully re-create works and realise some of their own creative ideas. They work within the structure of a given task with a limited range of choices and a clear sense of the end purpose of their activity.

EXP 5 Students creatively solve problems individually and in groups selecting from a range of skills and techniques, manipulating them and accessing appropriate technologies to satisfactorily complete given tasks. They engage in arts activities and complete arts works that are structured and show clearly developed ideas.

EXP 6 Students explore their ideas using processes such as observation and research. They reflect on their experiences and The Arts of different times and places, using them where appropriate. Students confidently plan their work in the Arts. They use critical processes in selecting and using skills, techniques, and technologies to communicate meaning.

EXP 7 Students demonstrate an understanding of arts contexts and practices and use skills, techniques, technologies, and processes to structure their arts works in specific styles or forms. They select and develop arts ideas to create and re-create arts works through individual and/or group exploration. They use critical problem solving, critical and social processes to develop personal solutions in their arts work. They set goals and confidently work towards them. The processes involved are more personal as students work purposefully to the intended result.

EXP 8 Students make arts works deliberately choosing from a diversity of ideas. They use arts languages to communicate personal ideas and interpret the ideas of others effectively for specific audiences or purpose. Students appraise the consequences of their decisions and are able to re-organise work in progress. They have a repertoire of skills, techniques, technologies and processes that demonstrate the inter-relationship between technical competence and the expressive qualities of the Arts.

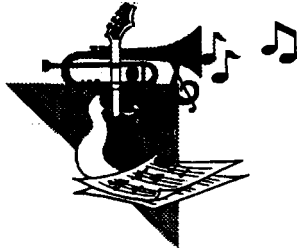
EXP 9 Students integrate their understanding of theory into their practice and manipulate ideas to make arts works. They integrate technical and aesthetic elements imaginatively, skilfully, and in a coherent personal style. Their arts works demonstrate independence, personal commitment, discipline, and control of the art form. They can step back from their arts works and see them with aesthetic distance.

PART 1

1. Where would you be most likely to hear this piece of music?
At a

- birthday party.
 orchestral concert.
 street parade.
 rock concert.

2. Explain what you heard in the music that made you pick this answer.

**PART 2**

3. If you moved to this part of the music you would

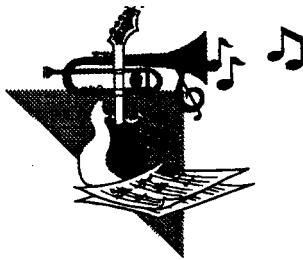
- march.
 skip.
 walk.
 run.

4. Explain what you heard in the music that made you pick this answer.

5. This music sounds

- sleepy.
- happy.
- sad.
- angry.

6. Explain what you heard in the music that made you pick this answer.



PART 3





7. Does the beat stay the same or does it change in this part of the music?

- same
- changes

8. Explain what you heard in the music that made you pick this answer.

PART 4

9. What shape does this section of music sound most like?

- 
- 
- 
- 

PART 5

10. What is the main instrument playing this part of the music?

- clarinet**
- flute**
- french horn**
- trumpet**

PART 6

11. Which of these rhythm patterns can you hear in this part?

- | | n | | | |
- | n | | | | n | z
- | n | | | | n | z
- | n | | | | n | z

**YEAR 3 MUSIC ANALYSIS TEST
MARKING KEY**

PART 1

1. Where would you be most likely to hear this piece of music? At a
a. birthday party b. orchestral concert c. street parade d. rock concert

1 mark b.

2. Explain what you heard in the music that made you pick this answer.

0 marks for:

- **irrelevant or incorrect response**,
eg it was loud, fast, different or nice, it sounded like people marching.
- **tautological response**, eg you would hear that kind of music at a concert.

1 mark for:

- **reference to the number of sounds**,
eg there are more sounds or there were many instruments.
- **personal experience**,
eg I have heard music like that at a concert.

2 marks for:

- **referring to specific instruments** - violins, trumpets etc or it was music that has a conductor.

PART 2

3. If you moved to this part of the music you would:
a. march b. skip c. walk d. run

1 mark for: b.

4. Explain what you heard in the music that made you pick this answer.

0 marks for:

- **tautological response**, eg I know it is.
- **incorrect response**,
eg it sounded like running.
- **response about movement**,
eg if I skipped that music would suit it.

1 mark for:

- **reference to music mood or sound**,
eg light and bouncy, light and happy.

2 marks for:

- **reference to music elements**,
eg because of the rhythm
 the beat is fairly fast.

5. This music sounds:

sleepy

happy

sad

angry

No score

6. Explain what you heard in the music that made you pick this answer.

0 marks for:

- **illogical or irrelevant response.**
- **tautological response**, eg it was happy.
- **personal reason**,
eg I liked it, it makes me feel good.

1 mark for:

- **reference to the sound of the music**,
eg the sound is fast and joyful.

2 marks for:

- **reference to musical elements or instruments**, eg the rhythm or beat, or the flutes have a happy sound. **or**
discussion of rhythm, instrumentation, texture, harmony.

PART 3

7. Does the beat stay the same or does it change in this part of the music?
same changes

1 mark for: **Changes.**

8. Explain what you heard in the music that made you pick this answer.

0 marks for:

- **tautological response**, eg because it changes.

1 mark for:

- **reference to movement**,
eg Skip — Walk
March, not marching.
- **reference to fast/slow without calling it beat**
or
- **reference to fast and slow beat**,
eg the beat is fast and changes to a slower beat.

2 marks for:

- **reference to rhythm change**,
eg Beat changes from
March — Waltz time
4/4 to 3/4.

PART 4

9. What shape does this section of music sound most like?

- A. B. C. D.

1 mark for:

- C.

PART 5

10. What is the main instrument playing this part of the music?

- a. clarinet b. flute c. French horn d. trumpet

1 mark for:

- D.

11. Which of these rhythm patterns can you hear in this part?

- A. B. C. D.

1 mark for:

- B.

PART 7

12. Explain how the music ends.

0 marks for:

- **incorrect or irrelevant response,**
eg like/dislike.

1 mark for:

- **simple general response,**
eg I like the way it ends because it's loud
It ends with a bang.

2 marks for:

- **explanation of how it is achieved,**
eg it gets louder, uses different instruments.

3 marks for:

- **above and mentions abrupt ending or identifies instruments and uses music-specific language,**
eg builds to crescendo then stops suddenly.

4 marks for:

- **above and discusses form, instrumentation/orchestration.**

13. a). Do you like this music?

- yes no

b). Explain what you heard in the music that made you pick this answer.

a). No score

b).

0 marks for:

- **irrelevant or tautological responses,**
eg I liked it or “nothing”.

1 mark for:

- **preference for this type of music,**
eg I like/dislike classical music.
- **reference to one musical element,**
eg I like music with a fast beat.

2 marks for:

- **reference to more than one musical element**
using everyday language.

3 marks for:

- **reference to more than one musical element**
using music-specific language or detailed
discussion of more than one element.

PART 1

1. Which group of instruments is playing this piece of music? Tick one group.

- brass band
- string quartet
- symphony orchestra
- concert band

2. Explain what you heard in the music that made you pick this answer.



PART 2

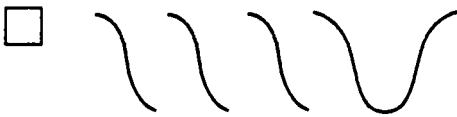
3. Does the beat stay the same or does it change in this part of the music?

- same
- changes

4. Explain what you heard in the music that made you pick this answer.

PART 3

5. What shape does this section of music sound most like?



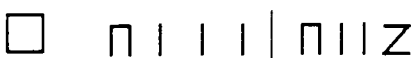
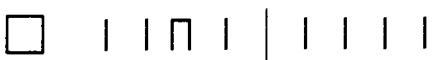
PART 4

6. What is the main instrument playing this part of the music?

- clarinet
- flute
- french horn
- trumpet

PART 5

7. Which of these rhythm patterns can you hear in this part ?

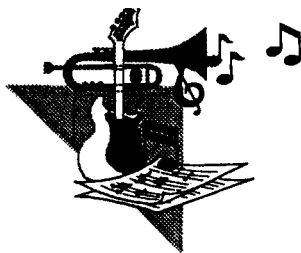


PART 6

8. What mood or feeling has the composer created in this piece of music?

- happiness
- sadness
- anger
- excitement

9. Explain what you heard in the music that helped to create the mood or feeling.



PART 7

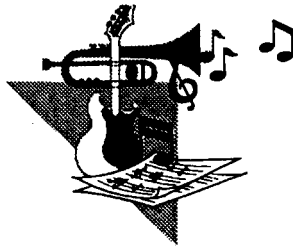
10. Explain how the music ends.

PART 8

11. In this piece of music the composer has combined different styles of music. What are they?

12. What is it in the music that helped you to decide?

13. What effect has this style of music had on Australian culture?



PART 9

14. Compare and contrast the two pieces of music you have heard, using the following headings:

a) **Instrumentation** (how the instruments are used)

b) **Expression** (dynamics, tempo, timbre)

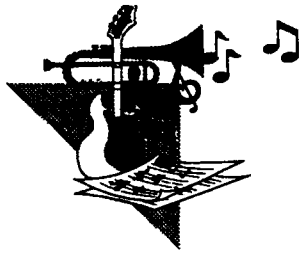
c) **Rhythm**

15. Which of these two pieces of music did you prefer?

first piece

second piece

Explain what you heard in the music that made you choose this piece.



PART 6

- | |
|---|
| 8. What mood or feeling has the composer created in this piece of music?
happiness sadness anger excitement |
|---|

No score

- | |
|--|
| 9. Explain what you heard in the music that helped to create the mood or feeling. |
|--|

0 marks for:

- **illogical or irrelevant response.**
- **tautological response**, eg it was happy.
- **personal reason**,
eg I liked it, it makes me feel good.

1 mark for:

- **reference to the sound of the music**,
eg the sound is fast and joyful.

2 marks for:

- **reference to musical elements or instruments**,
eg the rhythm or beat, or the flutes have a happy sound.

3 marks for:

- **discussion** of rhythm, instrumentation, texture, harmony.

PART 7

- | |
|--|
| 10. Explain how the music ends. |
|--|

0 marks for:

- **incorrect or irrelevant response.**
eg like/dislike.

1 mark for:

- **simple general response**,
eg I like the way it ends because it's loud.
It ends with a bang.

2 marks for:

- **explanation of how it is achieved**,
eg It gets louder, uses different instruments.

3 marks for:

- **above and mentions abrupt ending or identifies instruments and uses music-specific language**,
eg builds to crescendo then stops suddenly.

4 marks for:

- **above and discusses form, instrumentation/orchestration.**

PART 8

11. In this piece of music the composer has combined different styles of music. What are they?

1 mark each,
Rock music
Trad. Aboriginal.

12. What is it in the music that helped you to decide?

0 marks for:
• **irrelevant or tautological responses.**

1 mark for:
• **reference to instrumentation**, list of instruments/voice.

2 marks for:
• **reference to instrumentation and** reference to Aboriginal language and rhythm or beat.

3 marks for:
• **discussion of genre with details**, eg identification of beat and rhythms used, lack of melody etc.

13. What effect has this style of music had on Australian culture?

0 marks for:
• **irrelevant responses.**

1 mark for:
• **recognition of an effect**, eg positive effect.

2 marks for:
• **elaboration of the effect**, eg brings the Aboriginal and white cultures together, promotes understanding between 2 cultures.

3 marks for:
• **reference to sociological effects**, eg promotes tolerance and understanding, fusion of two cultures into a distinctively Australian style of music.

PART 9

14. Compare and contrast the two pieces of music you have heard, using the following headings:
- a) **Instrumentation** (how the instruments are used)
- b) **Expression** (dynamics, tempo, timbre)
- c) **Rhythm**

a) Instrumentation

0 marks for:

- **incorrect reference to instruments** or one piece only

1 mark for:

- **identification** of correct instruments in both pieces.
- **brief reference to style** used by instruments

2 marks for:

- **reference to the use** made of the instruments in both pieces, eg trumpets give a bouncy sound in the first piece (no contrast made).

3 marks for:

- **description and comparison** (or contrast) of the use made of the instruments in both pieces (specify instruments).

4 marks for:

- **comparison** including harmony, texture and orchestration **using music-specific language**.

b) Expression

0 marks for:

- **incorrect interpretation** of the term, eg reference to mood “sad”, “happy”.
- reference to one piece only

1 mark for:

- basic reference to **dynamics** or tempo for both pieces eg loud/soft in each piece.

2 marks for:

- description of **dynamics and tempo**, eg loud/soft, fast/slow in each piece.

3 marks for:

- above plus **comparison** of the way expression is achieved in each piece.

4 marks for:

- above plus discussion of **timbre, form, recurring themes and use of technical language**, eg crescendo/decrescendo.

c) Rhythm

0 marks for:

- **incorrect or irrelevant responses**.
- reference to one piece only

1 mark for:

- **basic reference to the type of rhythm**, eg light and bouncy in first piece etc.

2 marks for:

- **reference to the change/lack of change** of beat, eg from 4/4 to 3/4 in the first piece.

3 marks for:

- **reference to the effect of the changes** with some **comparison** of the two pieces.
May use **music-specific language**.

15. Which of these two pieces of music did you prefer?
first piece second piece

Explain what you heard in the music that made you choose this piece.

0 marks for:

- **irrelevant comments**,
eg I liked it or “nothing”.

1 mark for:

- **reference to the type** of music,
eg Like this type, don't like other type.
Need to mention name and type,
eg rock, classical, etc.
- brief reference to a musical element

2 marks for:

- reference to **one musical element** with some detail eg details about instrumentation.

3 marks for:

- reference to **more than one musical element** using everyday language.

4 marks for:

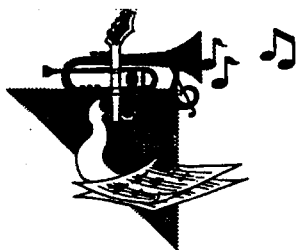
- reference to **more than one musical element with music-specific language**, including some discussion of the elements

PART 1

1. Which group of instruments is playing this piece of music? Tick one group.

- brass band
 string quartet
 symphony orchestra
 concert band

2. Explain what you heard in the music that made you pick this answer.



PART 2

3. Does the beat stay the same or does it change in this part of the music?

- same
 changes

4. Explain what you heard in the music that made you pick this answer.

PART 3

5. What shape does this section of music sound most like?



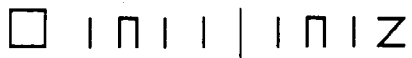
PART 4

6. What is the main instrument playing this part of the music?

- clarinet
- flute
- french horn
- trumpet

PART 5

7. Which of these rhythm patterns can you hear in this part?

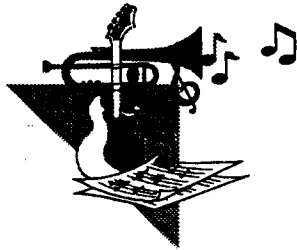


PART 6

8. What mood or feeling has the composer created in this piece of music?

- happiness
- sadness
- anger
- excitement

9. Explain what you heard in the music that helped to create the mood or feeling.



PART 7

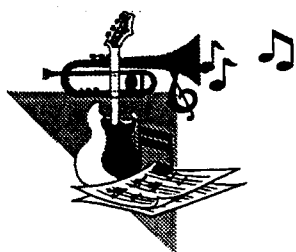
10. Explain how the music ends.

PART 8

11. In this piece of music the composer has combined different styles of music. What are they?

12. What is it in the music that helped you to decide?

13. What effect has this style of music had on Australian culture?



PART 9

14. Compare and contrast the two pieces of music you have heard, using the following headings:

a) **Instrumentation** (how the instruments are used)

b) **Expression** (dynamics, tempo, timbre)

c) **Rhythm**



PART 10

15. The composer of this music is an Australian who is describing the landscape. Describe the mood the composer has created with this piece of music.

16. What musical elements and effects has he used to achieve this?

YEAR 10 MUSIC ANALYSIS
MARKING KEY

PART 1

1. Which group of instruments is playing this piece of music?
Tick one group:
a. brass band b. string quartet c. symphony orchestra
d. concert band

1 mark for: c.

2. Explain what you heard in the music that made you pick this answer.

0 marks for:

- **incorrect or tautological responses**,
eg sounds like an orchestra.

1 mark for:

- **general reference** to instruments,
eg can hear strings and brass.
- **elimination** of other alternatives,
eg can't be ...

2 marks for:

- **reference to the variety** of the music.
- **identification of individual instruments**,
eg violins, flutes, trumpets.

3 marks for:

- **reference to families of instruments** as peculiar to an orchestra, eg violins and trumpets play in sections and then together.

PART 2

3. Does the beat stay the same or does it change in this part of the music?
same changes

1 mark for: Changes

4. Explain what you heard in the music that made you pick this answer.

0 marks for:

- **tautological response**, eg because it changes.

1 mark for:

- **reference to movement**,
eg Skip — Walk
March, not marching.

0 marks for:

- **illogical or irrelevant response.**
- **tautological response**, eg it was happy.
- **personal reason**,
eg I liked it, it makes me feel good.

1 mark for:

- **reference to the sound of the music**,
eg the sound is fast and joyful.

2 marks for:

- **reference to musical elements or instruments**, eg
the rhythm or beat, or the flutes have a happy sound.

3 marks for:

- **discussion** of rhythm, instrumentation, texture,
harmony.

PART 7

10. Explain how the music ends.

0 marks for:

- **incorrect or irrelevant response.**
eg like/dislike.

1 mark for:

- **simple general response**,
eg I like the way it ends because it's loud.
It ends with a bang.

2 marks for:

- **explanation of how it is achieved**,
eg It gets louder, uses different instruments.

3 marks for:

- **above and mentions abrupt ending or identifies instruments and uses music-specific language**,
eg builds to crescendo then stops suddenly.

4 marks for:

- **above and discusses form, instrumentation/orchestration.**

PART 8

11. In this piece of music the composer has combined different styles of music. What are they?

1 mark for

Rock music and Trad. Aboriginal.

12. What is it in the music that helped you to decide?

0 marks for:

- irrelevant or tautological responses.

1 mark for:

- reference to instrumentation only, list of instruments/voice

2 marks for:

- reference to instrumentation and reference to Aboriginal language and rhythm or beat.

3 marks for:

- discussion of genre with details, eg identification of beat and rhythms used, lack of melody etc.

13. What effect has this style of music had on Australian culture?

0 marks for:

- irrelevant responses.

1 mark for:

- recognition of an effect, eg positive effect.

2 marks for:

- elaboration of the effect, eg brings the Aboriginal and white cultures together, promotes understanding between 2 cultures.

3 marks for:

- reference to sociological effects, eg promotes tolerance and understanding, fusion of two cultures into a distinctively Australian style of music.

PART 9

14. Compare and contrast the two pieces of music you have heard, using the following headings:

- a) Instrumentation (how the instruments are used)
- b) Expression (dynamics, tempo, timbre)
- c) Rhythm

Instrumentation

0 marks for:

- **incorrect reference to instruments.**
- reference to one piece only

1 mark for:

- **identification** of correct instruments in both pieces.
- brief reference to style used by instruments

2 marks for:

- **reference to the use** made of the instruments in both pieces, eg trumpets give a bouncy sound in the first piece (no contrast made).

3 marks for:

- **description and comparison** (or contrast) of the use made of the instruments in both pieces (specify instruments).

4 marks for:

- **comparison** including harmony, texture and orchestration **using music-specific language.**

Expression

0 marks for:

- **incorrect interpretation** of the term, eg reference to mood “sad”, “happy”.
- reference to one piece only

1 mark for:

- basic reference to **dynamics** or tempo for both pieces
eg loud/soft in each piece.

2 marks for:

- description of **dynamics and tempo**,
eg loud/soft, fast/slow in each piece.

3 marks for:

- above plus **comparison** of the way expression is achieved in each piece.

4 marks for:

- above plus discussion of **timbre, form, recurring themes and use of technical language**, eg crescendo/decrescendo.

Rhythm

0 marks for:

- **incorrect or irrelevant responses.**
- reference to one piece only

1 mark for:

- basic reference to the **type of rhythm**,
eg light and bouncy in first piece etc.

2 marks for:

- **reference to the change/lack of change** of beat,
eg from 4/4 to 3/4 in the first piece.

3 marks for:

- **reference to the effect of the changes** with some **comparison** of the two pieces may use **music-specific language**.

PART 10

15. The composer of this music is an Australian who is describing the landscape. Describe the mood the composer has created with this piece of music.

0 marks for:

- **incorrect or irrelevant comments.**

1 mark for:

- **appropriate description of mood**,
eg gloomy, morbid; or
- **appropriate comments about landscape** without linking to mood.

2 marks for:

- **linking mood to landscape**,
eg lonely, desolate landscape.

3 marks for:

- **elaboration of mood** citing appropriate landscape, eg description of how the effect is achieved.

16. What musical elements and effects has he used to achieve this?

0 marks for:

- **irrelevant or incorrect responses.**

1 mark for:

- **reference to sounds or instruments not linked to mood.**

2 marks for:

- reference to sounds or instruments **linked to the mood** or images created.

3 marks for:

- above with **technical discussion** of how the mood is created, eg use of minor keys.

4 marks for:

- technical discussion referring to above plus **more overall discussion** as well as elements, eg form/structure.

17. Of the three pieces of music that you have heard today which one do you think is most effective in its use of musical elements?
first piece second piece third piece

Explain your reasons.

0 marks for:

- **irrelevant comments**,
eg I liked it or “nothing”.

1 mark for:

- reference to **the type** of music,
eg like this type, don't like other type.
Need to mention name and type,
eg rock, classical, etc.

2 marks for:

- reference to **one musical element**,
eg instrumentation.

3 marks for:

- reference to **more than one musical element** but
very little discussion.

4 marks for:

- reference to **more than one musical element**
with music-specific language, including some
discussion of the elements.

GENERAL INFORMATION

In 1996 the Western Australian Monitoring Standards program is assessing The Arts at the system level for the first time.

Assessment in The Arts is different from assessment in other learning areas. Accurate monitoring of performing arts must involve observation of students as they are actually involved in preparation and performance.

You will be administering the music process assessment task. Altogether the activities will take about 90 minutes. The assessment task may be administered over two sessions, ideally before and after a lunch break. However, it must be administered in one day. Your whole class will be involved in the following activities.

TIMING GUIDE

Session 1 outline (approx. 45 minutes)

- introduction and directed music warm-up (5 minutes)
- watch stimulus video (10 minutes)
class discussion
individual brainstorming
- group planning (15 minutes)
- group rehearsal (15 minutes)

BREAK

recess or lunch

Session 2 outline (approximately 45 minutes)

- final rehearsal (5 minutes)
- group performances for video (20 minutes)
- student critique of group performances (15 minutes)
- collection of all materials (5 minutes)

You will not be required to make formal assessment of your students' performance. All materials and video are to be returned to MSE for central marking. All planning sheets and critique sheets will be used as part of the assessment.

Prior to administering the assessment task:

- Become familiar with these administration guidelines
- Match student names with numbers on the organising sheet
- Identify nominated students in your class and organise students into small groups (see page 4 for guidelines on small group organisation)
- Organise suitable space, free of furniture and outside disturbances, for the assessment task such as a music room or empty classroom
- Provide as many sound sources as possible for students, such as traditional classroom music instruments, tin cans, bottles, plastic bags, rulers, small containers of rice/ beans/etc for shaking
- Have a video player and television set positioned for all class members to see clearly
- Organise a support teacher for videotaping of performances. It is the support teacher's responsibility to:
 - Mark out area for videotaping performances (see page 11 for guidelines on videotaping. You may detach this page of the booklet for the support teacher.)
 - Organise video camera

During the assessment task:

Please read aloud to the students all instructions enclosed in a box like this

- Help students having difficulty following the instructions but **do not help students with actual task**

After the assessment task:

- Collect all materials from students
- Check identification of group planning envelopes and individual booklets
- Return all materials, including all student booklets, planning sheets, videotapes, unused student booklets and administration guidelines to MSE in the return envelope provided

ADMINISTERING THE MUSIC ASSESSMENT

Time allowance: approximately 90 minutes

Small group organisation

It is essential to organise the groups prior to the assessment period, preferably the day before.

In your assessment package is an organising sheet. Match your alphabetical class list with these numbers.

Every fourth number has an asterisk*. Students matched to these asterisked numbers are students around whom groups of four are to be formed. Groups of three or five are also acceptable where there are odd numbers. Allow these students to choose their own groups. Intervene, however, if you believe performance will be adversely affected by student choice.

Teacher will need:

a blank videotape (provided by MSE)
a video camera
a video player and television set
stimulus video
sound sources
this booklet
manuscript paper (provided by MSE)
an instrument with which to make a signal sound during the warm-up (triangle, drum etc)

Each student will need:

a pencil or pen and eraser
a firm surface on which to write (a file or book)
student booklet

Each group will need:

a group envelope containing the Group Plan Sheet

SESSION 1

Introduction and warm-up (5 minutes)

When students are gathered quietly in the assessment area explain that you will be reading out the instructions because all classes doing these activities have to receive the same information.

Distribute the booklets and inform students that names are not needed on any of these materials.

Then say:

*Many students will be doing the task you are doing today.
The work you do today will help parents and other interested people see how well students in this state are progressing in music. It is very important that you do your best.*

You will need to listen very carefully so that you understand what to do.

First you are going to do a practice, making different sounds. Then you will be planning a composition using different sounds that you can find. You will have some time to think of ideas by yourself and then to plan and rehearse in a group before performing it for a video camera. After that you will have a short time to write down some comments about your group's performance. Each step will be explained as you proceed through the assessment activity. Any questions?

Answer any questions as necessary, then begin the warm-up. Use the outline on pages 9 and 10 of this booklet or use your preferred warm-up sequence which will help students focus on creative use of sounds and different music elements. It is important that such elements as expression, tempo, dynamics, pitch and harmony are included in the warm-up.

Class discussion and brainstorming (10 minutes)

Then say:

The sound composition you are going to create today will be based on ideas you have when you look at a short video.

Show the video. Hold a quick discussion to highlight what the students saw in the video. Try to emphasise that the video shows life before a storm; a storm approaching; the storm; storm effects; and after the storm. Awareness of these different aspects may help to give shape to the students' composition.

Then say:

Open up your assessment booklet and turn to the page with the title "Planning". Have your pen or pencil ready. You will look at the video, again then you are to write down ideas about different sounds you think could be made that might describe parts of the video. You can also draw quick pictures to describe these sounds if you prefer.

Show the video again, then say:

The top section of the planning sheet is where you can jot down ideas that come to you.

Now remember what you saw in the video. Imagine yourself in the video. What sounds can you hear? Jot down your ideas now.

It is important to record some ideas because ideas from this sheet will be used by you and your group to help make your composition. Also, markers will use this sheet to help understand and judge your work.

Allow a minute or so for brainstorming, then say:

Now, on the bottom part of the sheet, you are to write two sound ideas chosen from the top of the sheet, that you think will help describe what you saw on the video. Next to each idea write how you think you could make the sound. Be imaginative. Think about different ways of using your voice, musical instruments, body percussion, or scraping, shaking, hitting or blowing other objects like paper, plastic, a pencil case and so on. There are some musical instruments available if you need them. You only have a few minutes to do this before sharing your ideas with your group.

Group planning and rehearsal (30 minutes)

After a few minutes direct students to their groups of four (three or five in the case of odd numbers). Distribute one group envelope containing the Group Plan Sheet to each group.

Instruct groups to insert their group name (A, B, C, D etc) on:

- the group envelope
- the Group Plan Sheet
- their student booklet

Then say:

In your groups you are going to work out a composition of sounds that your group will perform. You will find it useful if each group member reads out ideas to the others. You will then have many ideas to begin with.

Before you start on your group work, turn to the page in your assessment booklet called "Ideas to help make your composition". Read them to yourself while I read them aloud.

Ideas to help you make your composition

Your composition should:

- *be about 1 minute long*
- *have a feeling or a mood*
- *be performed for an audience*

Your composition could also

- *have expression with variety of loud parts and soft parts*
- *have a beginning, a middle and an end*
- *have rhythm patterns*
- *have melody patterns*

While reading out this list, explain points where necessary and remind students that the list is there for them to look at while they are working on their composition.

Remember, it is the group's task to work out a composition that will describe in sound the feeling you had and what you saw in the video. You should spend some time jotting down your group's plan on the big planning sheet provided. This will help you when it comes time to perform your composition. Markers will be using what you record here to help understand and judge your work.

You will have time for rehearsal and trying out your composition once you have your ideas planned out. Are there any questions?

Deal with any questions or problems. Supervise groups to ensure that some written plan is made and that students then move on to rehearsal. If students wish to use traditional notation provide them with manuscript paper where necessary. Allow modification of written plan if development occurs during the rehearsal period. Students should record changes on their plan..

SESSION 2

Final rehearsal (5 minutes)

Remind students to look at the "Ideas....." page, to make sure their composition includes enough of the points required.

Videotaping group performances (20 minutes)

Work in collaboration with the support teacher to videotape group performances. It is very important that all students remain within the camera view finder area. **Also please make sure that instruments which make loud sounds are placed furthest from the camera (microphone), otherwise these sounds will dominate and distort sound reproduction.** The support teacher should identify groups clearly at the start of each performance by stating "This is Group..".

Note: It is extremely important for marking that groups are clearly identified.

Critique of performances (15 minutes)

Say:

Turn to the page in your assessment booklet with the title "Comments About your Composition ". The markers of this test want to know your ideas about your group's performances. Your answers to these questions will help them know this. I will read the questions to you while you read it yourself, then answer the questions.

COMMENTS ABOUT YOUR COMPOSITION

1. Tick one of the following which was important in your composition.

- | | |
|--|---|
| <input type="checkbox"/> melody | <input type="checkbox"/> rhythm |
| <input type="checkbox"/> expression
(louds and softs) | <input type="checkbox"/> instruments
(different types) |

Explain why it was important.

2.. *If you had more time to rehearse your composition, what is one thing you would do to improve it?*

Remember, it is only your group's performance that you are going to write comments on. These sheets will also be used to help judge your work.

Collection of materials (5 minutes)

- Instruct each group to place their student booklets, group plan sheet and manuscript (if used) into their group envelope.
- Collect group envelopes and place in order A, B, C, D, etc.
- Place in return envelope together with:
 - videotape and stimulus video
 - unused student booklets and administration guidelines
- Return to MSE for marking in the return envelope provided

Thanks to you and your class for your participation in this assessment task

IMPROVISATIONAL WARM-UP

The following suggested music warm-up includes activities which should focus children's thinking on creative use of sound and different musical elements. It is essential that children participate in such a warm-up before beginning to work out their composition. You may use this structure or one of your own. Whatever warm-up is used, children are to be given opportunity to briefly explore such musical elements as dynamics, tempo, melody, expression and harmony.

The following activity does not require instruments. You may vary it as you wish.

It is suggested that a small hand drum or triangle is used to make signal sounds between activities.

Spend only a short time on each activity. Remember that the whole warm-up is to take about 5 minutes.

Say:

What you are going to do today is make different sounds and play with those sounds to make a composition to describe a scene you will see on video.

Musical compositions can be made from many different sounds. You can use objects found in the room and be imaginative with them to make sound by scraping, hitting, blowing or shaking them. You can use our classroom musical instruments, use your voice, or make sounds with your hands or feet.

Before you start composing you are now to explore a few sounds made with the voice and body percussion.

Quickly find a place to stand by yourself and we will experiment with some of these sounds. Stop each time you hear the signal.

Make a sound with your foot by scraping it on the floor.

Signal

Now stamp your feet slowly, all together (conduct with a slow beat)

Signal

Now a little faster (gradually increase the tempo, keeping all children together).

Signal

Rub your hands together, slowly at first then gradually increase the pace.

Signal

Rub your hands on the floor, slowly then faster (only if smooth surface).

Divide class into three groups.

Group 1: stamp to regular marching beat

Group 2: clap hands to marching beat

Group 3: pat floor

Have 3 groups make sounds together to conducted beat.

Signal

Discuss, and take suggestions from class, how the three groups might sound different rhythms, or accent different beats while all performing together. then try out ideas.

Point out that this effect gives the sound 'thickness', but it has no melody.

Say:

As a group hum a soft low note

Signal

Now hum a middle range note..... increase the volume.....decrease the volume.

Group 1: Hum a low note

Group 2: Hum a middle range note

Group 3: Hum a high note

Conduct by changing dynamics (softs and louds)

Now return to the first exercise you did in three parts, but this time add the hum sound, one group at a time.

Signal

Say:

What other sounds can we make with the voice instrument?

Take 2 or 3 suggestions and try them. Combine them, varying the dynamics and tempo.

Say:

That is the end of the warm-up. Now you are to move on to planning your compositions.

VIDEOTAPING INSTRUCTIONS

SUPPORT TEACHER (SESSION 2)

- Choose position for camera. Video against a plain background without windows. Position camera so that light source is from in front, or at side, but **not** behind performance.
- Mark out suitable area for videotaping performances (approximately 3m x 2m).
- While the students are practising their dances at the beginning of this session please check the video camera to ensure that the equipment is correctly focused on the performing area.
- The camera should be in a fixed position to cover the performing area and no attempt should be made to pan or zoom.
- Groups should be videotaped in order A,B,C etc.
At the start of each group performance identify the group stating clearly into the microphone the letter name, eg. "This is Group A".
- Have camera running 5 - 10 seconds before group is identified.
- Videotape each performance and pause the camera between performances.
- Please assist the class teacher to facilitate the rapid changeover of groups and to maintain silence from the audience while each performance is in progress.
- Please check that all performances have been recorded, rewind the tape at the end of the session and ensure that it is included in the return package.
- If you have used a small video cassette attach it to the blank VHS tape provided. This is to enable identification.

Prior to administering the assessment task:

- Become familiar with these administration guidelines
- Match student names with numbers on the organising sheet
- Identify nominated students in your class and organise students into small groups (see page 4 for guidelines on small group organisation)
- Organise suitable space, free of furniture and outside disturbances, for the assessment task such as a music room or empty classroom
- Provide as many sound sources as possible for students, such as traditional classroom music instruments, tin cans, bottles, plastic bags, rulers, small containers of rice/ beans/ etc for shaking. **You may wish your students to use their own particular instrument, in which case they should be given notice in advance of the assessment day.**
- Organise a support teacher for videotaping of performances. It is the support teacher's responsibility to:
 - Mark out area for videotaping performances (see page 12 for guidelines on videotaping. You may detach this page for the support teacher.)
 - Organise video camera

During the assessment task:

Please read aloud to the students all instructions enclosed in a box like this

- Help students having difficulty following the instructions but **do not help students with actual task**

After the assessment task:

- Collect all materials from students
- Check identification of group planning envelopes and individual booklets
- Return all materials, including student booklets, planning sheets, videotape, stimulus pictures, unused student booklets and administration guidelines to MSE in the return envelope provided

ADMINISTERING THE MUSIC ASSESSMENT

Time allowance: approximately 110 minutes

Small group organisation

It is essential to organise the groups prior to the assessment period, preferably the day before.

In your assessment package is an organising sheet. Match your alphabetical class list with these numbers.

Every fourth number has an asterisk*. Students matched to these asterisked numbers are students around whom groups of four are to be formed. Groups of three or five are also acceptable where there are odd numbers. Allow these students to choose their own groups. Intervene, however, if you believe performance will be adversely affected by student choice.

Teacher will need:

a blank videotape (provided by MSE)

a video camera

sound sources

this booklet

manuscript paper (provided by MSE)

instrument with which to make a signal sound during the warm-up (triangle, drum etc)

Each student will need:

a pencil or pen and eraser

a firm surface on which to write (a file or book)

student booklet with colour print inserted

Each group will need:

a group envelope containing the Group Plan Sheet

SESSION 1

Introduction and warm-up (5 minutes)

When students are gathered quietly in the assessment area explain that you will be reading out the instructions because all classes doing these activities have to receive the same information.

Distribute the student booklets and inform students that names are not needed on any of these materials.

Then say:

Many students will be doing the task you are doing today. The work you do today will help parents and other interested people see how well students in this state are progressing in music. It is very important that you do your best.

You will need to listen very carefully so that you understand what to do.

First you are going to do a warm-up, then you will be planning a composition using different sounds. You will have some time to brainstorm ideas by yourself and then to plan and rehearse in a group before performing it for a video camera. After that you will have a short time to write down some comments about your group's performance. Each step will be explained as you proceed through the assessment activity.

Any questions?

Answer any questions as necessary, then begin the warm-up. Use the outline on pages 9 and 10 of this booklet or use your preferred warm-up sequence which will help students focus on creative use of sounds and different music elements. It is important that such elements as expression, tempo, dynamics, pitch and harmony are included in the warm-up

Class discussion and brainstorming (10 minutes)

Then say:

The sound composition you are going to create today will be based on ideas you have when you look at a particular picture included with your assessment booklet. Open your booklet, turn to the page headed "Planning". Take out the picture and have your pen or pencil ready.

Now look closely at the picture. Imagine yourself in the picture. What sounds can you hear?

The top section of the page is where you can brainstorm: jot down any ideas that come to you.

It is important to record some ideas because ideas from this sheet will be used by you and your group to help make your composition. Also, markers will use this sheet to help understand and judge your work.

It might be necessary to lead a brief discussion about the picture and possible ideas about sound. You could ask questions such as:

What time of day do you think it is in the picture? How many things can you see in the picture which could make sound? What feelings do you get when you look into the picture? What possible sounds could give those feelings?

Allow about 5 minutes for brainstorming, then say:

Now, on the bottom part of the sheet, you are to write three sound ideas, chosen from the top of the sheet, that you think will express what you see in the picture. Next to each idea, write how you think you could make the sound. Be imaginative. Imagine ways of making interesting sounds by using your voice, musical instruments, body percussion or scraping, shaking, hitting or blowing objects such as paper, plastic, a bunch of keys, a pencil case, and so on. You may be inventive with traditional

instruments as well as using them in traditional ways. There are some musical instruments available if you need them. You only have a few minutes to do this before sharing your ideas with your group.

Group planning and rehearsal (40 minutes)

After a few minutes direct students to their groups of four (three or five in the case of odd numbers). Distribute one group envelope containing the Group Plan Sheet to each group.

Instruct groups to insert their group name (A, B, C, D etc) on:

- the group envelope
- the Group Plan Sheet
- their student booklet

Then say:

In your groups you are going to work out a composition of sounds that your group will perform. Each group member is to read out ideas to the others. You will then have many ideas to begin with.

Before you start on your group work, turn to the page in your assessment booklet headed "Ideas to help make your composition" Read them to yourself while I read them aloud.

IDEAS TO HELP YOU MAKE YOUR COMPOSITION

Your composition should:

- *be between 1 and 2 minutes long*
- *express a feeling or a mood*
- *have a score using a form of notation*
you may use musical notation or you may write it down using your own words and symbols

Your composition could also:

- *have expression with variety of*
tempo (beat)
dynamics (loudness, softness, silence)
include timbre (tone quality - rough, sweet, etc) if students are familiar with this term
- *have rhythm,*
eg rhythmic patterns
body percussion
accents
include ostinato and syncopation if students are familiar with these terms

- *have texture (layers of sound) -*
- *have pitch,*
eg vocal sounds
instrumental sounds
melodic patterns
- *have form - write down your plan so that it will have a clear beginning, middle and end.*
- *be performed - remember you are to perform for an audience and video camera*

Remember, it is the group's task to work out a composition to express the picture. You should spend some time jotting down your group's plan on the big planning sheet provided. Markers will be using what you record here to help understand and judge your work. It is very important that your group works out a clear way to notate your composition and create a score.

You will have time to try out your composition and to rehearse it once you have your ideas planned out. Remember to continually refer to the Ideas page for help. Are there any questions?

Deal with any questions or problems. Supervise groups to ensure that some written plan or score is made and that students then move on to rehearsal. If students wish to use traditional notation provide them with manuscript paper where necessary. Allow modification of written plan if development occurs during the rehearsal period. Students should record changes on their plan.

SESSION 2

Final rehearsal (10 minutes)

Remind students to look at the "ideas....." page to make sure that their compositions include enough of the points required.

Videotaping group performances (25 minutes)

Work in collaboration with the support teacher to videotape group performances. It is very important that all students remain within the camera view finder area. **Also please make sure that instruments which make loud sounds are placed furthest from the camera (microphone), otherwise these sounds will dominate and distort sound reproduction.** The support teacher should identify groups clearly at the start of each performance by stating "This is Group..".

Note: It is extremely important for marking that groups are clearly identified.

Critique of performances (15 minutes)

Turn to the page in your assessment booklet headed "Comments About your Composition". The markers of this test want to know your ideas about your group's composition. I will read the questions to you while you read through them and then you can answer them.

COMMENTS ABOUT YOUR COMPOSITION

1. Tick one of the following which was important in your composition.

- | | | |
|---|----------------------------------|----------------------------------|
| <input type="checkbox"/> dynamics | <input type="checkbox"/> harmony | <input type="checkbox"/> texture |
| <input type="checkbox"/> instrument variety | <input type="checkbox"/> form | <input type="checkbox"/> tempo |
| <input type="checkbox"/> pitch/melody | <input type="checkbox"/> rhythm | |

Explain why it was important.

2. If you had more time to rehearse your composition, what is one thing you do to improve it? Be as specific as you can.

Remember, it is only your group's performance that you are going to write comments on. These sheets will also be used to help judge your work. You have about 15 minutes to complete this.

Collection of materials (5 minutes)

- Instruct each group to place their student booklets, group plan sheet and manuscript (if used) into their group envelope.
- Collect group envelopes and place in order A, B, C, D, etc.
- Place in return envelope together with:
 - videotape and stimulus pictures
 - unused student booklets and administration guidelines
- Return to MSE for marking in the return envelope provided

Thanks to you and your class for your participation in this assessment task.

IMPROVISATIONAL WARM-UP

The following suggested music warm-up includes activities which should focus children's thinking on creative use of sound and different musical elements. It is essential that children participate in such a warm-up before beginning to work out their composition. You may use this structure or one of your own. Whatever warm-up is used, children are to be given opportunity to briefly explore such musical elements as dynamics, tempo, melody, expression and harmony.

The following activity does not require instruments. You may vary it as you wish.

It is suggested that a small hand drum or triangle is used to make signal sounds between activities.

Spend only a short time on each activity. Remember that the whole warm-up is to take about 5 minutes.

Say:

What you are going to do today is make different sounds and play with those sounds to make a composition to describe a scene you will see on video.

Musical compositions can be made from many different sounds. You can use objects found in the room and be imaginative with them to make sound by scraping, hitting, blowing or shaking them. You can use our classroom musical instruments, use your voice, or make sounds with your hands or feet.

Before you start composing you are now to explore a few sounds made with the voice and body percussion.

Quickly find a place to stand by yourself and we will experiment with some of these sounds. Stop each time you hear the signal.

Make a sound with your foot by scraping it on the floor.

Signal

Now stamp your feet slowly, all together (conduct with a slow beat)

Signal

Now a little faster (gradually increase the tempo, keeping all children together).

Signal

Rub your hands together, slowly at first then gradually increase the pace.

Signal

Rub your hands on the floor, slowly then faster (only if smooth surface).

Divide class into three groups.

Group 1: stamp to regular marching beat

Group 2: clap hands to marching beat

Group 3: pat floor

Have 3 groups make sounds together to conducted beat.

Signal

Discuss, and take suggestions from class, how the three groups might sound different rhythms, or accent different beats while all performing together. then try out ideas.

Point out that this effect gives the sound 'thickness', but it has no melody.

Say:

As a group hum a soft low note

Signal

Now hum a middle range note..... increase the volume.....decrease the volume.

Group 1: Hum a low note

Group 2: Hum a middle range note

Group 3: Hum a high note

Conduct by changing dynamics (softs and louds)

Now return to the first exercise you did in three parts, but this time add the hum sound, one group at a time.

Signal

Say:

What other sounds can we make with the voice instrument?

Take 2 or 3 suggestions and try them. Combine them, varying the dynamics and tempo.

Say:

That is the end of the warm-up. Now you are to move on to planning your compositions.

VIDEOTAPE INSTRUCTIONS

SUPPORT TEACHER (SESSION 2)

- **Choose position for camera. Video against a plain background without windows. Position camera so that light source is from in front, or at side, but not behind performance.**
- **Mark out suitable area for videotaping performances (approximately 3m x 2m).**
- **While the students are practising their dances at the beginning of this session please check the video camera to ensure that the equipment is correctly focused on the performing area.**
- **The camera should be in a fixed position to cover the performing area and no attempt should be made to pan or zoom.**
- **Have camera running 5 - 10 seconds before group is identified.**
- **At the start of each group performance identify the group stating clearly into the microphone the letter name, eg. "This is Group A".**
- **Videotape each performance and pause the camera between performances.**
- **Rewind the tape at the end of the session.**

GENERAL INFORMATION

In 1996 the Western Australian Monitoring Standards program is assessing The Arts at the system level for the first time.

Assessment in The Arts is different from assessment in other learning areas. Accurate monitoring of performing arts must involve observation of students as they are actually involved in preparation and performance.

You will be administering the music process assessment task. Altogether the activities will take about 115 minutes and the assessment task should be administered during one session. Your whole class will be involved in the following activities.

TIMING GUIDE

Session outline
(approx. 115 minutes)

- introduction, class discussion and brainstorming (15 minutes)
- group planning (20 minutes)
- group rehearsal (35 minutes)
- group performances for video (25 minutes)
- student critique of group performances (15 minutes)
- collection of all materials (5 minutes)

You will not be required to make formal assessment of your students' performance. All materials and video are to be returned to MSE for central marking. All planning sheets and critique sheets will be used as part of the assessment.

Prior to administering the assessment task:

- Become familiar with these administration guidelines
- Match student names with numbers on the organising sheet
- Identify nominated students in your class and organise students into small groups (see page 4 for guidelines on small group organisation)
- Organise suitable space, free of furniture and outside disturbances, for the assessment task, such as a music room or empty classroom.
- Provide as many sound sources as possible for students, such as traditional classroom instruments, tin cans, bottles, plastic bags, rulers, small containers of rice/ beans/etc for shaking. **You may wish your students to use their own particular instrument, in which case they should be given notice in advance of the assessment day.**
- Organise a support teacher for videotaping of performances. It is the support teacher's responsibility to:
 - Mark out area for videotaping performances (see page 9 for guidelines on videotaping. You may detach this page for the support teacher.)
 - Organise video camera

During the assessment task:

Please read aloud to the students all instructions enclosed in a box like this

- Help students having difficulty following the instructions but **do not help students with actual task**

After the assessment task:

- Collect all materials from students
- Check identification of group planning envelopes and individual booklets
- Return all materials, including all student booklets, stimulus pictures, planning sheets, videotape, unused student booklets and administration guidelines to MSE in the return envelope provided

ADMINISTERING THE MUSIC ASSESSMENT

Time allowance: approximately 115 minutes

Small Group Organisation

It is essential to organise the groups prior to the assessment period, preferably the day before.

In your assessment package is an organising sheet. Match your alphabetical class list with these numbers.

Every fourth number has an asterisk*. Students matched to these asterisked numbers are students around whom groups of four are to be formed. Groups of three or five are also acceptable where there are odd numbers. Allow these students to choose their own groups. Intervene, however, if you believe performance will be adversely affected by student choice.

Teacher will need:

a blank videotape (provided by MSE)
a video camera and recorder
this booklet
sound sources
manuscript paper (provided by MSE)

Each student will need:

a pencil or pen and eraser
a firm surface on which to write (a file or book)
student booklet with colour print inserted

Each group will need:

a group envelope containing the Group Plan Sheet

Introduction, class discussion and brainstorming (15 minutes)

When students are gathered quietly in the assessment area explain that you will be reading out the instructions because all classes doing these activities have to receive the same information.

Distribute student booklets and inform students that names are not needed on any of these materials.

Say:

Many students will be doing the task you are doing today.

The work you do today will help parents and other interested people see how well students in this state can work with concepts in music. It is very important that you do your best.

First you will be asked to examine a copy of a painting. You will then brainstorm some of your ideas on your own before joining your group to plan, rehearse and perform a composition reflecting your ideas drawn from the painting. Your performance will be videotaped. After that you will have a short time to write comments about your group's performance. Each step will be explained as you proceed through the assessment activity.

Any questions?

Answer any questions as necessary then say:

The sound composition you are going to create today will be based on ideas you have when you look at a picture included with your assessment booklet. Open the booklet and turn to the page titled "Planning". Take out the picture. Have your pen or pencil ready.

Now look closely at the picture.Imagine yourself in the picture..... What sounds can you hear?

The top section of the planning page is where you are to brainstorm. Jot down any ideas that come to you.

It is important to record some ideas because ideas from this sheet will be used by you and your group to help make your composition. Also, markers will use this sheet to help understand and judge your work.

It might be necessary to lead a brief discussion about the picture and possible ideas about sound. You could ask questions such as:

What time of day do you think it is in the picture? How many things can you see in the picture which could make sound? What feelings do you get when you look into the picture? What possible sounds could give those feelings?

Allow about 5 minutes or so for brainstorming, then say:

Now, on the bottom part of the sheet, you are to write three sound ideas, chosen from the top of the sheet, that you think will express what you see or feel after looking at the picture. Next to each idea write how you think you could make the sound. Musical compositions can be made from many different sounds. Be imaginative. Imagine ways of making interesting sounds by using musical instruments, body percussion, or scraping, shaking, hitting or blowing objects such as paper, plastic, a bunch of keys, a pencil case, and so on. You should also remember that your voice is an instrument. You may also be inventive with traditional musical instruments as well as using them in usual ways. There are some musical instruments available if you need them. You only have a few minutes to do this before sharing your ideas with your group.

Group planning and rehearsal (55 minutes)

After a few minutes direct students to their groups of four (three or five in the case of odd numbers). Distribute one group envelope containing the Group Plan Sheet to each group.

Instruct groups to insert their group name (A, B, C, D etc) on:

- the group envelope
- the Group Plan Sheet
- their student booklet

Then say:

In your groups you are going to work out a composition of sounds that your group will perform. Each group member is to read out ideas to the others. You will then have many ideas to begin with.

Before you start on your group work, turn to the page in your assessment booklet headed "Ideas to help you make your composition". Read them to yourself while I read them aloud.

IDEAS TO HELP YOU MAKE YOUR COMPOSITION

Your composition should:

- *be between 1 and 2 minutes long*
- *express a feeling or a mood*
- *have a score using a form of notation, eg traditional or graphic, with performing instructions*

Your composition could also:

- *have expression with variety of*
 - tempo*
 - dynamics*
 - timbre*
- *have rhythm, eg*
 - rhythmic patterns*
 - ostinato,*
 - body percussion*
 - accents*
 - syncopation*
- *have harmony*
- *have texture*

Critique of performances (15 minutes)

Say:

Take your assessment booklet and turn to the page titled "Comments about your composition". You are now to write down your thoughts on your own group's performances.

You have fifteen minutes to complete this task.

Read through the critique questions with the students to check for understanding. Explain where necessary.

COMMENTS ABOUT YOUR COMPOSITION

1. Indicate with a tick which of the following were important in helping your group shape your composition.
- | | | |
|---|----------------------------------|----------------------------------|
| <input type="checkbox"/> dynamics | <input type="checkbox"/> harmony | <input type="checkbox"/> texture |
| <input type="checkbox"/> instrument variety | <input type="checkbox"/> form | <input type="checkbox"/> tempo |
| <input type="checkbox"/> rhythm | <input type="checkbox"/> melody | <input type="checkbox"/> timbre |

Choose two of these and write briefly how they contributed to your composition.

2. If you had more time to rehearse your composition, what would you do to improve the performance? Be as specific as you can.

Remember, it is only your group's performance that you are going to write comments on. These sheets will also be used to help judge your work.

Collection of materials (5 minutes)

- Instruct each group to place their student booklets, group plan sheet and manuscript (if used) into their group envelope.
- Collect group envelopes and place in order A, B, C, D, etc.
- Place in return envelope together with:
 - videotape and stimulus pictures
 - unused student booklets and administration guidelines
- Return to MSE for marking

Thanks to you and your class for your participation in this assessment task.

VIDEOTAPING INSTRUCTIONS

SUPPORT TEACHER (SESSION 2)

- Choose position for camera. Video against a plain background without windows. Position camera so that light source is from in front, or at side, but not behind performance.
- Mark out suitable area for videotaping performances (approximately 3m x 2m).
- While the students are practising their dances at the beginning of this session please check the video camera to ensure that the equipment is correctly focused on the performing area.
- The camera should be in a fixed position to cover the performing area and no attempt should be made to pan or zoom.
- Groups should be videotaped in order A, B, C etc. At the start of each group performance identify the group stating clearly into the microphone the letter name, eg. "This is Group A".
- Have camera running 5 - 10 seconds before group is identified.
- Videotape each performance and pause the camera between performances.
- Please assist the class teacher to facilitate the rapid changeover of groups and to maintain silence from the audience while each performance is in progress.
- Please check that all performances have been recorded, rewind the tape at the end of the session and ensure that it is included in the return package.
- If you have used a small video cassette attach it to the blank VHS tape provided. This is to enable identification.



GROUP ASSESSMENT: The same group marks are given to each individual in a group.

Arts language

(This is an overall, on-balance judgement of the group's ability to communicate through music)

1. How effectively did the group communicate their composition/message in the language of music?

0 marks	for no evidence , i.e. no communication of ideas, eg meaningless or isolated, non-musical sound effects
1 mark	for beginning to develop - attempts to convey their message and reflect message through soundscape
2 marks	for sound development - evidence of attempt to create mood and/or expression to communicate a message
3 marks	for well developed - message is clearly conveyed using sensory experiences
4 marks	for highly developed - a highly engaging, meaningful performance, with depth

(It may be helpful in marking this criterion to use sound only with no visuals)

Group Planning

2. How effectively did the group plan the composition?

0 marks	for no evidence of planning
1 mark	for beginning to develop - some attempt to relate performance to stimulus, eg list of sounds and sound sources, assigning roles/tasks
2 marks	for sound development - linking sound source/instrument and description of sound, eg triangle tinkles to make rain, drum to make thunder clap, clarinet for rusty windmill squeaking
3 marks	- as for '2' plus evidence of attempt to produce a score
4 marks	for well developed - development of a score in conventional or unconventional form which provides clear structure of composition which correlates with final performance and which could be followed by others
5 marks	for highly developed - a well developed musical score using correct terminology and a variety of musical elements, eg melody, rhythm, dynamics, tempo

Stimulus

3. How effectively has the performance reflected the stimulus?

0 marks	for no evidence of stimulus, eg students obviously not relating sounds to storm, painting etc.
1 mark	for beginning to develop - some suggestion of an attempt to reflect isolated events from the stimulus, eg use of instruments as sound effects as in narrative form, eg thunder, horse click clack
2 marks	for sound development - attempt to represent stimulus through appropriate use of musical elements to link sound effects to create some reflection of mood of storm or painting
3 marks	for well developed - good reflection of mood of stimulus throughout sequence through use of dynamics, tempo, melody etc to reflect mood
4 marks	for highly developed - obviously well organised use of instruments, harmony, rhythm, dynamics, form etc to create a clear representation of the mood of the stimulus

Mood

4. How effectively has the group expressed mood?

0 marks	for no evidence - dynamics, tempo etc constant - series of sound effects with no musical qualities
1 mark	for beginning to develop - slight variation in dynamics attempted to create some feeling of mood, eg loud and soft
2 marks	for sound development - evidence of use of sound sources/instruments to create mood using variety in dynamics, melody or tempo
3 marks	for well developed - evidence in planning and performance of organised structure to reflect mood through variety in dynamics, melody, tempo, rhythm etc
4 marks	for highly developed - planning and performance provide evidence of an appealing composition which clearly illustrates mood of stimulus - inclusion of terminology, eg Forte, pianissimo, crescendo etc

Form

5. How effectively has the group expressed form?

0 marks	for no evidence - a continuous, non-decipherable series of sounds without change
1 mark	for beginning to develop - evidence of attempt at beginning, middle, end - as for narrative
2 marks	for sound development - obvious change in expression, instrumentation, to denote beginning, middle, end - organisation within group evident
3 marks	for well developed - well organised structure to show form through a variety of elements, eg dynamics, tempo, rhythm, texture etc - obvious organisation within group
4 marks	for highly developed - evidence of organisation and leadership in planning and presentation which results in an appealing musical piece with a variety of elements to create a well planned form, eg canon, coda, theme etc

Variety of instruments/sound sources

6. How effectively has the group used a variety of instruments/sound sources?

0 marks	for no evidence - no imagination used to vary sounds to link with stimulus
1 mark	for beginning to develop - some attempt made to vary sounds to link with stimulus, eg cymbal/thunder
2 marks	for sound development - obvious link to stimulus through planning and performance and variation in instruments, eg xylophone for sunshine, drum for thunder, triangle for rain, but not much evidence of linking for quality of tone
3 marks	for well developed - good balance of instruments/sound sources to create good tone and texture to reflect the stimulus
4 marks	for highly developed - pleasing combination of instruments, silence, voices etc to create an appealing tone with use of texture

Expression

7. How effectively has the group used expression?

0 marks	for no evidence - no expression - even - all loud or all soft
1 mark	for beginning to develop - slight changes in dynamics - loud/soft
2 marks	for sound development - obvious variation in dynamics, tempo and/or melody in an attempt to reflect mood
3 marks	for well developed - effective use of dynamics, tempo, rhythm, melody, harmony, tone, etc to reflect mood - some evidence of organisation in planning as well as performance
4 marks	for highly developed - exceptional use of elements to create a pleasing sense of expression which clearly conveys mood - inclusion of appropriate variety of dynamics, tempo, rhythm, melody, harmony, tone, texture, legato, staccato etc - evidence of organisation/leadership in planning and performance

Harmony

8. How effectively has the group used harmony?

0 marks	for no evidence - single, random, isolated sounds
1 mark	for beginning to develop - sounds are simultaneous and linked, with slight contrast to reflect stimulus
2 marks	for sound development - evidence at an attempt at texture, some layers of sound and contrast
3 marks	for well developed - evidence in planning and performance to organise sounds to provide contrast of harmony with two or more melodies and clear evidence of texture
4 marks	for highly developed - a pleasing sense of harmony, obviously planned and organised, through use of melodies, texture, contrast and a sense of complementary sounds in the construction of form

Audience

9. How well did the group perform for an audience?

0 marks	for no evidence - no awareness of audience - not sure when to start, no organisation of positions - do not relate to other performers
1 mark	for beginning to develop - some evidence of organisation of positions, signal to start etc
2 marks	for sound development - thought has gone into appropriate positioning of performers and organisation of performance, eg signal to start, and performers relate to each other for cues
3 marks	for well developed - obvious organisation of appropriate positioning of performers in relation to instruments for audience, appointed leader for cues, some acknowledgement of audience
4 marks	for highly developed - polished performance with organisation/cooperation between performers and all performers well positioned with leader providing appropriate cues - confident acknowledgement of audience

INDIVIDUAL ASSESSMENT

Individual planning

P. How effectively did the individual student explore ideas to reflect the stimulus?

0 marks	for no evidence - no attempt or irrelevant to the stimulus and task
1 mark	for beginning to develop - has attempted to display one or two ideas which reflect the stimulus but with little imagination
2 marks	for sound development - ideas are literal but quite imaginative and link with the stimulus to creative sound ideas
3 marks	for well developed - non-literal, impressionistic, abstract ideas transformed into musical ideas which reflect atmosphere

COMMENTS (Use of elements)

C1. How effectively did the student identify and justify the importance of selected musical elements in the composition?

0 marks	for no evidence - selection made but no justification
1 mark	for beginning to develop - tautological or irrelevant justification that does not demonstrate understanding, eg "we did good melody"
2 marks	for sound development - justification which indicates understanding of the term but does not link to composition/performance, eg "melody - because it had a good tune"
3 marks	for well developed - justification which indicates understanding and which is linked to the composition/performance, eg "We used high notes in the melody to show sunshine and low notes to show the storm"
4 marks	for highly developed - justification which indicates understanding, which is linked to the composition, and which is discussed in relation to other elements in the context of the whole work, eg "we used a minor key for our melody with a slow tempo to create a mood of eeriness for the deserted farm"

COMMENTS (improvements to the composition/performance)

C2. How effectively did the student show awareness for making improvements to the overall composition (notation and performance)?

0 marks	for no evidence - irrelevant or tautological response, eg "We would make it better" or "we would make lots of changes," or no indication of room for improvement, eg "none" or "nothing"
1 mark	for beginning to develop - responses with no explanation or justification, eg "we would have more loud and soft"
2 marks	for sound development - responses which relate more specifically to parts of the performance, eg "we would have started off very softly and built up to a loud ending"
3 marks	for well developed - responses which identify areas of weakness or a need for more purpose in the composition or show awareness of a need for more shape or form, eg "next time we would have a softer tone at the beginning, using only the flute and then build up layers with other instruments"
4 marks	for highly developed - responses as for '3' with correct use of technical/music- specific language "Forte, pianissimo, decrescendo, modulation, pentatonic scales" etc

ADMINISTERING THE MUSIC ANALYSIS ASSESSMENT

Time allowance: approximately 45 minutes

Teacher will need:

the stimulus cassette
this booklet
audio cassette player

Each student will need:

a sharpened pencil and eraser
student booklet

Introduction (5 minutes)

When students are gathered quietly in the assessment area explain that you will be reading out the instructions because all classes doing these activities have to receive the same information.

Distribute the student booklets. Then say:

Today you will be helping to gather important information about how students are progressing at school and it is very important that you do your best. You will need to listen to all my instructions very carefully so that you understand what to do.

Throughout this test you will be listening to a music tape and then you will answer questions about it.

Analysis task : (40 minutes)

You will find that the test is in sections called Part 1, Part 2, etc. Each part is clearly labelled on the students' test papers. Each of these parts corresponds to a section of music on the tape. There is a pause of 5 seconds between each part on the tape. A voice will announce the end of each part. You should then pause or stop the tape.

Note: Because students work at different rates there will be unavoidable waiting time for some students. It is advisable to minimise this waiting time by noting when the majority of the class has reached the point to play the next part on the tape, and having all the class restart at this point. Those students who may need more time to complete answers can return to unfinished answers when all the tape has been played.

Say:

In a few moments I am going to play a piece of music for you. It is important that you listen very carefully as you are going to answer some questions about it. The music is in parts on the tape. You will be answering a group of questions after you hear each part. Don't try to answer any questions until the music stops and we read the questions again. I will tell you when to write your answers. Now look at your test paper and you will see the questions in Part 1. I will read these questions while you follow.

1. *Where do you think you are most likely to hear this piece of music?
at a
 birthday party
 orchestral concert
 street parade
 rock concert*
2. *Explain what you heard in the music that made you pick this answer.*

Answer any comprehension questions, ensuring that all students understand what to do.

Then say:

I am going to play the music for this part now. Think about where you might hear this music as you listen. When the music stops we will read the questions again, then you are to write your answers. Now listen carefully to the music.

Play Part 1 on the tape. A voice will announce the end of Part 1. At this point stop or pause the tape. (There is a pause of five seconds between each segment of music.) Then read questions 1 and 2 and allow time for students to write their answers between each one.

When you consider that the majority of the students have completed the answers, continue. Say:

Part 2 is made up of questions 3, 4, 5 and 6. Follow as I read them aloud.

3. *If you moved to this part of the music would you
 march skip walk run*
4. *Explain what you heard in the music that made you pick this answer.*
5. *This music sounds
 angry happy sad sleepy*
6. *Explain what you heard in the music that made you pick this answer.*

Now listen carefully to Part 2 and think about the answers to the questions.

Play Part 2 on tape. Then read each question separately, allowing time between each one for the students to write their response.

When you consider the majority of the students have completed their response to a question, move on to the next question.

After question 6 say:

Part 3 is made up of questions 7 and 8. Let's read them together.

7. *Does the beat stay the same or does it change?*
same changes

8. *Explain what you heard in the music that made you pick this answer.*

Now think about the beat as you listen to this part.

Play Part 3 on the tape. Read each question separately, again allow time in between for the students to write their responses.

When you consider the majority of the students have completed questions 7 and 8, continue to Part 4. Say:

Part 4 is made up of question 9. I will read it now.

9. *What shape does this section of music sound most like?*

Look carefully at the shapes as you listen to the music and decide which one the music sounds most like. Now listen carefully.

Play Part 4 on tape. Allow a reasonable time for question 9 to be completed.

Then say:

Part 5 also consists of only one question. Let's read question 10 in Part 5 before we listen to more of the tape.

10. *What is the main instrument playing this part of the music?*
Clarinet flute french horn trumpet

Listen carefully to the instruments as I play part 5.

Play Part 5 on tape. Then read question 10 again and allow a reasonable time for question 10 to be completed.

Say:

Part 6 is question 11. Watch and listen carefully as I read it. Question 11 says:

11. *Which of these rhythm patterns can you hear in this part ?*

Look at the four rhythm patterns. You will need to listen to the tape very carefully while you look at the rhythm patterns and decide which one you can hear on this part of the tape.

Ensure that all students have looked at the rhythm patterns and then play Part 6 on the tape. Then say:

Now tick the box next to the rhythm pattern you heard.

Allow students to complete question 11. Then say:

We are up to the last questions now. As before we will read the questions then listen to the tape and then write the answers. Question 12 says:

12. Explain how the music ends. (You will need to listen very carefully to the ending of the music to answer this question)

Question 13 asks you:

13. Do you like this music? yes no

Explain what you heard in the music that made you pick this answer.

There is no right or wrong answer here, it is your own opinion which is needed.

Also you must explain what you heard in the music that made you pick this answer.

Now, don't forget to listen carefully to the ending of the music as I play the tape.

Play Part 7 of the tape and read each question again allowing time between each for the answers to be written as before.

If time permits, allow students to complete any unfinished responses on the paper.

Collect student booklets.

YEAR 7 MUSIC ANALYSIS

Time allocation	<u>50 minutes</u>
------------------------	-------------------

- | | |
|-------------------------------|------------|
| • Introduction | 5 minutes |
| • Completion of analysis task | 45 minutes |

Prior to administering the assessment task:

- Become familiar with these administration guidelines and the audio cassette stimulus
- Organise suitable space, free of outside disturbances, for the assessment task
- Organise student seating as for a test situation
- Organise an audio cassette player
- Ensure that sound quality is good and can be heard clearly in all parts of the classroom

During the assessment task:

<i>Please read aloud to the students all instructions enclosed in a box like this</i>

- Help students having difficulty following the instructions but **do not help students with actual task**
- You may answer reading comprehension questions if students do not understand a question, but **do not give explanations of specific terminology related to music**

After the assessment task:

- Collect all materials from students.

ADMINISTERING THE MUSIC ANALYSIS ASSESSMENT

Time allowance: approximately 50 minutes

Teacher will need:

the stimulus cassette
this booklet
audio cassette player

Each student will need:

a sharpened pencil and eraser
student booklet

Introduction

(5 minutes)

When students are gathered quietly in the assessment area explain that you will be reading out the instructions because all classes doing these activities have to receive the same information.

Distribute the student booklets. Then say:

Today you will be helping to gather important information about how students are progressing at school and it is very important that you do your best. You will need to listen to all my instructions very carefully so that you understand what to do.

Throughout this test you will be listening to a music tape and then you will answer questions about it.

Analysis task :

(45 minutes)

You will find that the test is in sections called Part 1, Part 2, etc. Each part is clearly labelled on the students' test papers. Each of these parts corresponds to a section of music on the tape. There is a pause of 5 seconds between each part on the tape. A voice will announce the end of each part. You should then pause or stop the tape.

Note: Because students work at different rates there will be unavoidable waiting time for some students. It is advisable to minimise this waiting time by noting when the majority of the class has reached the point to play the next part on the tape, and having all the class restart at this point. Those students who may need more time to complete answers can return to unfinished answers when all the tape has been played.

Say:

In a few minutes I will play a piece of music for you. It is important that you listen very carefully as you are going to answer some questions about it. The music is in parts on the tape. You will be answering a group of questions after you hear each part. Don't try to answer any questions until after the music stops. Look at your test paper and you will see Part 1. Now I would like you to read the two questions in part 1. Ask about any words that you are not sure of.

Allow time for students to read the questions in Part 1 which are as follows:

Part 1

1. Which group of instruments is playing this piece of music?

brass band

string quartet

symphony orchestra

concert band

2. Explain what you heard in the music that made you pick this answer.

Answer any comprehension questions without giving explanations of specific terminology related to music. Then say:

Listen carefully to this piece of music and think about what group of instruments is playing it as you listen. Don't write anything yet.

Play Part 1 on the tape. A voice will announce the end of Part 1. At this point stop or pause the music. (There is a five second silence between each segment of music.) Say:

Now you may answer the questions in Part 1 only.

Allow a **reasonable** time for completion of this Part, i.e. until you consider the majority of students have completed their answers.

Then say:

Part 2 is made up of questions 3, and 4, Read these questions through before I play the next part on the tape.

Allow the students time to look at the questions in Part 2 which are as follows:

Part 2

3. Does the beat stay the same or does it change?

same

changes

4. Explain what you heard in the music that made you pick this answer.

Answer comprehension questions if necessary. Then say:

Listen carefully now as I play Part 2 and remember you are listening for the beat.

Play Part 2 on the tape. Then say:

You may now answer the questions in Part 2 only.

Allow a **reasonable** time for completion of this part, moving around the room and supervising as for a test situation. When you consider the majority of students have completed their answers say:

Now look at Part 3 before I play that part of the tape. You will see some shapes. This time you are to decide which of these shapes the music sounds most like.

Part 3

5. *What shape does this section of music sound most like?
(see student booklet)*

When students have looked at the shapes say:

Now listen carefully to the tape as I play Part 3. Remember you are matching one of these shapes to the sound of the music.

Play Part 3 of the tape and then allow time, as before, for students to complete answers.

Say:

Look now at Part 4 and read question 6 before I play the next part of the tape.

Part 4

6. *What is the main instrument playing this part of the music?
clarinet flute french horn trumpet*

Say:

Listen carefully now to Part 4 and decide which is the main instrument playing this part.

Play Part 4 of the tape and then allow time, as before, for students to answer the question.

Then say:

Now we are moving on to Part 5 which is about rhythm. You are to look at the rhythm patterns on your paper and match one of them with the rhythm on the tape.

Part 5

7. *Which of these rhythm patterns can you hear in this part ?
(see student booklet)*

Allow time for students to look at the rhythm patterns, then say:

Now listen carefully to Part 5. Remember you are listening for the rhythm.

Play Part 5 on the tape and allow time, as before, for students to answer the question.

Then say:

Now read questions 8 and 9 in Part 6 before I play the next part.

Part 6

8. *What mood or feeling does this piece of music create for you?*

*happiness
sadness
anger
excitement*

9. *Explain what you heard in the music that helped to create the mood or feeling.*

Allow time for students to read the questions. Then say:

Now listen to the tape and think about the mood of the music as you listen.

Play Part 6 of the tape and allow time, as before, for students to answer the questions.

Then say:

Now look at question 10 in Part 7.

Part 7

10. *Explain how the music ends.*

Allow students time to look at the question, then say:

This time, you will need to listen very carefully to the ending of the music.

Play Part 7 of the tape and then allow time, as before, for students to complete their answers.

Say:

On the next part of the tape you will hear a different piece of music. Questions 11, 12 and 13 in Part 8 refer to the new piece. Read the questions in Part 8 now, before I play this part on the tape.

Part 8

11. *In this piece of music the composer has combined different styles of music. What are they?*

12. *What is it in the music that helped you to decide?*

13. *What effect has this style of music had on Australian culture?*

Allow time for the students to read questions 11, 12 and 13. Then say:

Now listen carefully to the new piece of music and think about the questions as you listen.

Play Part 8 of the tape and allow students time, as before, to complete their answers.

Then say:

The last part of the tape is Part 9. This part refers to both pieces you have heard today. Read the questions in part 9 and then I will play the tape.

Allow students time to read questions 14 and 15 in Part 9.

Part 9

14. *Compare and contrast the two pieces of music you have heard, using the following headings:*

- a) *Instrumentation (how the instruments are used)*
- b) *Expression (dynamics, tempo, timbre)*
- c) *Rhythm*

15. *Which of these two pieces of music did you prefer?*

first piece

second piece

Explain what you heard in the music that made you choose this piece.

Then say:

Listen carefully now, as I play the last part of the tape and think about questions 14 and 15 as you listen.

Play Part 9 of the tape and allow students time, as before, to complete their answers. If time permits allow students to complete any unfinished responses on the paper.

Collect student booklets

YEAR10 MUSIC ANALYSIS

Time allocation	<u>60 minutes</u>
------------------------	-------------------

- | | |
|-------------------------------|------------|
| • Introduction | 5 minutes |
| • Completion of analysis task | 55 minutes |

Prior to administering the assessment task:

- Become familiar with these administration guidelines and the audio cassette stimulus
- Organise suitable space, free of outside disturbances, for the assessment task
- Organise student seating as for a test situation
- Organise an audio cassette player
- Ensure that sound quality is good and can be heard clearly in all parts of the classroom

During the assessment task:

<i>Please read aloud to the students all instructions enclosed in a box like this</i>

- Help students having difficulty following the instructions but **do not help students with actual task**
- You may answer reading comprehension questions if students do not understand a question, but **do not give explanations of specific terminology related to music**

After the assessment task:

- Collect all materials from students.

ADMINISTERING THE MUSIC ANALYSIS ASSESSMENT

Time allowance: approximately 55 minutes

Teacher will need:

the stimulus cassette
this booklet
audio cassette player

Each student will need:

a sharpened pencil and eraser
student booklet

Introduction

(5 minutes)

When students are gathered quietly in the assessment area explain that you will be reading out the instructions because all classes doing these activities have to receive the same information.

Distribute the student booklets. Then say:

Today you will be helping to gather important information about how students are progressing at school and it is very important that you do your best. You will need to listen to all my instructions very carefully so that you understand what to do.

Throughout this test you will be listening to a music tape and then you will answer questions about it.

Analysis task :

(55 minutes)

You will find that the test is in sections called Part 1, Part 2, etc. Each part is clearly labelled on the students' test papers. Each of these parts corresponds to a section of music on the tape.

Note: Because students work at different rates there will be unavoidable waiting time for some students. It is advisable to minimise this waiting time by noting when the majority of the class has reached the point to play the next part on the tape, and having all the class restart at this point. Those students who may need more time to complete answers can return to unfinished answers when all the tape has been played.

There is a pause of 5 seconds between each part on the tape. A voice will announce the end of each part.

Say:

In a few minutes I will play a piece of music for you. It is important that you listen very carefully as you are going to answer some questions about it. The music is in

parts on the tape . You will be answering a group of questions after you hear each part. Don't try to answer any questions until after the music stops. Look at your test paper and you will see Part 1. Read through the questions in Part 1 and ask about any words that you are not sure of.

Allow time for the students to read the questions in Part 1 which are as follows:

Part 1

1. *Which group of instruments do you think is playing this piece of music?*

brass band

string quartet

symphony orchestra

concert band

2. *Explain what you heard in the music that made you pick this answer.*

Answer any comprehension questions without giving explanations of specific terminology related to music. Then say:

Listen carefully to this piece of music and think about what group of instruments is playing it as you listen. Don't write anything yet.

Play Part 1 on the tape. A voice will announce the end of Part 1. At this point stop or pause the music. (There is a five second silence between each segment of music.)

Say:

Now you may answer the questions in Part 1 only.

Allow a **reasonable** time for completion of this Part, ie. until you consider the majority of students have completed their answers. Supervise as for a test situation.

Then say:

Part 2 is made up of questions 3 and 4. Read these questions through before I play the next part on the tape.

Allow the students time to look at the questions in Part 2 which are as follows:

Part 2

3. *Does the beat stay the same or does it change in this part of the music?*

same

changes

4. *Explain what you heard in the music that made you pick this answer.*

Answer comprehension questions if necessary.

Then say:

Listen carefully now as I play Part 2 and remember you are listening for the beat.

Play Part 2 of the tape. Then say:

You may now answer the questions in Part 2.

Allow a **reasonable** time for completion of part 2, supervising as before. When you consider the majority of students have completed their answers say:

Now look at Part 3 before I play that part of the tape. You will see some shapes. This time you are to decide which of these shapes the music sounds most like.

Part 3

5. *What shape does this section of music sound most like?*
(see student booklet)

When students have looked at the shapes say:

Now listen carefully to the tape as I play Part 3. Remember you are matching one of these shapes to the sound of the music.

Play Part 3 of the tape and then allow time, as before, for students to complete answers. Then say:

Look now at Part 4 and read question 6 before I play the next part of the tape.

Part 4

6. *What is the main instrument playing this part of the music?*
clarinet flute french horn trumpet

Say:

Listen carefully now to Part 4 and decide which is the main instrument playing this part.

Play Part 4 of the tape and then allow time, as before, for students to answer the question. Then say:

Now we are moving on to Part 5 which is about rhythm. Look at Question 7 and you will see some rhythm patterns on your paper. Match one of them with the rhythm on the tape.

Part 5

7. *Which of these rhythm patterns can you hear in this part?*
(see student booklet)

Allow time for students to look at the rhythm patterns, then say:

Now listen carefully to Part 5. Remember you are listening for the rhythm.

Play Part 5 on the tape and allow time, as before, for students to answer the question. Then say:

Now read questions 8 and 9 in Part 6 before I play the next part.

Part 6

8. *What mood or feeling has the composer created in this piece of music?*
happiness
sadness
anger

excitement

9. Explain what you heard in the music that helped to create the mood or feeling.

Allow time for students to read the questions. Then say:

Now listen to the tape and think about the mood of the music as you listen.

Play Part 6 of the tape and allow time, as before, for students to answer the questions. Then say:

Now look at question 10 in Part 7.

Part 7

10. Explain how the music ends.

Allow students time to look at the question. Then say:

This time, you will need to listen very carefully to the ending of the music.

Play Part 7 of the tape and then allow time, as before, for students to complete their answers. Then say:

On the next part of the tape you will hear a different piece of music. Questions 11, 12 and 13 in Part 8 refer to the new piece. Read the questions in Part 8 now, before I play this part on the tape.

Part 8

11. In this piece of music the composer has combined different styles of music. What are they?

12. What is it in the music that helped you to decide?

13. What effect has this style of music had on Australian culture?

Allow time for the students to read questions 11, 12 and 13. Then say:

Now listen to the new piece of music and think about the questions as you listen.

Play Part 8 of the tape and then allow students time, as before, to complete their answers.

Then say:

The next part of the tape is Part 9. This part refers to both pieces you have heard so far. Read the question in part 9 and then I will play the tape.

Allow students time to read question 14 in Part 9.

Part 9

14. Compare and contrast the two pieces of music you have heard, using the following headings:

- a) Instrumentation (how the instruments are used)
- b) Expression (dynamics, tempo, timbre)
- c) Rhythm

Then say:

Listen now, as I play this part of the tape and think about question 14 as you listen.

Play Part 9 of the tape and allow students time, as before, to complete their answers.

Then say:

Part 10 refers to the third and last piece on the tape. Read questions 15, 16 and 17 now before I play the tape.

Allow time for students to read the questions in Part 10.

Part 10

15. *The composer of this music is an Australian who is describing the landscape. Describe the mood the composer has created with this piece of music.*
16. *What musical elements and effects has he used to achieve this?*
17. *Of the three pieces of music that you have heard today which one do you think is most effective in its use of musical elements?*
first piece
second piece
third piece
Explain your reasons

Play Part 10 of the tape and allow students time, as before, to complete their answers.

If time permits allow students time to complete any unfinished responses on the paper.

Collect student booklets.

TEST-OF-FIT (summary statistics)

	Item-Person Interaction		Persons	
	Items	Std Error	Location	Std Error
Mean	0.000	-1.971	-1.286	-0.257
SD	0.898	7.957	1.175	1.257
Skewness				
Kurtosis				
Correlations		0.000		0.035

Complete data degrees of freedom = 4281.07
 data degrees of freedom = 43.96

Complete

Item-Trait Interaction		Person separation index	
Total Item Chi Sq	3577.143		
0.896			
Total Degree Freedom	88.000	Cronbach	N/A
Total ChiSq Probability	0.000		

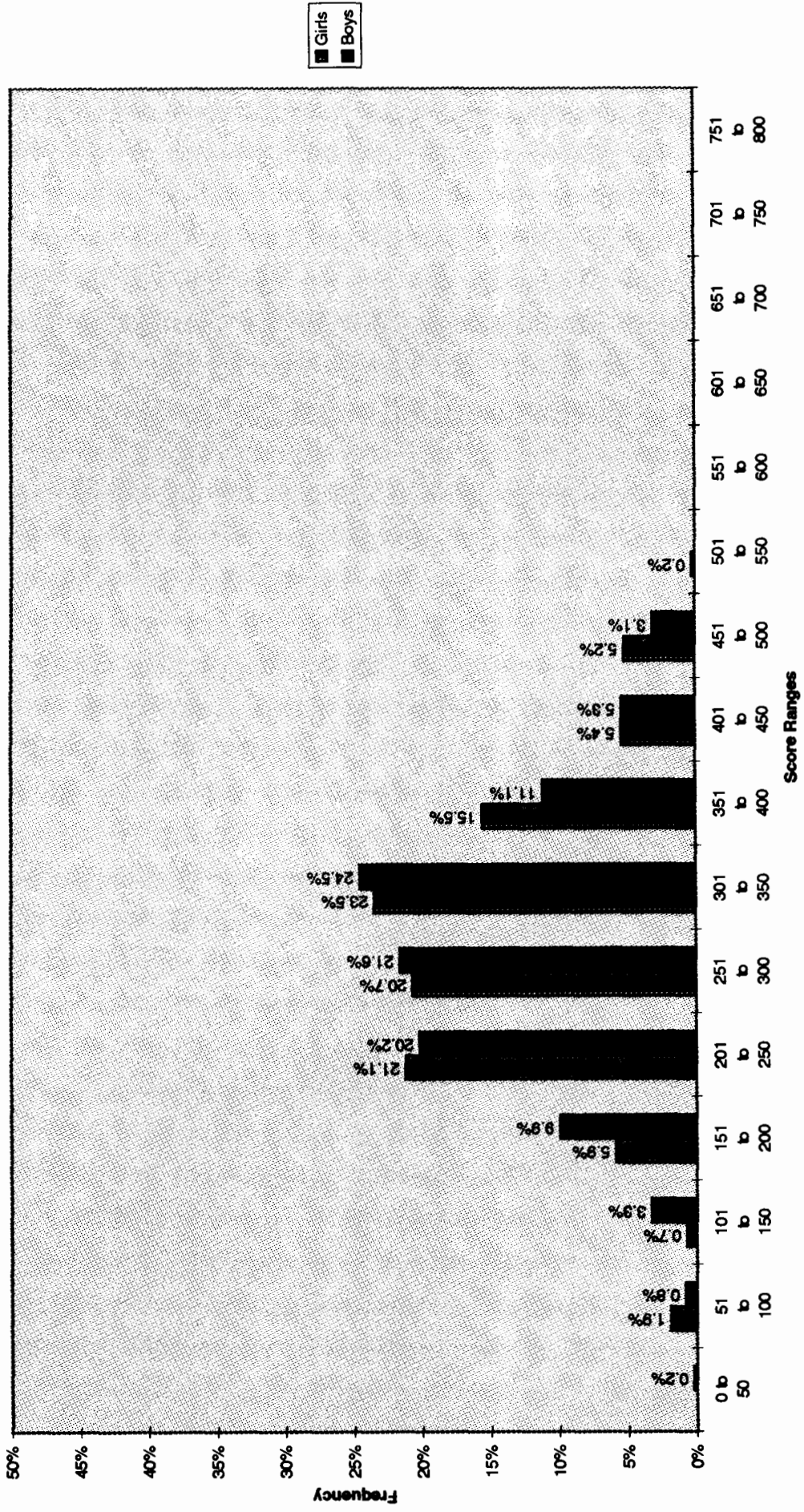
Test of Fit Power EXCELLENT
 Analysis Title: music analysis 6 may 1998
 INDIVIDUAL ITEM-FIT

Label	Location	SE	Fit	ChiSq	Probability
Ex001 MA02	-0.858	0.04	9.701	164.076	0.000
Ex002 MA04	-0.280	0.04	4.397	92.117	0.000
Ex001 MA02	-.356	.356			
Ex002 MA04	-.802	.802			
Ex003 MA06	-1.459	.907	.552		
Ex004 MA07	.000				
Ex005 MA08	-.185	-1.204	1.389		
Ex006 MA12	-3.088	-.574	1.449	2.213	
Ex007 MA13	-2.421	.855	1.566		
Ex008 MA15	-2.024	-.071	2.094		
Ex009 MA17	-1.899	-.138	2.037		
Ex010 MA18	-.018	.018			
Ex011 MA19	-3.398	.051	3.347		
Ex012 MA20	-1.703	-.389	2.093		
Ex013 MA21	-2.087	.029	.224	1.834	
Ex014 MA22	-1.816	-.906	.595	2.128	
Ex015 MA23	-1.945	-.286	1.017	1.213	
Ex016 MA24	-2.839	-.282	1.232	1.889	
Ex017 MA25	-3.037	.263	2.774		
Ex018 MA26	-3.221	-.940	.353	3.809	
Ex019 MA27	-1.988	-.226	.788	1.426	
Ex020 MP01	-3.291	-.272	1.324	2.239	
Ex021 MP02	-3.426	-.271	1.595	1.866	.235
Ex022 MP03	-3.422	-.261	1.500	2.183	
Ex023 MP04	-2.775	-.868	3.541	.101	
Ex024 MP05	-3.019	-.676	2.214	1.481	
Ex025 MP06	-3.443	-.396	1.502	2.337	
Ex026 MP07	-2.128	-1.013	.715	2.426	
Ex027 MP08	-3.904	-.571	2.542	1.933	
Ex028 MP09	-3.843	.170	1.688	1.984	
Ex029 MP10	-4.096	.447	3.649		
Ex030 MP11	-4.179	-.407	1.621	2.965	
Ex031 MP12	-3.266	.464	.971	1.831	
Ex032 MP13	-3.609	.508	3.102		
Ex033 MP14	-4.316	-1.050	1.068	4.298	
Ex034 MP15	-3.439	-.712	1.144	3.008	
Ex035 MP16	-1.580	-1.474	.191	2.864	
Ex036 MP17	-2.717	-.886	.226	3.378	
Ex037 MP18	-3.450	-.590	1.277	2.763	
Ex038 MP19	-3.350	-1.654	-.548	1.106	4.447
Ex039 MP20	-2.824	-.624	1.159	2.289	
Ex040 MP21	-2.257	-1.264	.390	3.130	
Ex041 MP22	-2.766	-.862	.613	3.015	
Ex042 MP23	-2.602	-1.076	1.161	2.517	
Ex043 MP24	-1.714	-1.170	-.173	3.057	
Ex044 MP25	-2.736	-1.271	1.430	2.577	
Ex045 MP26	-3.826	-.850	1.598	3.078	

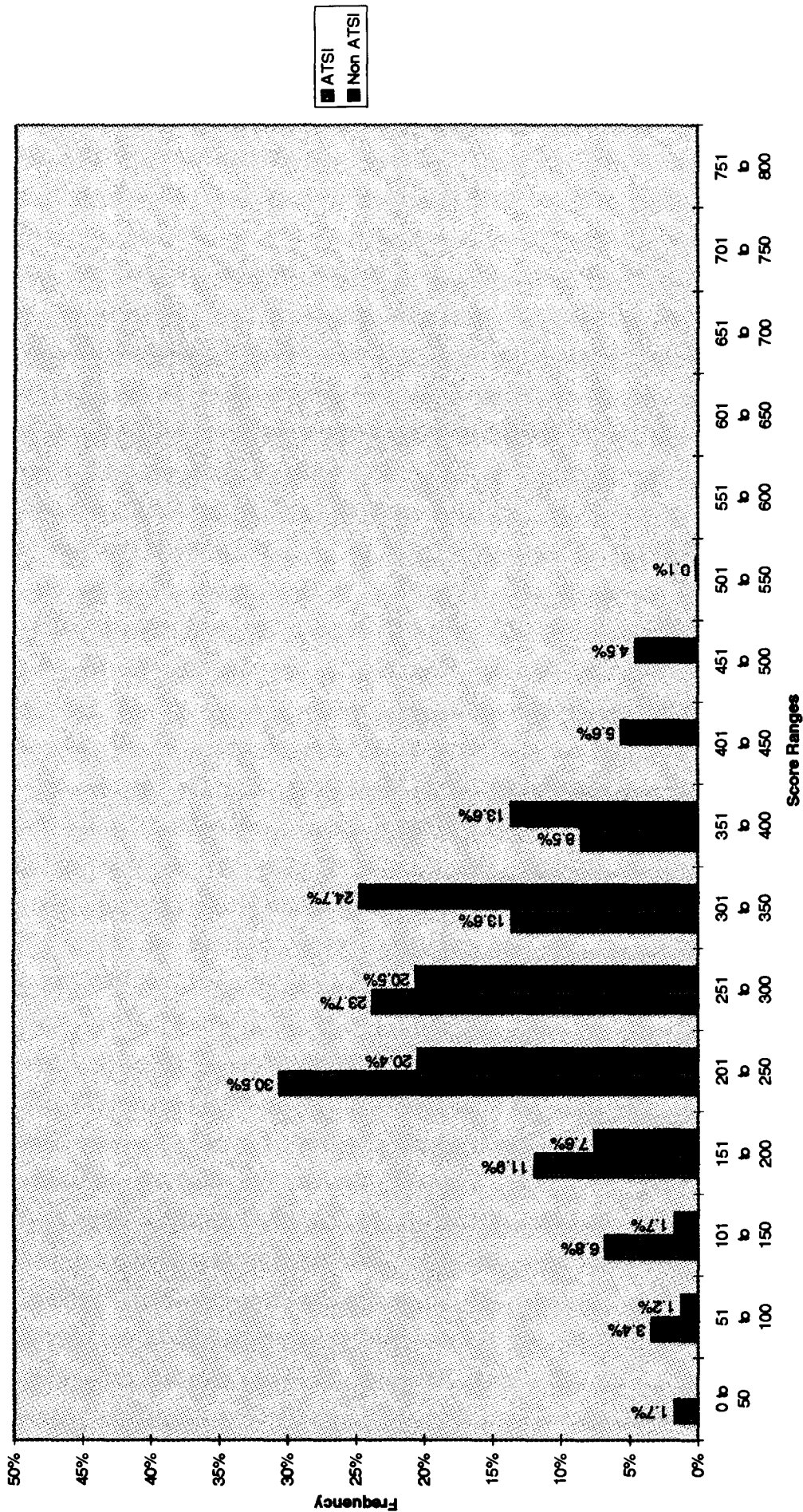
2 Item Parameters

Item Code	Location		Scale		Skewness		Kurtosis	
	Estm	SE	Estm	SE	Estm	SE	Estm	SE
Ex001 MA02	-.858	0.037	.356	0.053	.000	0.000	.000	0.000
Ex002 MA04	-.280	0.043	.802	0.053	.000	0.000	.000	0.000
Ex003 MA06	.553	0.046	.503	0.027	-.227	0.024	.000	0.000
Ex004 MA07	-3.771	0.051	.000	0.000	.000	0.000	.000	0.000
Ex005 MA08	.263	0.022	.394	0.019	.301	0.014	.000	0.000
Ex006 MA12	.566	0.025	.896	0.012	-.073	0.005	-.006	0.002
Ex007 MA13	.268	0.044	.997	0.026	-.214	0.020	.000	0.000
Ex008 MA15	.270	0.032	1.030	0.024	.018	0.014	.000	0.000
Ex009 MA17	.125	0.031	.984	0.024	.034	0.013	.000	0.000
Ex010 MA18	-2.089	0.034	.018	0.052	.000	0.000	.000	0.000
Ex011 MA19	.881	0.039	1.686	0.028	-.013	0.014	.000	0.000
Ex012 MA20	.662	0.032	.949	0.023	.097	0.015	.000	0.000
Ex013 MA21	1.060	0.031	.598	0.015	-.021	0.006	.028	0.004
Ex014 MA22	1.099	0.030	.667	0.015	.026	0.007	-.005	0.003
Ex015 MA23	1.373	0.033	.539	0.015	-.061	0.007	-.006	0.004
Ex016 MA24	.371	0.038	.785	0.019	-.079	0.008	.002	0.004
Ex017 MA25	.435	0.069	1.453	0.069	-.066	0.023	.000	0.000
Ex018 MA26	1.013	0.056	1.119	0.045	.049	0.010	.026	0.005
Ex019 MA27	.810	0.048	.563	0.031	-.047	0.011	.003	0.005
Ex020 MP01	.128	0.039	.909	0.019	-.088	0.008	.006	0.004
Ex021 MP02	-.217	0.038	.473	0.014	-.133	0.005	-.003	0.001
Ex022 MP03	.204	0.041	.929	0.019	-.103	0.008	.003	0.004
Ex023 MP04	.981	0.040	.652	0.016	-.223	0.008	-.086	0.004
Ex024 MP05	.356	0.038	.820	0.017	-.128	0.008	-.035	0.004
Ex025 MP06	.158	0.040	.962	0.019	-.092	0.008	.001	0.004
Ex026 MP07	.502	0.034	.769	0.016	.025	0.008	-.005	0.004
Ex027 MP08	.270	0.043	1.031	0.021	-.164	0.008	-.029	0.003
Ex028 MP09	.236	0.046	.950	0.021	-.155	0.008	.011	0.004
Ex029 MP10	.488	0.052	1.936	0.033	-.112	0.019	.000	0.000
Ex030 MP11	-.025	0.045	1.173	0.025	-.101	0.008	.009	0.003
Ex031 MP12	.684	0.048	.790	0.019	-.120	0.009	.030	0.005
Ex032 MP13	-1.313	0.037	1.678	0.037	-.127	0.012	.000	0.000
Ex033 MP14	-.428	0.039	1.398	0.034	-.001	0.007	.019	0.003
Ex034 MP15	-.054	0.038	1.060	0.024	-.036	0.007	.007	0.003
Ex035 MP16	-.251	0.059	.750	0.037	.107	0.012	-.005	0.005
Ex036 MP17	.242	0.059	.970	0.045	.055	0.011	.023	0.005
Ex037 MP18	-.917	0.030	1.025	0.022	-.057	0.006	.005	0.002
Ex038 MP19	.022	0.026	.918	0.015	.046	0.003	.009	0.001
Ex039 MP20	-.857	0.028	.856	0.019	-.045	0.006	-.002	0.002
Ex040 MP21	-.150	0.027	.891	0.017	.073	0.006	.004	0.003
Ex041 MP22	-.655	0.028	.941	0.020	.021	0.006	.011	0.002
Ex042 MP23	-.763	0.029	.880	0.019	-.007	0.006	-.013	0.002
Ex043 MP24	-.087	0.025	.766	0.016	.112	0.006	.015	0.003
Ex044 MP25	-.758	0.030	.932	0.020	-.013	0.006	-.023	0.002
Ex045 MP26	-.546	0.032	1.158	0.024	-.062	0.006	-.004	0.002

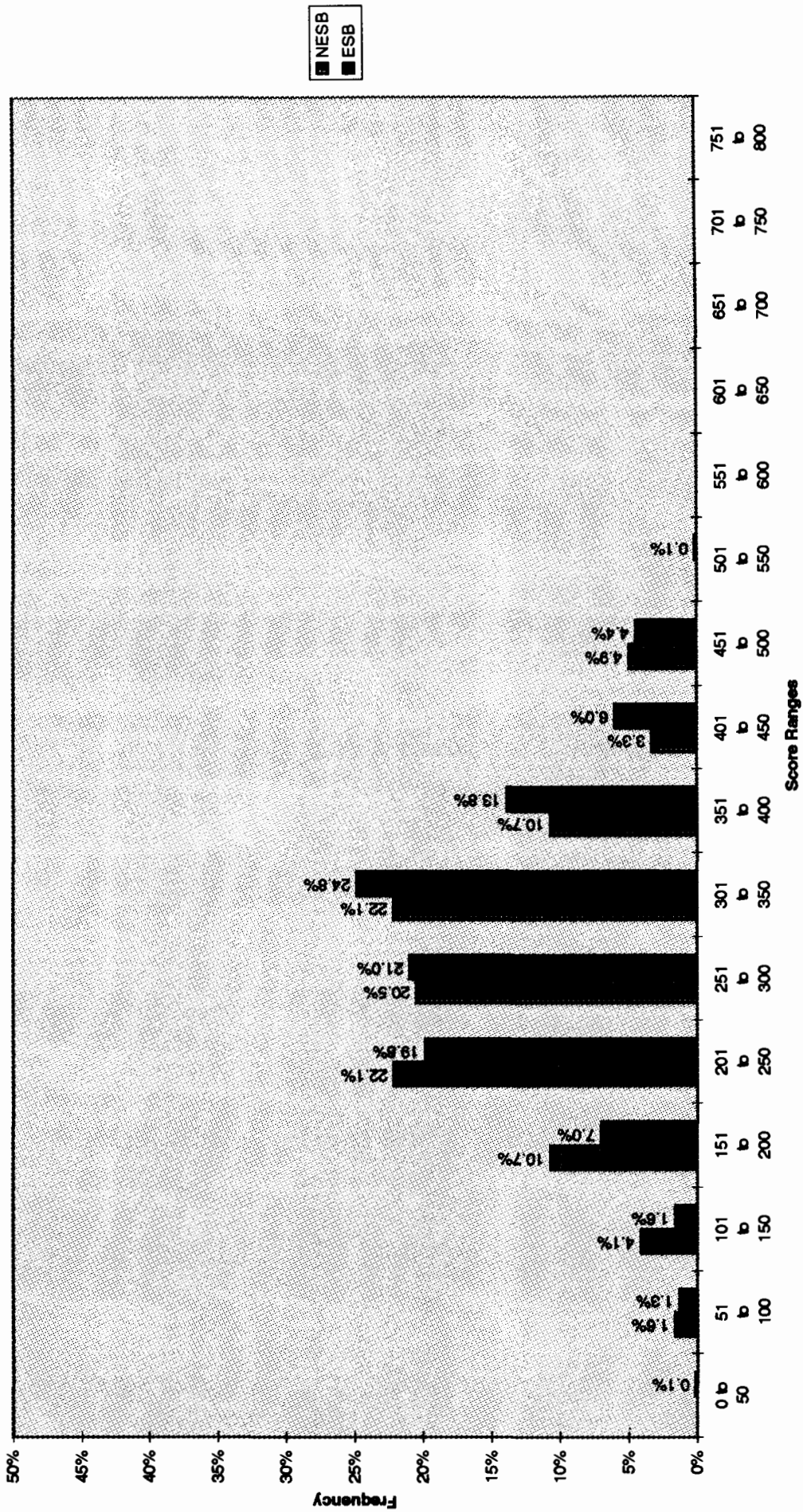
Year 3 Girls vs Year 3 Boys



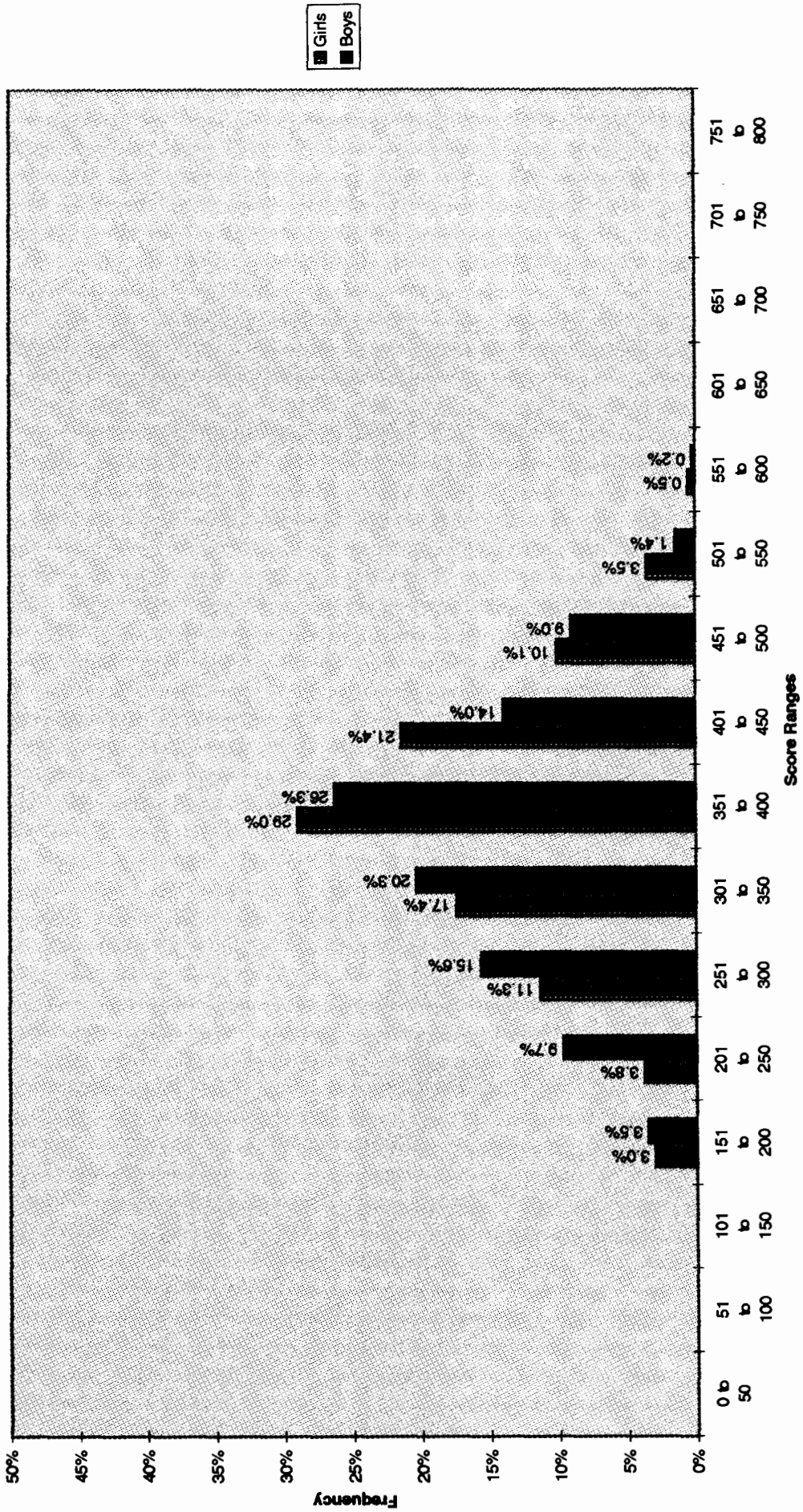
Year 3 ATSI vs Year 3 Non ATSI



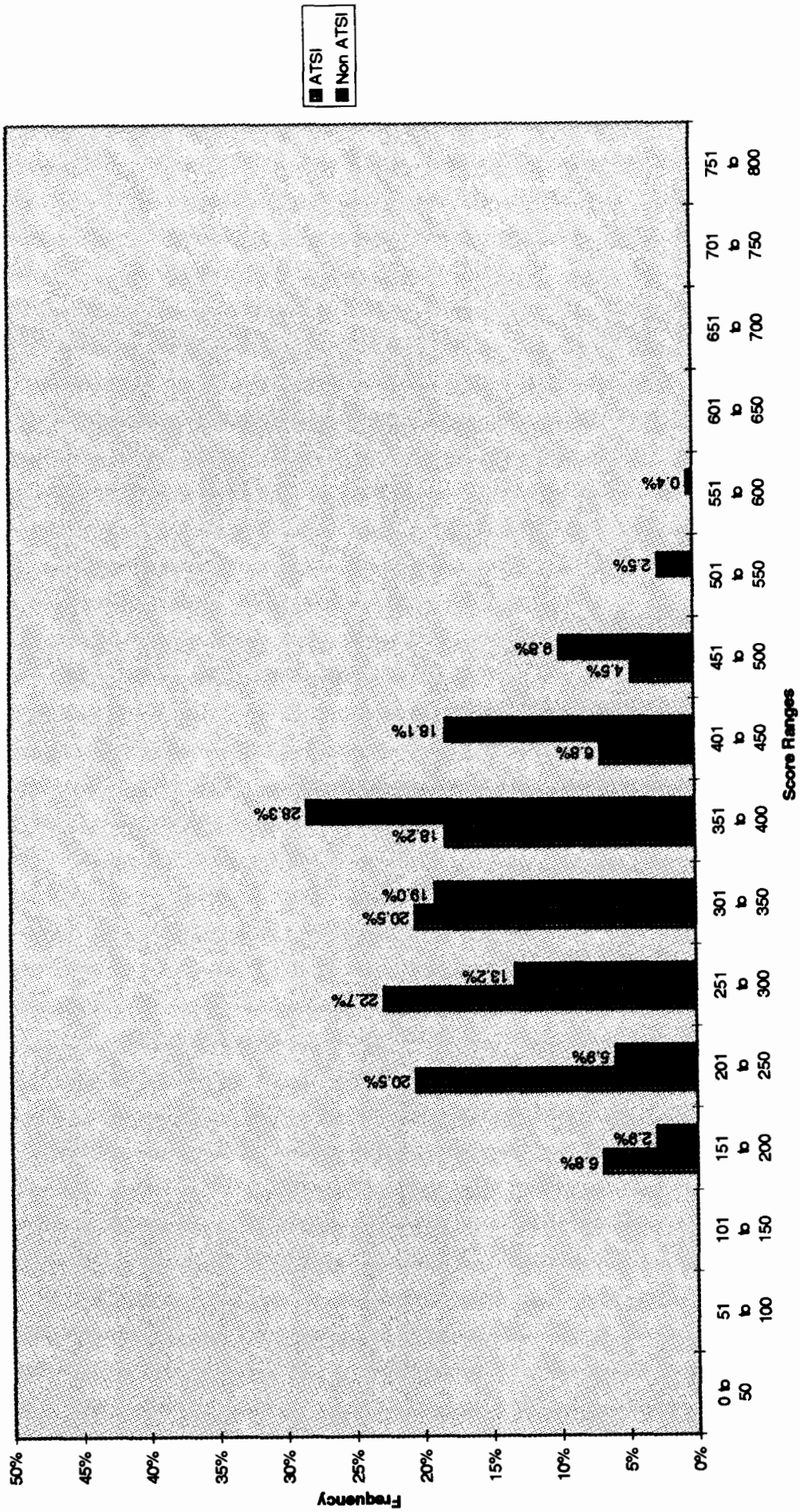
Year 3 NESB vs Year 3 ESB



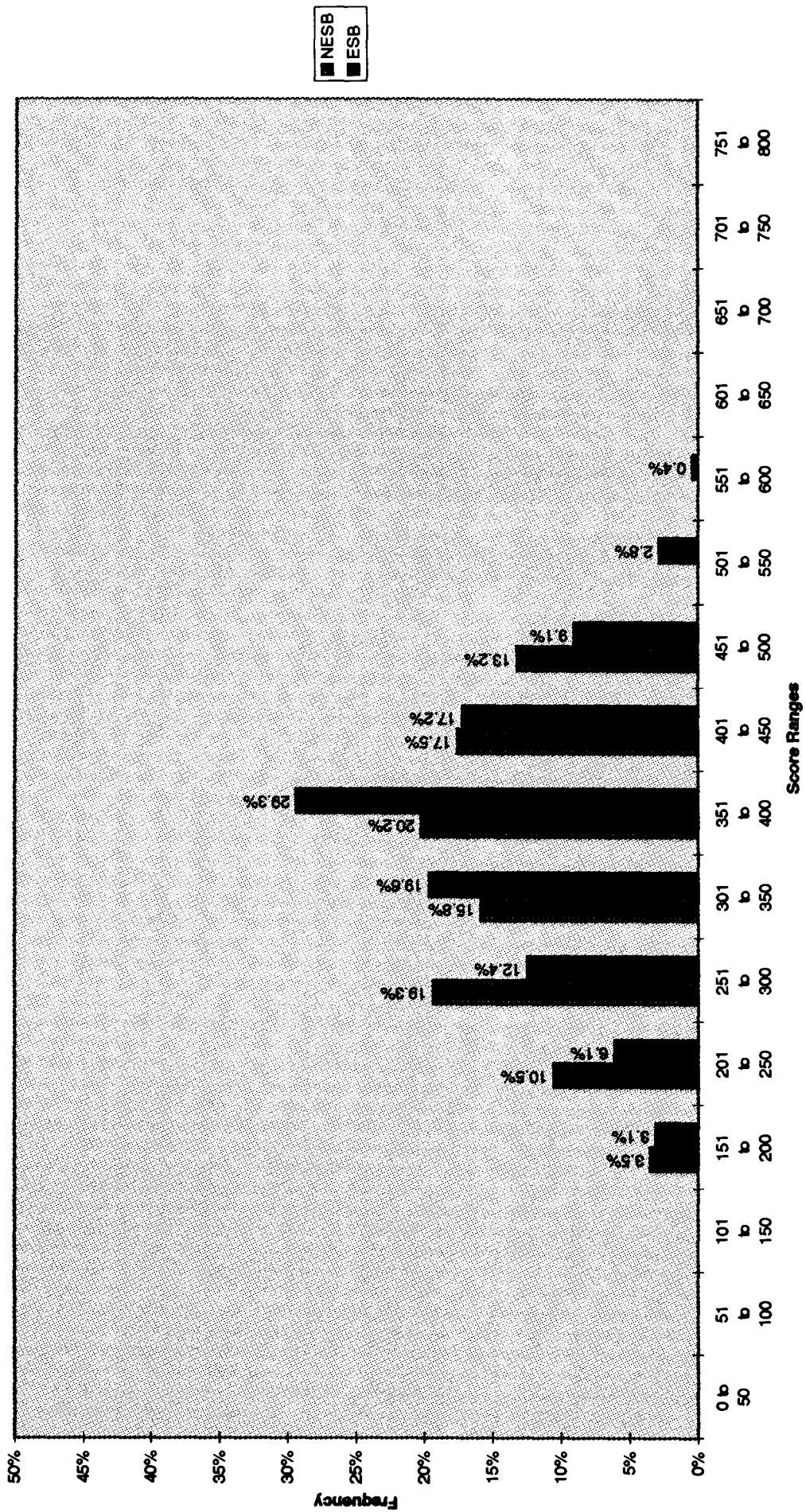
Year 7 Girls vs Year 7 Boys



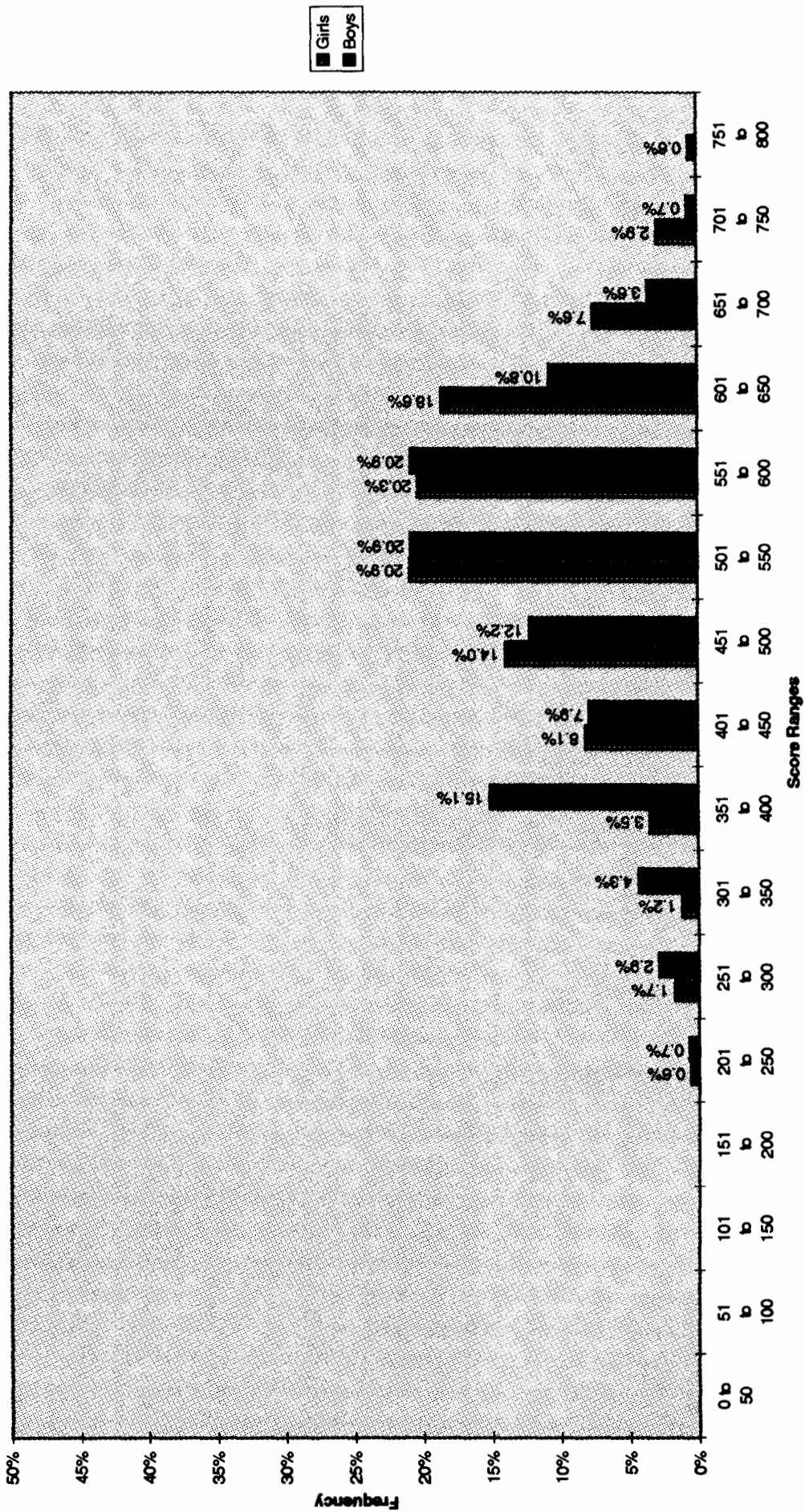
Year 7 ATSI vs Year 7 Non ATSI



Year 7 NESB vs Year 7 ESB

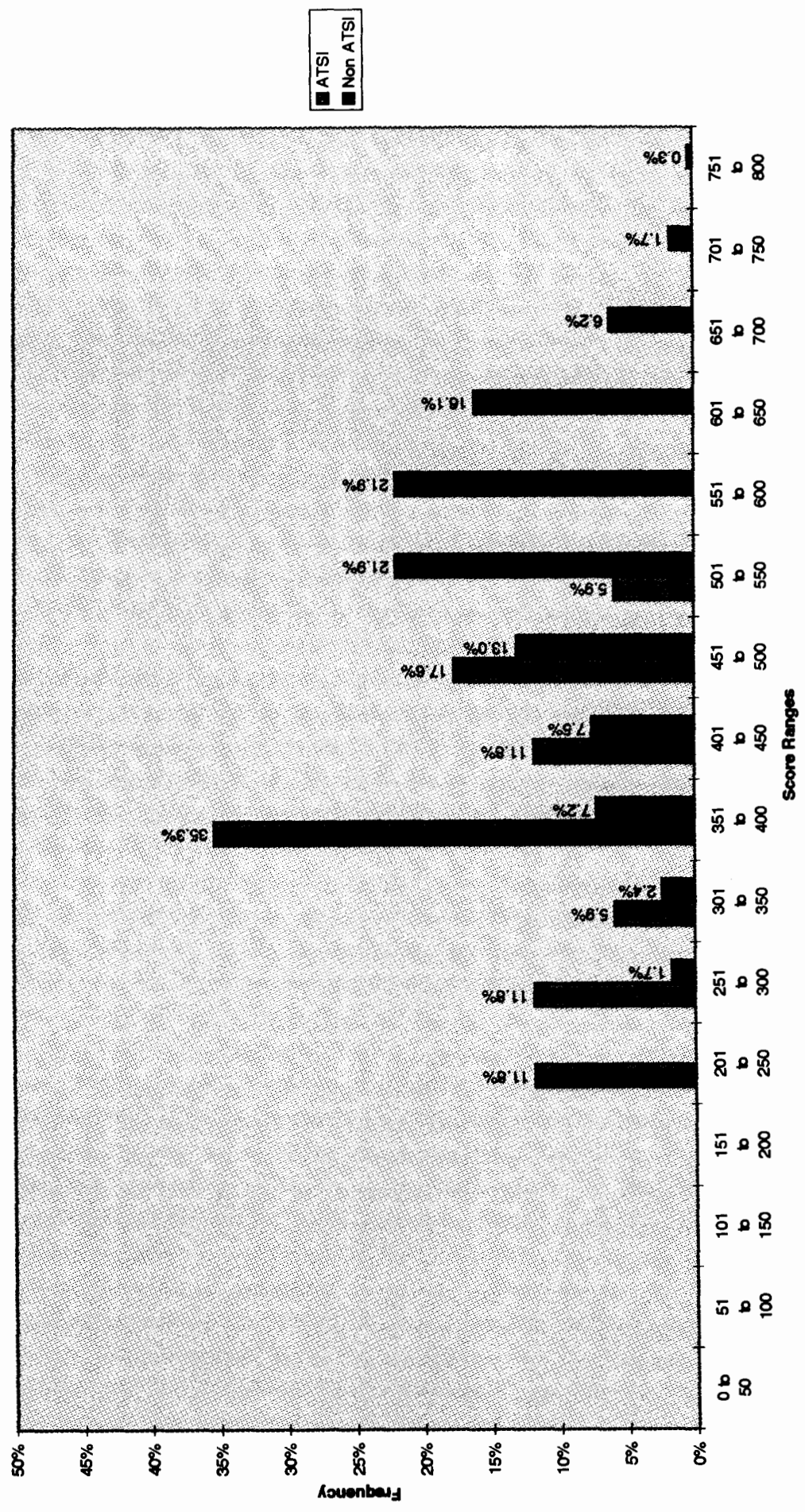


Year 10 Girls vs Year 10 Boys

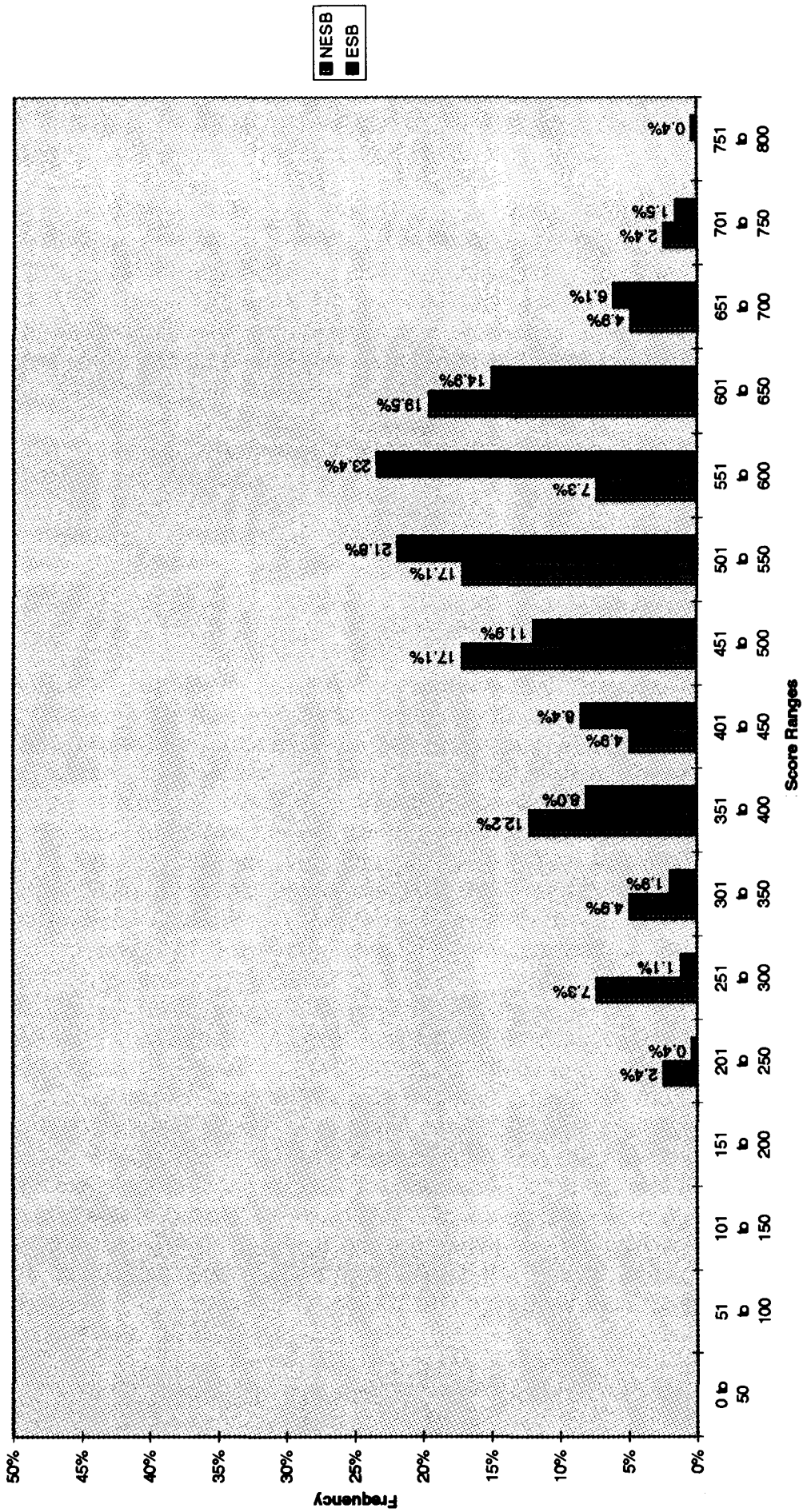


1996 MUSIC

Year 10 ATSI vs Year 10 Non ATSI



Year 10 NESB vs Year 10 ESB



THE EXPRESSING STRANDS OF THE ARTS STUDENT OUTCOME STATEMENTS (EDUCATION DEPARTMENT OF WESTERN AUSTRALIA, 1996)

The 'Expressing' strands describe the use of skills, techniques and technologies in music in exploring, developing, creating and communicating through students' musical activities and musical works. The first strand of the model is *Creating, exploring and developing ideas*, and the second is *Using skills, techniques, technologies and processes*. The two strands are developed into eight ordered levels of achievement from low (level 1) to high (level 8). The statements for the two strands at Level 1 are:

Uses play, sensory experiences and imagination as starting points for arts activities; and

Uses simple arts skills and processes in sharing their arts activities informally with others (Education Department of Western Australia, 1996, p.1).

At Level 1, students are self-focused and use their play and sensory experiences as a basis for making and sharing arts activities. They work from the familiar and make connections to their arts activities, using simple skills to explore, create, develop and communicate ideas in informal settings. They use simple processes involving planning, expressing and reflecting in an exploratory way, not always being clearly aware of the learning that is taking place. At Level 1 children draw significantly on their experiences of play where they take on roles, explore stories, and experiment with sounds, patterns and movements. At the same time, they begin to make choices about using some of these elements in more directed ways, sometimes shaping them for sharing with others.

Musical experiences at Level 1 include performing variations to known activities and songs, inventing new actions for songs and suggesting ways of moving to the beat of songs. Students begin to participate alone in a given role within the context of a singing game, echoing simple patterns. They can match a key feature with movement such as marching or swaying to a beat, or patschen clapping to accents in twos. They can match sounds and symbols through pictures or graphic notation and follow a simple score by one to one correspondence. They reproduce environmental sounds with voice, body or instruments such as high pitched noises for a fire siren, clicking noises with their tongues for horses galloping or trotting (Education Department of Western Australia, 1996, p.9).

The statements for the two strands at Level 2 are:

Builds on their experiences in exploring, shaping and
Communicating Ideas for personal creative expression; and

Uses the skills, technologies and simple arts processes that
they have learned in making and sharing arts activities
informally with others (Education Department of Western
Australia, 1996, p.1).

In progressing to Level 2, students begin to move beyond self focus and to utilise the skills learnt by their active participation at Level 1, in directed activities and arts works. They develop the ability to recognise features and elements of arts work. The depth of understanding of features and elements is evident through their responses and participation. Students are able to make conscious simple choices to use appropriate elements in guided activities in the making of musical works and show that they organise and arrange those elements to create a final product. They present their musical works to a familiar audience, such as classmates or parents.

Students at Level 2 sing more accurately in whole class songs and singing games. They improvise an answering phrase in a structured game situation and are able to work with a partner in a structured task such as playing improvised phrases on percussion instruments or creating their own accompanying patterns, using a limited range. Skills also include the creation of body percussion with a partner to accompany a song or recorded music. They can not necessarily use conventional musical notation, but they can read simple scores, such as soundscape or graphic notation, and can notate their own rhythms, melodies and accompaniment patterns using these methods (Education Department of Western Australia, 1996, p.14).

The statements for the two strands at Level 3 are:

Explores and develops ideas and feelings through arts activities using
given contexts; and

Uses a range of skills, techniques, technologies and processes in
communicating for an audience or purpose (Education Department of
Western Australia, 1996, p.1).

By Level 3, students have acquired a degree of technical skill and facility which enables them to successfully re-create musical works and realise some of their own creative ideas. However, they still work within the structure of a given task with a limited range of choices and a clear sense of the end purpose of their activity. They are now able to sing simple two part songs, rounds and canons. Students at this level can improvise ostinato accompaniments for known works and compose short, simple structured

instrumental or vocal works. For example, they may improvise and compose tunes on tuned percussion instruments, recorder or keyboard over relevant given chords such as a 12 bar blues bass or simple repeated chord progression.

Students can now aurally recognise and describe musical features such as simple rhythmic and melodic patterns, tempo and structure of music. They use and interpret signs and symbols representing pitch, duration of sound and dynamics (Education Department of Western Australia, p.18).

The statements for the two strands at Level 4 are:

Uses creative problem solving to explore and develop ideas for individual and group solutions to given tasks; and

Selects and manipulates a range of skills, techniques, technologies and processes to demonstrate and share solutions to given tasks (Education Department of Western Australia, 1996, p.1).

Students at Level 4 can creatively solve problems both individually and in groups, selecting from a range of skills and techniques. They manipulate these skills and access appropriate technologies to complete given tasks. They show clearly developed ideas as they engage in musical activities and complete structured musical works. They demonstrate that they carefully choose, combine and manipulate more than one element in given tasks using problem solving skills to achieve certain effects. They require less teacher assistance and direction as they experiment and identify decision making points in their creation and re-creation of music works. They consider the purpose and needs of an audience and present their works to suit these needs.

At this level, students' musical works capture characteristic qualities of a piece and interpret elements of pitch, rhythm, dynamics and phrasing in composition. They explore major and minor tonalities, textures, forms and mediums in composing and performing short instrumental or vocal works. In their invention, they explore combinations of sounds from the environment, chords, ostinati, and incorporate known structures such as ternary or binary form. At this level, students can add harmonic or rhythmic accompaniments to known songs and instrumental works with a simple harmonic structure (Education Department of Western Australia, 1996, p.22).

The statements for the two strands at Level 5 are:

Uses creativity and originality to explore and develop ideas for individual and group solutions to given tasks drawing on links to the arts of different times and places; and

Chooses appropriate arts skills, techniques, technologies and processes to shape and share meaning through arts activities (Education Department of Western Australia, 1996, p.1).

At Level 5, students use processes such as observation and research to explore their ideas. They reflect on their experiences of music of different times and places, using them where appropriate. They confidently plan their musical works, using critical processes in selecting and using skills, techniques and technologies to communicate meaning. At this level, students are discriminating in their selection of music skills, techniques, technologies and processes, even though they are still working through given tasks. They show proficiency in problem solving skills and can draw on links to music of different times and places, as well as their personal experiences, in the creation of their works. They share meaning through musical activities.

Students at Level 5 employ a variety of structural devices such as repetition, variation and contrast within their original compositions. They are able to demonstrate through performance of a known or original work, the ability to control basic expressive qualities of sound and an understanding of simple musical styles. Students can use conventional notation to notate short melodic and rhythmic patterns heard in a musical context and they interpret pitch, rhythm, dynamics, phrasing of music with minimal guidance in preparing a work for performance. They plan, rehearse and perform musical works as a member of an ensemble, demonstrating the ability to work sensitively while performing as a leader or group performer (Education Department of Western Australia, 1996, p.25).

The statements for the two strands at Level 6 are:

Draws on a wide range of ideas, contexts and past and contemporary practice in arts activities to create and re-create arts works; and

Structures arts works by applying skills, techniques, technologies and processes to specific styles and forms (Education Department of Western Australia, 1996, p.1).

At Level 6, students clearly demonstrate an understanding of musical contexts and use skills, techniques, technologies and processes to structure their musical works in specific styles or forms. They select and develop ideas to create and re-create musical works through individual and/or group exploration. They set goals and use critical problem solving, critical and social processes to develop personal solutions in their work. Students' musical statements are more personal and evolve through individual and group exploration of ideas.

Students plan and present musical works to different audiences via critical creative processes and can record their work in the form of a score which can be re-created by others. In their compositions, students can use

rhythmic and melodic features from different contexts such as African or Asian cultural rhythms. They can also sight read and interpret musical notation with accuracy in the areas of pitch, rhythm, dynamics, phrasing and overall structure and have the ability to listen to music and follow the score, identifying important musical details.

Students need little guidance in stylistic interpretation to compose and notate songs or accompaniments for known works. They can choose a medium and perform a repertoire of works that require good control of the expressive qualities of sound and an understanding of musical structures and styles. They can perform as a member of a group, demonstrating the ability to maintain an independent part and blending sensitively within the ensemble as conductor, leader or group performer (Education Department of Western Australia, 1996, p.32)

The statements for the two strands at Level 7 are:

Aligns ideas with artistic purpose and articulates why and how their choices were made; and

Manages and refines skills, techniques, technologies and processes In the Arts to communicate for a specific audience or purpose (Education Department of Western Australia, 1996, p.1).

At Level 7, students deliberately choose from a diversity of ideas to make musical works. They use musical language to communicate personal ideas and interpret the ideas of others effectively for specific audiences or purpose. They are able to critically appraise the consequences of their decisions and re-organise their work in progress. Students have a repertoire of skills, techniques, technologies and processes that demonstrate the inter-relationship between technical competence and the expressive qualities of music. They are able to rationalise their choices in the creation of musical works and communicate their works effectively for a specific audience or purpose. They employ a wide variety of techniques to reflect on their works and to record their thinking and making process. They display a knowledge of the various forms of performance and use their skills and knowledge to present their works in ways that evoke intended and desired audience response.

At Level 7, students arrange existing works to suit a specified medium. They can use appropriate technology to create an original work using some musical characteristics of a work they have listened to or performed. Use of sound sources is diverse and includes electronic media. They aurally identify musical elements and describe how the use of a particular element can vary according to the social and/or cultural environment in which the work was written. Students plan and present community group performances that display a sensitive understanding of the musical style, occasion, audience and performance venue. They need little guidance in stylistic interpretation (Education Department of Western Australia, 1996, p.35).

The statements for the two strands at Level 8 are:

Selects and manipulates ideas, arts theory and practice to make Arts works that show personal commitment and control of the art form; and

Integrates technical and structural elements to control the chosen medium using a range of skills, techniques, technologies and processes (Education Department of Western Australia, 1996, p.1).

At Level 8, students integrate their understanding of theory into practice and manipulate ideas to make musical works, integrating technical and aesthetic elements imaginatively, skilfully, and in a coherent personal style. Their musical works demonstrate independence, personal commitment, discipline, and control. They are able to view their musical works with aesthetic distance. They demonstrate musical expressiveness, a sense of artistry and understanding of musical techniques appropriate to individual styles. Students can identify, distinguish and describe musical devices and techniques, such as variety, repetition, contrast, tension and resolution, when listening to music. At this level, they perform a repertoire of contrasting works, displaying control and knowledge of the expressive qualities and an understanding of the musical structures, subtle expressive features and styles. Students take an active and responsible role in selecting, rehearsing and presenting a musical program, displaying a sensitive understanding of the musical style, occasion, audience and venue (Education Department of Western Australia, 1996, 38).

THE APPRECIATING STRANDS OF *THE ARTS STUDENT OUTCOME STATEMENTS* (EDUCATION DEPARTMENT OF WESTERN AUSTRALIA, 1996)

The 'appreciating' strands require students to respond to, reflect on, and evaluate their own musical works and the work of others, using their aesthetic understanding. Students understand that music is shaped by historical, social and economic contexts and use this understanding both in their own work and when responding to the work of others. The 'appreciating' strands are, *Responding, reflecting and evaluating* and *Understanding the role of the Arts in Society*.

The statements for these strands at Level 1 are:

Responds to arts works and activities in personal ways showing interest in the response of others; and

Identifies arts experiences in their own lives (Education Department of Western Australia, 1996, p.3).

Students at Level 1 make simple personal responses to their own musical works and activities and those of others showing engagement or enjoyment. They express their opinions directly but also show acceptance of the music of others. Students are self-focused and recognise music in their own worlds and in their immediate communities. They bring with them musical experiences from a variety of backgrounds. These experiences depend on children's cultural heritage, their living environments and their individual exposure to music. This impacts upon their depth of knowledge and understandings of music and their ability to respond and reflect on their own work and the works of others (Education Department of Western Australia, 1996, p.11).

Students often display spontaneity through physical movements such as dancing, jumping or clapping to music and, frequently, they include the singing of songs in their play activities. They articulate their likes and dislikes of music, expressing their personal preferences. They recognise that music is in their everyday lives and describe, in their own words, what they hear. They identify songs such as *Happy Birthday* at birthday celebrations. Students respond to the mood of music with free movement and identify some elements of music through body movement and shapes. They identify soft and loud sounds through singing games and respond to music through writing, drawing or contributing to discussions (Education Department of Western Australia, 1996, p.12).

The statements for the sub-strands at Level 2 are:

Outlines features of their own and others' arts works and activities using simple arts terminology relating their responses to these features; and

Outlines how the Arts are used for a range of different purposes in their everyday lives and familiar other cultures (Education Department of Western Australia, 1996, p.3).

Students at Level 2 respond critically by making subjective observations about elements of their own musical works and those of others. They use a variety of simple given frameworks in making their judgements. They recognise a range of different purposes for music such as entertainment, ceremonial and advertising. They move on from the characteristically sensory responses of Level 1 by giving descriptions of content and features and making simple critical judgements, using simple language and symbols in their oral and written work and participating in guided class discussions and small group works using simple musical terms (Education Department of Western Australia, 1996, p.15).

Students' reflection upon music works includes features such as melody, harmony, instruments used, form and expression. They identify the purpose of a work and how the purpose affects the way it should be performed. For example, they realise that some Anzac day songs should be sung in a sombre manner (Education Department of Western Australia, 1996, p.16).

The statements for the sub-strands at Level 3 are:

Outlines key features of arts works and activities, giving reasons for their responses using appropriate arts terminology and critical processes;

- a. Uses their understandings of the arts in their community and other cultures and times in making and sharing their own arts activities and arts works; and
- b. Identifies the contribution of the arts and artists in their immediate community (Education Department of Western Australia, 1996, p.3).

At Level 3 students recognise and identify the important features of musical works. Reflection is facilitated by their use of appropriate terminology and critical processes which enables them to articulate their reasons for personal responses. They identify the contributions of music in communities and consciously explore using distinctive features of known arts works from other cultures and times in their expressive activities (Education Department of Western Australia, 1996, p.19).

Students at Level 3 are able to describe obvious features that assist in shaping a musical work such as repetition, form, gradual and sudden changes in dynamics and texture. They aurally recognise identified musical features used in a musical work such as rhythmic and melodic patterns, tempo, instrumentation, timbre, dynamics and structure. They listen to music from other cultures and times, such as Aboriginal music, rock n roll or classical

music, associating some of the characteristics from these styles to a particular culture, and aurally identifying selected music of different styles (Education Department of Western Australia, 1996, p.20).

The statements for the sub-strands at Level 4 are:

- a. Makes critical observations about arts works and activities using given criteria;
 - b. Recognises and accepts that different people have different points of view and personal responses;
-
- a. Recognises similarities and differences and makes links between the Arts from different times and places; and
 - b. Recognises and understands the contributions the Arts and artists make to Australian society (Education Department of Western Australia, 1996, p.3).

At Level 4 students respond to, and reflect on, their own and others' musical works using a given set of criteria such as structured questioning. They recognise and accept others' views and opinions, as well as similarities and differences in music works of different times, cultures and places. They show awareness and understanding of the role of music and artists in Australian society (Education Department of Western Australia, 1996, p.23).

Students use critical frameworks through identified criteria for responding to musical experiences, accepting that their opinions and observations may not be shared by others and respecting the responses of others. They make comparisons and connections between music of different cultures, places and times, identifying music in Australian society and its contribution to economic growth and development (Education Department of Western Australia, 1996, p.24).

The statements for the sub-strands at Level 5 are:

- Uses arts terminology and critical frameworks to analyse and express informed opinions about arts works and activities;
- a. Identifies and discusses distinguishing features of arts works which locate them in a particular time, place or culture; and
 - b. Identifies and discusses the distinguishing features of arts works and activities in contemporary Australian society (Education Department of Western Australia, 1996, p.3).

At Level 5 students use formal critical processes such as analysis to express informed opinions about musical works and, at the same time they respect differences of opinions. They recognise, discuss and use distinguishing features of musical works that come from a particular place,

time or culture as well as those that are identifiably Australian (Education Department of Western Australia, 1996, p.27).

Students are able to locate information about aspects of a musical work from a variety of sources and are able to use this information in their critical analysis, independently. They can formulate a set of critical questions and give broad general responses that indicate links across musical forms. They show some understanding of the nature of music and its uses in particular societies or different cultures, in particular recognising distinguishing features that make them identifiably local, national and international (Education Department of Western Australia, 1996, p.27).

Students are able to listen to a performance and talk about stylistic elements that locate it in a time, place or culture. They can also discuss the musical accuracy of their own performance of a composition, recognising and accepting suggestions and opinions. Using stated criteria, they give reasons for their preferred performances.

At this level they can discuss a variety of roles played by musicians in society, identifying different interpretations. They can also discuss distinguishing musical characteristics of works they have composed and/or performed, such as rondo form or minuet and trio as played in movements of symphonies, concertos and instrumental sonatas (Education Department of Western Australia, 1996, p.28).

The statements for the sub-strands at Level 6 are:
Identifies, analyses and interprets features of arts works and activities expressing and discussing responses to them; and

- a. Shows an understanding of how the arts are shaped by particular historical, social, economical and political contexts and values and how these change over time; and
- b. Identifies career opportunities in and related to the Arts
(Education Department of Western Australia, 1996, p.3)

At level 6, students use critical processes to describe, analyse and interpret musical works, giving personal points of view and interpretations. They understand and discuss how musical works communicate ideas and both reinforce and challenge social, cultural and artistic values. Students understand the importance of historical, social, economic and political contexts and analyse, describe and interpret musical works from these perspectives. They recognise that these change over time (Education Department of Western Australia, 1996, p.31).

At this level, they discuss the manner in which identified musical elements are used to create unity and contrast or the element of surprise in works heard or performed. They listen to popular songs and discuss the harmonic and rhythmic tension and release and how successful they consider the use of such elements. They supply program notes for a work

by an influential composer, displaying knowledge of the social and historical contexts of the work and of its importance to the development of musical traditions. Students listen to and discuss and research the difference in sound quality between contemporary instruments and their predecessors such as comparing a harpsichord to a piano/keyboard, or an acoustic guitar to an electric guitar (Education Department of Western Australia, 1996, p.32).

The statements for the sub-strands at Level 7 are:

Uses processes of critical analysis to support interpretations and personal judgements about arts works and activities; and

Discusses the effect of continuity and change in local, national and international arts (Education Department of Western Australia, 1996, p.3).

Students at Level 7 use formal processes of analysis and interpretation to make judgements about their own and others' musical works and experiences. They discuss how music stays the same and how it changes over time, focusing their understanding from a Western Australian perspective, an Australian perspective and an international perspective. They understand the importance of the relationships between cultural issues and music practice, exploring these issues and influences in local, national and international contexts. They also explore historical and contemporary issues (Education Department of Western Australia, 1996, p.35).

Students at this level are able to compare and evaluate musical works from the same genre by contrasting themes. They listen to, and critically discuss, their performances and the performances of others and they express a reasoned opinion about interpretations. They aurally identify and describe the similarities and distinguishing features in works by composers working in similar cultural contexts and historical periods such as Bach and Handel. Students can identify and describe the musical features of a work that fuses two cultures or styles, such as jazz fusion or Aboriginal rock fusion.

Students discuss how contexts can change the acceptance of a musical work, such as taking a piece of concert music which has been used in an advertisement and examining the difference in the acceptance of the music. They examine aspects of the contemporary music industry in Australia, exploring ways in which musicians create and reflect social values in their music (Education Department of Western Australia, 1996, p.36).

The statements for the sub-strands at Level 8 are:

Critically reflects on meanings and values associated with particular arts works and activities; and

a. Researches arts works from a variety of contexts, understanding how histories are constructed in the arts and

how their own expression and appreciation of the arts is shaped by them; and

b. Critically examines the ways the arts challenge and shape values and are influenced by prevailing values (Education Department of Western Australia, 1996, p.3).

At level 8 students discuss how music is influenced by the values of society, culture, historical periods and national identity. They relate changes in musical expression to changing and improved instrumental technology, developments in playing techniques and/or the introduction of new sound sources, including electric and electronic instruments. At this level students analyse the role of music in influencing public opinion and they describe how social or political issues have influenced and challenged a performer's interpretation of a composition. They examine the diverse contribution music can make to the social and economic structure of a society and analyse the impact of commercialism on artistic expression. (Education Department of Western Australia, 1996, p.40).

In order to test the strands related to both *Expressing* and *Appreciating*, it was necessary to ensure that the students had the opportunity to display their understanding of music through the use of an array of symbol systems. This includes both the non-verbal language of the arts to express an idea and the most common system of language, the spoken or written symbol (Mercer & Church, 1998).

To display their knowledge and skills in the strands of *Expressing*, students had the opportunity to use both non-verbal arts language in the performance of their musical compositions, and written language to illustrate their planning and reflection. To display their knowledge and skills in the strands of *Appreciating*, it was necessary for students to receive and read the specific language of music within the stimulus, and then to translate it into written language (Mercer & Church, 1998, p.2). It must be appreciated that, while students might be constrained in their use of written language to fully interpret the subtleties of the art message, this is difficult to avoid in a testing situation.