The impact of testing on students:

Australian students’ perspectives on NAPLAN and internal assessments.

Mark William Dowley
BEd (Phys. Ed.), MEd (International Policy), PGCertEdRes
ORCID: 0000-0002-3213-1529

Submitted in partial fulfilment of the requirements for the degree of Doctor of Education

June, 2019

Melbourne Graduate School of Education

The University of Melbourne
Abstract

National and state testing policies have become an increasingly common feature of the policy landscape in education, both in developed and developing countries. Testing policies can generate a range of emotional responses among students, including high levels of stress. Alternatively, students’ emotional responses may not be discretely associated with large-scale standardised tests, but instead generalise to any testing situation. This study aimed to compare student responses and perceptions of assessment in both NAPLAN and internal tests. This study used an anonymous survey to gather data from 206 Year 7 and Year 9 Australian students on their perceptions of the importance their parents and teachers placed on doing well in tests, and their own self-reported responses to both NAPLAN and their internal tests. We found that the students in this study placed more value on internal tests than NAPLAN and students were also more likely to be confident in internal tests and bored for NAPLAN. A small percentage of students reported negative physical responses, such as crying or feeling sick to both types of tests, however, there were no significant differences between NAPLAN and internal tests in the number of students reporting negative physical responses. Furthermore, individuals who placed a high value on a given assessment and have greater emotional stability were more likely to experience positive responses to assessment. The findings suggest that NAPLAN does not cause significant negative responses in the majority of students. Implications for schools and policymakers are discussed.
Declaration

I, Mark Dowley, certify that:

(i) The thesis comprises only my original work towards the D.Ed.

(ii) Due acknowledgement has been made in the text to all other material used.

(iii) The thesis is fewer than 55,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Mark Dowley
Preface

This thesis was edited in accordance with the Australian Standards for Editing Practice as follows:

(i) Standard D: Language and Illustrations: and

(ii) Standard E: Completeness and Consistency.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship for the offset of my tuition fees.
Acknowledgements

My thanks go to Associate Professor Suzanne Rice and Associate Professor Lawrie Drysdale of the University of Melbourne, who supervised the thesis and were a pleasure to work with. Associate Professor Graham Hepworth of the University of Melbourne Statistical Consulting Centre also helped me with my statistical knowledge. I am grateful for the educators and students who made this research possible. My deepest thanks to my family who have encouraged me through all my years of education. Finally, my thanks go to Sarah for her support, patience and belief in me.
## Contents

Abstract ........................................................................................................................................... ii

Declaration ...................................................................................................................................... iii

Preface ........................................................................................................................................ iv

Acknowledgements ..................................................................................................................... v

Contents ......................................................................................................................................... vi

List of Tables .................................................................................................................................. ix

List of Abbreviations .................................................................................................................... xii

### Chapter 1: Introduction ........................................................................................................... 1

1.1 Research Questions .................................................................................................................. 2

1.2 Key Terms ................................................................................................................................ 2

1.2.1 Assessment .......................................................................................................................... 2

1.2.2 NAPLAN ............................................................................................................................... 2

1.2.3 Internal assessments .............................................................................................................. 3

1.2.4 Value ..................................................................................................................................... 3

1.2.5 Expectancy ........................................................................................................................... 3

1.2.6 Motivation ............................................................................................................................ 4

1.2.7 Academic Self-Concept ........................................................................................................ 4

1.2.8 Responses ............................................................................................................................ 5

1.3 The Importance of the Study .................................................................................................... 5

1.3.1 Limitations of the study ...................................................................................................... 7

1.3.2 Summary of the Chapters .................................................................................................. 8

### Chapter 2: Literature Review .................................................................................................. 9

2.1 Introduction ............................................................................................................................. 9

2.2 Global Education .................................................................................................................... 9

2.2.1 From knowledge to accountability ..................................................................................... 10

2.3 NAPLAN as a High-Stakes Test ............................................................................................ 13

2.3.1 Benefits of high-stakes testing ........................................................................................... 15

2.3.2 Limitations of High Stakes Testing ................................................................................... 16

2.3.3 Negative impacts of high stakes testing programs ........................................................... 21

2.4 Impact on Curriculum ............................................................................................................ 22

2.4.1 Narrowing of curriculum .................................................................................................... 23

2.4.2 Decreased time in other subjects ..................................................................................... 24

2.4.3 Potential positive curriculum changes ................................................................................ 26

2.5 Impact on Teaching ............................................................................................................... 27

2.5.1 Negative impacts on teaching ........................................................................................... 28

2.5.2 Intended benefits for teaching ........................................................................................... 30

2.6 Impact on Students ............................................................................................................... 31

2.6.1 Potential equity issues ......................................................................................................... 32

2.6.2 The intrapersonal level ....................................................................................................... 33

2.7 The Stakeholders and the Focus of the Literature .................................................................. 35

2.7.1 The students’ perspective ................................................................................................. 37

2.8 Conclusion ............................................................................................................................... 41
Chapter 3: Methodology and Sampling

3.1 Introduction ................................................................. 44
3.1.1 Background on the Australian School System .................. 44
3.2 Sampling: Using Students as the Critical Reference Group .......... 46
3.2.1 Participant recruitment and administration of the survey ........... 48
3.2.2 Sample ........................................................................... 48
3.2.3 Ethics ............................................................................ 53
3.3 The Study Design .............................................................. 53
3.3.1 The instrument .............................................................. 54
3.3.1.1 The construct and rationale for questionnaire items ............. 55
3.3.2 Pilot testing ................................................................. 58
3.4 Data Analysis ................................................................. 58
3.4.1 Sample limitations ......................................................... 59

Chapter 4: Student Perceptions of Assessment: NAPLAN and Internal Tests ........... 61
4.1 Student Perceptions of NAPLAN ........................................... 61
4.1.1 Student responses to NAPLAN ........................................ 61
4.1.1.1 Negative emotional responses: Feeling bored, nervous or stressed .... 63
4.1.1.2 Positive emotional responses: Confident, enthusiastic and excited .... 64
4.1.1.3 Negative physical responses .......................................... 64
4.1.1.4 Open-ended responses ................................................. 64
4.1.2 Student expectations of performance in NAPLAN ............... 66
4.1.3 Value of doing well in NAPLAN ....................................... 67
4.1.4 Motivation for NAPLAN tests ....................................... 69
4.1.5 Feedback from NAPLAN ............................................... 71
4.1.5.1 Uses of feedback ....................................................... 72
4.1.6 Conclusions about student responses to NAPLAN ............. 73
4.2 Student perceptions of internal testing ................................... 74
4.2.1 Student responses to internal testing ................................... 74
4.2.1.1 Open-ended responses ............................................... 76
4.2.2 Student expectations of performance in internal tests ............ 76
4.2.3 Value placed on doing well in internal tests ....................... 77
4.2.4 Student motivation for internal tests ................................ 79
4.2.5 Feedback from internal tests .......................................... 79
4.3 NAPLAN and Its Impact on School Experiences .................... 81
4.4 Summary ................................................................. 87

Chapter 5: Comparing Responses to NAPLAN and Internal Tests ......................... 89
5.1 A Comparison of Student Responses to NAPLAN and Internal Tests ........ 89
5.1.1 Differences in open-ended responses .................................. 91
5.2 Comparing Value in NAPLAN and Internal Tests ................... 92
5.3 Comparing Expectations in NAPLAN and Internal Tests ............ 96
5.4 Comparing Feedback from NAPLAN and Internal Tests ............ 100
5.5 Summary ................................................................. 103

Chapter 6: The Relationships between Expectancy, Value and Student Responses to Assessment .......................................................... 105
6.1 Expectancy and Responses to Assessment ................................ 106
6.1.1 Student Expectancy and Responses to NAPLAN .................. 106
6.1.2 Student Expectancy and Responses to Internal Tests ............ 106
6.2 Student Value and Responses to Assessment .......................... 107
6.2.1 The Relationship between Student Value and Responses to NAPLAN .... 108
6.2.2 The Relationship between Student Value and Responses to Internal Tests
6.2.3 Relationship between Parent, Teacher and Student Value of NAPLAN
6.2.4 Relationship between Parent, Teacher and Student Value of Internal Tests
6.2.5 Academic self-concept and expectancy in NAPLAN
6.3 Comparing responses between students situated low and high on value, motivation and academic self-concept
6.3.1 Comparing low and high value groups and responses to NAPLAN
6.3.2 Differences between groups: Internal Tests
6.4 Motivation
6.4.1 Motivation and responses to NAPLAN
6.4.2 Motivation and responses to internal testing
6.5 Academic Self-Concept
6.5.1 High and Low Academic self-concept groups and responses to NAPLAN
6.5.2 High and Low Academic self-concept and responses to internal tests
6.6 Gender
6.7 Year Level Differences
6.7.1 Year level and differences in value of NAPLAN
6.7.2 Year level and differences in value of internal tests
6.7.3 Year level and differences in Expectancy for NAPLAN and Internal Tests
6.7.4 Year level and expectancy for internal tests
6.7.5 Year level and motivation for NAPLAN
6.7.6 Year level and motivation for internal tests
6.7.7 Differences in Value for NAPLAN and Internal Testing within Year Levels
6.7.8 Year level and responses
6.8 Personality
6.9 Summary

Chapter 7: Discussion of Findings, Implications for Policy, and Further Research
7.1 Responses to Assessment in the Australian Context
7.1.1 Research question 1: How do students experience and respond to NAPLAN?
7.1.2 Research question 2: What differences are there in students’ responses to internal tests and to NAPLAN?
7.1.3 Research question 3: To what extent do responses to these two types of assessment differ by year level, academic self-concept and the value placed on the assessment?
7.1.4 Research question 4: What are the relationships between student perceptions of teacher and parent value placed on doing well in the tests, and student responses to assessment?
7.2 Conclusion
7.2.1 Limitations and Implications for Further Research
7.2.2 Implications for policy

References

Appendix A: Ethics Approval
Appendix B: Plain Language Statement
Appendix C: Consent Forms—Principals and Parents
Appendix D: Survey Instrument
List of Tables

Table 1 Characteristics of Schools Participating in the Study ..................................................49
Table 2 Student Participation by Gender ..................................................................................50
Table 3 Student Participation by Year Level ............................................................................51
Table 4 Student Participation by Sector ..................................................................................51
Table 5 Student Participation by Language Background .........................................................52
Table 6 Student Reported Responses to NAPLAN ..................................................................63
Table 7 Thematic Categorisation of the Responses to the Open-Ended Section of the Survey .........................................................................................................................65
Table 8 Student expectations for performing well in NAPLAN .............................................66
Table 9 Student perceptions of the value placed on NAPLAN by stakeholder .....................68
Table 10 Student motivation for NAPLAN ..............................................................................70
Table 11 Student perceptions of the source of NAPLAN feedback ......................................71
Table 12 Student perceptions of the feedback about NAPLAN from Parents. Students were able to select more than one answer. ..........................................................72
Table 13 Student perceptions of the feedback about NAPLAN from Teachers. Students were able to select more than one answer. ........................................................................73
Table 14 Student responses to internal tests ..........................................................................75
Table 15 Student Expectations of Performance on Internal Mathematics and English Tests ..........................................................................................................................77
Table 16 How important is it to you(r) [teachers/parents/students] for you to do well in your internal tests? ........................................................................................................78
Table 17 Student motivation for internal English/Mathematics Tests ..................................79
Table 18 Student Perceptions of Feedback on Internal Tests from Parents ...................80
Table 19 Student Perceptions of Feedback on Internal Tests from Teachers .........................81
Table 20 Student perceptions of the value their school places on doing well in NAPLAN ...82
Table 21 Student Perceptions of Practice Tests for NAPLAN completed at School ..........83
Table 22 Number of Students who have Completed NAPLAN Workbooks at Home ....83
Table 23 Number of classes cancelled to practise for NAPLAN ........................................84
Table 24 Student understanding of the uses of NAPLAN in their school ..........................85
Table 25 Student Perceptions of the Use of Internal Test Results .......................................85
Table 26 Student perceptions of the importance of assessment tests ................................86
Table 27 Students’ Reported Responses to NAPLAN and Internal Tests .................................90
Table 28 Differences in Open-Ended Responses to NAPLAN and Internal tests. Percentages of the entire sample. ..................................................................................92
Table 29 Students’ Perception of Teacher Value on NAPLAN Compared to Internal Tests ..93
Table 61 Student Motivation for Internal English Tests by Year Level .............................. 141
Table 62 Student Motivation for Internal Mathematics by Year Level .............................. 142
Table 63 Comparing Student Value of NAPLAN and Internal Tests in Year 7 Students ...... 143
Table 64 Comparing Student Value of NAPLAN and Internal Tests in Year 9 Students ...... 144
Table 65 Student Responses to NAPLAN by Year Level ...................................................... 145
Table 66 Student Responses to Internal Tests by Year Level ............................................... 146
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>ACARA</td>
<td>Australian Curriculum, Assessment and Reporting Authority</td>
</tr>
<tr>
<td>ICSEA</td>
<td>Index of Community Socio-Educational Advantage</td>
</tr>
<tr>
<td>LBOTE</td>
<td>Language Background Other Than English</td>
</tr>
<tr>
<td>NAPLAN</td>
<td>National Assessment Program—Literacy and Numeracy</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>PISA</td>
<td>Programme for International Student Assessment</td>
</tr>
<tr>
<td>SES</td>
<td>socio-economic status</td>
</tr>
<tr>
<td>TIPI</td>
<td>Ten Item Personality Inventory</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

This chapter outlines the aim and purpose of the study. It describes the rationale, significance and methodology. The limitations and delimitations are outlined. A brief outline of each chapter and the structure of the thesis is provided at the conclusion.

Each year, all Australian school students in years 3, 5, 7 and 9 undergo testing as part of the National Assessment Program—Literacy And Numeracy (NAPLAN). This study explores how students experience this assessment, and the broader impact of NAPLAN on their experiences of school.

NAPLAN is a contentious issue in Australian schools. The debate about the value and impact of the program reaches a peak twice a year. First, in May, when students complete the tests and again, later in the year when the results are released. The media report sensationalised stories about the negative impact on students and argue that the tests reveal an underperforming education system. Sound bites of politicians punctuate these stories, either claiming success or blaming someone else (often teachers) for the outcomes on the national tests.

This study aims to understand how NAPLAN affects one of the key stakeholders in education: the students. It explores how the NAPLAN program affects the students as they prepare and complete the tasks. This research aims to value the student’s perspective, by measuring student perceptions of how NAPLAN affects them and their lives in school. The study also compares student reactions to NAPLAN with their reactions to the internal assessments students complete and explores differences between different groups of students.
1.1 Research Questions

There are four key research questions guiding this study:

1. How do students experience and respond to NAPLAN?
2. What differences are there in students’ responses to internal tests and to NAPLAN?
3. To what extent do responses to these two types of assessment differ by year level, academic self-concept and the value placed on the assessment?
4. What are the relationships between student perceptions of teacher and parent value placed on doing well in the tests, and student responses to assessment?

1.2 Key Terms

In this section, key terms are defined to provide a scope to the boundaries of this research.

1.2.1 Assessment

Griffin (2017) defines assessment as the search for evidence of learning. There are three core ideas in this definition, it is a search, it focuses on evidence and the rationale of assessment is to measure or promote learning (Griffin, 2017). Assessment includes the tools teachers use to gather evidence and measure student learning. Within the broader category of assessment, there are standardised tests, a test with uniform questions, administration and scoring (Koretz, 2008). In this study, the assessment under consideration is internal test forms developed and administered by a school, and the national standardised tests that comprise NAPLAN.

1.2.2 NAPLAN

NAPLAN comprises five tests completed over three days, by all Year 3, 5, 7 and 9 students in Australia. The tests measure skills in four domains: reading, writing, numeracy and language conventions (spelling, grammar and punctuation). The content of each test is
informed by the Australian Curriculum, which sets out the expectations for what all young Australians should be taught (Australian Curriculum, Assessment and Reporting Authority, [ACARA], 2018). Questions are multiple choice or require a short written response. The exception is the writing test, in which students are expected to write a sustained text (ACARA, 2018).

1.2.3 **Internal assessments**

Internal assessments refer to formal assessment tasks conducted in school that are not NAPLAN, that are used for the purposes of assessing student learning by teachers, and are understood to be used in this way by students. These include the common teacher-devised tests that occur during the school year. Examples include a Year 7 fractions test, an oral presentation or a writing task, such as an essay completed under test conditions.

1.2.4 **Value**

This study draws on expectancy value theory (Eccles, 1983). The construct of value comprises attainment value, intrinsic value and utility value. Eccles (1983) defined attainment value as the importance of doing well on a given task. Intrinsic value is the enjoyment one gains from doing the task (Deci & Ryan, 1985), while utility value or usefulness refers to how a task serves a purpose for the individual’s future needs (Deci & Ryan, 1985). For this study, I limit my discussion of value to attainment value, or the importance of doing well on an assessment task.

1.2.5 **Expectancy**

Expectancy is a future-oriented belief about how well an individual will complete upcoming tasks (Eccles, 1983). Expectancy of success is similar to self-efficacy in social-cognitive theories of motivation or competence in self-determination theory (Cook & Artino, 2016). It is a combination of ability beliefs and expectancies for success. Ability beliefs focus on present ability, while expectancies focus on the future. These beliefs are shaped by an
individual’s goals, self-concept and their perceptions of task difficulty (Cook & Artino, 2016). I define expectancy as students’ beliefs about how well they will complete upcoming tasks—in this case, internal tests and NAPLAN. For this sample, expectancy was measured by asking students how well they expected to perform in their assessments.

1.2.6 Motivation

Motivation is a complex construct and can be understood through the many contemporary motivation theories. These include motivation attribution theory (Weiner, 2000), social cognitive theory (Bandura, 1997; Schunk, Meece & Pintrich, 2008; Zimmerman & Cleary, 2006), goal orientation (Dweck, 2013) and self-determination theory (Deci and Ryan, 1985). In expectancy-value theory, motivation is defined as the persistence and vigour in completing tasks (Eccles & Wigfield, 2002). This is similar to Skinner, Kindermann and Furrer (2009) who conceptualise motivation as effort exertion and persistence. In this study, I defined motivation as the self-reported effort students applied to performing well on the assessment tasks.

1.2.7 Academic Self-Concept

Academic self-concept is an individual’s perceptions about themselves in achievement situations (Bong & Skaalvik, 2003). It consists of fairly stable perceptions of the self that are past-oriented and are used to predict outcomes including motivation, emotion and performance (Bong & Skaalvik, 2003). Academic self-concept differs from other similar constructs because it is more general. For example, self-efficacy in self-determination theory, or expectancy in expectancy-value theory is more context specific and varies by task (Cook & Artino, 2016). Assessing academic self-concept relies heavily on social comparative information and the most commonly used method for measuring academic self-concept is self-reports (Marsh, 1999). In this study, students were asked about their perceptions of their
relative achievement in each subject using a variation on the Academic Self-Description Questionnaire (Marsh, 1999).

1.2.8 Responses

Research into high-stakes testing found that students experience a variety of responses. These responses include increased anxiety and stress (Brown et al., 2004; Lewis, 2000), negative physical responses, including freezing during the test (Paris & McEvoy, 2000), an inability to sleep the night before (Flores & Clark, 2003) and crying (Madaus et al., 2009). For this study, responses were the self-reported reactions students had to the assessment tests. Based on previous research into parent perceptions of student responses (Wyn et al., 2014) we assessed emotional responses (such as feeling nervous), attitudinal responses (such as feeling confident) and physical responses (such as crying).

1.3 The Importance of the Study

Achieving high educational standards is regarded as essential to global competitiveness (Lingard, 2010). The fear of becoming internationally uncompetitive has been used as the justification for countries to reform education systems. The Australian Federal Government and other stakeholders express concern that their education systems are failing to prepare students for the workplace. In Australia, these fears—combined with declining performance in international assessments, such as the Programme for International Student Assessment (PISA)— have led to several changes, including an increase in policy borrowing, a focus on the economic purposes of education, an increase in accountability measures and the implementation of high-stakes testing programs (Lingard, 2010). The proposed benefits of testing programs include the ability to measure the effectiveness of interventions (Alexander, 2010), increase consistency between schools and states (Clarke, Shore, Rhoades, Abrams, Miao & Li, 2003), improve efficiencies and solve the problem of the unaccountable teacher by making them accountable for student performance data.
(Lingard, Thompson & Sellar, 2016). These proposed advantages been used to justify the establishment of high-stakes testing programs internationally, but it has yet to be determined if the perceived benefits outweigh the unintended negative consequences.

International research has demonstrated that high-stakes testing influences the enacted curriculum (Clarke et al., 2003; Jones 2007). High-stakes testing policies also lead to changes in pedagogy (Klenowski, 2011), as teachers retreat to rote learning practices to prepare students for testing, rather than investing in innovative deep learning practices. Further, it is argued that high stakes testing has affected student wellbeing in many ways (Rice, Dulfer, Polesel & O’Hanlon, 2015; Rogers, Barblett and Robinson, 2016). For example, high-stakes testing may increase dropout rates (Lewis, 2000), have negative impacts on disadvantaged students, and be used as a biased sorting function for educational advancement (Moon, Brighton, Jarvis & Hall, 2007). Additionally, high-stakes testing can contribute to a lack of student engagement if teachers present material in repetitive ways (Darling-Hammond, 2010). The introduction of high-stakes testing has also been linked to increased anxiety and stress among students (Darling-Hammond, 2010).

However, many of these negative consequences have been reported by teachers on behalf of the students. How accurate these perceptions are, and whether they are shared by the students themselves is unknown. Further, it is unclear whether some of the negative emotional student responses reported in regard to high stakes testing are to do with the testing program, or are a function of any situation in which students are formally assessed. For example, research into the impact of NAPLAN on middle school students concluded that more research is required to “ensure that the negative experiences reported by some participants are directly related to NAPLAN preparation and testing and not generally experienced test anxiety.” (Swain, Pendergast & Cumming, 2018, p. 338).
The current study sought to determine from students’ perspectives how they experience and respond to NAPLAN. It also sought to determine whether student responses were particular to NAPLAN or were more general responses to testing, and to understand more clearly relationships between responses and students’ values and expectancies. To achieve this, 206 students across four schools in the state of Victoria, Australia were surveyed. Their expectations, motivation for, and valuing of NAPLAN and internal tests were examined in conjunction with their emotional responses to the tests.

The findings suggest that attitudes and responses to assessment vary between students. The students in this study placed more value on internal tests than NAPLAN. The students were also more likely to be confident in internal tests and bored for NAPLAN, and a small percentage of students reported negative physical responses such as crying or feeling sick to both types of tests. Individuals with higher levels of academic self-concept and greater emotional stability were more likely to experience positive responses to assessment. This additional knowledge of the impact of NAPLAN on students can potentially inform the policy debate about the purpose and consequences of high stakes testing programs.

1.3.1 Limitations of the study

The findings of this study revealed the perceptions of 206 students from four secondary schools in the state of Victoria. The schools represented are two government and two Catholic. They represent metropolitan, rural and regional areas. One of the Catholic schools is a boys’ school and the other three are co-educational.

The students completed a survey where they were asked their perceptions of the NAPLAN ‘pen and paper’ tests and internal testing, and perceptions about the value their parents and teachers placed on both forms of assessment. This study was a snapshot of student
perceptions of assessment and the general responses students experience when completing assessment tests.

1.3.2 Summary of the Chapters

Chapter One covers the purpose, background and significance of the thesis. The study is placed in the context of a competitive environment where many countries are engaging in high-stakes testing to raise student achievement. In Australia a national testing program was introduced for Year levels 3, 5, 7, and 9. This thesis investigates student perceptions of both the national test (NAPAN) and internal assessment processes.

Chapter Two discusses the related theory, literature and research on testing to date, and their relationship with the current study.

Chapter Three describes the quantitative methodology for the gathering and analysis of data and the sample for the study.

Chapters Four, Five and Six report the results from the survey. Chapter Four describes student perceptions of NAPLAN and internal tests, Chapter Five compares responses between the two types of assessment and Chapter Six describes the relationships between expectancy, value and student responses to assessment.

Chapter Seven summaries and discusses the findings in relation to the literature and highlights the key findings. The contribution of the findings is outlined and areas for future research are identified.
Chapter 2: Literature Review

2.1 Introduction

This study investigated an important educational issue: national testing in schools. Across many developed countries, there have been political shifts towards neoliberal policy actions. These policy actions include a focus on measurable outcomes, competition and choice as tools for improvement in education (Levin & Fullan, 2008). This has fostered a culture of testing and assessment and the consequential implementation of national assessment programs (Lingard, 2010).

2.2 Global Education

Australia has borrowed educational reforms from England and the United States. This has led to an increased use of market mechanisms in Australian education. Underpinning this direction in education is the assumption that competition between schools and increased pressure from governments will raise the standards of educational achievement (Levin & Fullan, 2008; Lingard, 2010). This competition is fostered through data collection and increased management by governing authorities (Ozga, 2009). Lingard (2010) notes that this shift in focus from inputs and processes to outputs and outcomes has resulted in increased pressure on educational institutions and educators. Educational outcomes—whether at the school, system or student level—are being identified, measured and made public (Levin & Fullan, 2008; Lingard, 2010). Lingard (2010) argues that, what is easily measured in outcomes rarely encompasses the full breadth of schooling aims: the social, physical, emotional, spiritual, environmental and political needs of students.

Current educational reforms prioritise curriculum development, proficiency in basic competencies (i.e., literacy and numeracy) and measurement of student assessment outcomes (Sahlberg, 2011). This measurement of outcomes, usually by standardised tests, has led to
what Lingard (2010) referred to as “governing by numbers”. Internationally, these numbers are produced by large-scale comparative assessments, such as the Organisation for Economic Co-operation and Development’s (OECD) PISA, the International Association for the Evaluation of Educational Achievement’s Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS).

As education is central to global competitiveness, the fear of becoming internationally uncompetitive has been used as the justification for the reform of education systems (Lingard, 2010). For example, the American education system was restructured early in the 21st century on the assumption that the public schools were in dire need of comprehensive reform (Moon et al., 2007). The concern was that education systems were failing to prepare students for the workplace and thus, preventing Americans from competing internationally.

It could be argued that in Australia, these fears—intensified by Australia’s declining performance in international assessments such as PISA—have influenced national and state educational policy decisions. Policy trends and actions include an increased focus on the economic purposes of education, an increase in accountability measures and the implementation of high-stakes testing programs.

2.2.1 From knowledge to accountability

Traditionally, testing has been used as a tool to identify student knowledge and learning. Students with strengths in desired areas—such as literacy and numeracy—were rewarded for their performance through either better employment or further educational opportunities. For example, as early as 165 BCE, China used a system of recommendation and examinations to appoint government officials (Gregory & Clarke, 2003). In the early 1900s, Cambridge University also used examinations. First, these tests determined who would be allowed entry to the University; subsequent tests acted as a summative measurement of mastery of particular skills at the completion of a course (Koretz, 2008). The
primary aim of testing was to assess student understanding; however, high-stakes testing arguably shifts the focus of testing from determining individual students’ levels of knowledge to school and teacher accountability for student results.

Within accountability policy frameworks, student achievement data are used as an indicator of educational quality. Key performance indicators—that is, test results—usually in literacy and numeracy are used as a proxy for the educational effectiveness of schools and teachers. These measures of effectiveness, conducted through national assessment programs, are used to make judgements about how well schools and teachers are carrying out their functions (Polesel et al., 2012).

Clearly, there is a need for some forms of accountability, given the impact that quality of education has on both societies and individuals. Education is an investment in human capital, which not only plays a critical role in economic performance, but also brings key individual and social benefits, such as better health, improved wellbeing and increased social and political engagement (Rychen & Salganik, 2003). The OECD (2014a) has shown that the literacy and numeracy skills students achieve by Year 9 will considerably affect their life outcomes. For these reasons, it is important that systems, schools and teachers are accountable for student learning. In Australia, the education system is held accountable through the National Assessment Program (NAP).

The movement towards the current National Assessment Program had its origins in the changes to the English education system under Margaret Thatcher in the 1980s. England’s Education Reform Act of 1988 established a national curriculum and arranged for the publication of league tables to compare school performances (Gregory & Clarke, 2003). The British Government argued that this would provide increased accountability for educational institutions and better control over teachers and schools. Testing of literacy and numeracy began in New South Wales, Australia in 1990. Over the next 10 years, all states and
territories introduced testing programs, in part to establish minimum literacy and numeracy standards. These standards served a dual purpose of being a diagnostic test to identify students at risk of not meeting literacy and numeracy standards and also promote consistency in literacy and numeracy standards across the country (Education and Employment Reference Committee, 2014). In 2006, the Australian Government (a Liberal–National Coalition) began to trial national tests under the leadership of Minister for Education Brendan Nelson. In 2007, a Labor government was elected and inherited the program. It grew under the Minister for Education and future prime minister, Julia Gillard. The first nationwide set of tests, NAPLAN, was conducted in 2008 for all Year 3, 5, 7 and 9 students. NAPLAN measures students’ proficiency in reading, writing, language conventions (spelling, grammar and punctuation), and numeracy through five separate tests, administered across three days.

NAPLAN is part of the larger National Assessment Program, which covers all testing endorsed by the Ministerial Council for Education, Early Childhood Development and Youth Affairs (2011). In addition to the population testing conducted through NAPLAN, the National Assessment Program includes sample testing in domains such as information and communications technology knowledge and incorporates participation in international sample-based tests, such as PISA.

In 2010, the responsibility for the development and administration of NAPLAN was transferred to the Australian Curriculum, Assessment and Reporting Authority (ACARA). A further development occurred the same year with the publication of schools’ results on the newly-established My School website (ACARA, 2011). My School presents a range of information about each school, including enrolments, the school’s Index of Community Socio-Educational Advantage (ICSEA) value, school income and expenditure, and NAPLAN scores. Each school’s results are displayed against 59 similar schools in a format that was designed to be easily read by the public (Gorur, 2015).
At the time, the Labor government argued that the *My School* website provided extensive information on Australian schools and introduced transparency and accountability into the Australian school system (ACARA, 2010). The publication of NAPLAN results was designed to support parents’ access to information about the quality of schooling and assist their decision-making in choosing schools for their children. These results became the primary measure of performance for individual schools and education systems more broadly (Klenowski & Wyatt-Smith, 2012). The Labor government, through ACARA, used the national educational assessments to make educators accountable to the government and the public. As a result, schools were judged on their published results. While the *My School* website was not designed to facilitate the construction of league tables, media outlets quickly developed school ranking or league tables following the release of NAPLAN results (Gorur, 2015). At the time of writing, the Liberal–National Coalition government (2014–present) has continued with both NAPLAN and the MySchool website.

### 2.3 NAPLAN as a High-Stakes Test

In the Australian context, it is important to consider whether NAPLAN constitutes a high-stakes testing program. In the United States, Johnson, Johnson, Farenga and Ness (2008) define high-stakes tests as those that have consequences for student success, teacher accountability, the reputation of schools or the funding of schools. Through the *My School* website—and its use by schools, media and the public—it can be argued that NAPLAN has consequences for the reputation of schools. The public reporting and interpretation of NAPLAN results support the argument that NAPLAN is a high stakes test as poor NAPLAN results could change the public perception of a school’s quality. It is plausible that these results could lead to the following chain of events. Poor NAPLAN results would be publicly available of the myschool web-site, this could lead to parents moving their child to another school, causing a drop in enrolments, this would impact the number of teachers required at
the school and eventuate with teacher redundancies. It is not inconceivable that this could begin a negatively reinforcing cycle and have long term implications for the sustainability of the school. In this scenario, NAPLAN could be considered high stakes for school leaders.

A subsequent effect could be increased pressure from school leaders for their teacher to improve NAPLAN results. This increases the stakes for individual teachers as the NAPLAN results of their students may impact their leader’s perception of their effectiveness. The stakes attached to NAPLAN are far lower than other accountability measures such as school reviews or OFSTED visits, there are moderate, indirect stakes for teachers.

At the student level, NAPLAN is generally low stakes as there are no direct consequences on the basis of test results and the tests are not used as a hurdle for progression to the next year level. The stakes may be higher if schools use NAPLAN data for entry to select schools or to determine positions in extension classes.

These consequences potentially filter down through school leaders and teachers. Researchers from Australia have concluded that the impact of NAPLAN on stakeholders in educational systems meets the criteria of Johnson et al. (2008) and allows NAPLAN to be classified as a high-stakes test (Dulfer, Polesel & Rice, 2012; Lingard, 2010; Lobascher, 2011). Despite being a high-stakes test, there are some key differences between NAPLAN and testing programs in other countries.

The NAPLAN program varies from United States and British testing regimes in some important ways. For example, individual teachers are protected, to the extent that there are no individual reports at the teacher level, although the reports provided to schools can give data at the class level. NAPLAN also provides individual student results to parents, giving them information about their child’s progress. Polesel, Dulfer & Turnbull (2012) note that Australian schools do not face closure due to poor results and in some cases may be offered additional support if they are deemed to be underperforming. As noted, the *My School*
website also makes it difficult for individuals to create league tables of schools; however, this has not prevented the development and publication of league tables through media organisations (Mockler, 2015). The My School website also works to contextualise student results in NAPLAN by providing information on school funding and the socio-economic makeup of the school, and by aiming to compare school performance to other schools with similar student profiles.

2.3.1 **Benefits of high-stakes testing**

Research into high-stakes testing acknowledges several intended benefits. These benefits include increased information for the diagnosis of student, teacher or school achievement issues (ACARA, 2010; Phelps, 2006), and the ability to measure the effectiveness of interventions (Alexander, 2010). It has also been noted that such testing can provide students with information about their current levels of knowledge (Sloane & Kelly, 2003), send clear signals to schools about what is important in the curriculum (Amrein & Berliner, 2002), increase curriculum consistency between schools and state educational systems (Clarke et al., 2003) and improve the alignment between the taught curriculum and national standards (Alexander, 2010). Another perceived benefit of high-stakes testing is that measuring outcomes may have the capacity to improve efficiencies and solve the problem of the unaccountable teacher by holding them accountable for student performance data (Lingard et al., 2016).

Assessment programs such as NAPLAN can also provide a way for policymakers to understand how the system is performing over the longer term. As NAPLAN is a national test undertaken on a regular basis, policymakers can use longitudinal data to track changes and compare subgroups of students by gender, geographic location and school type (Wu, 2015). NAPLAN also allows students to be assessed using the same instrument (Gorur, 2015) and
gives educational stakeholders a great deal of moderately reliable information about large groups of students.

High stakes testing also allows for individual school performance to be measured. Specifically, NAPLAN is strongly shaping how schools determine student progress. Gorur (2015) presents evidence that NAPLAN results and indicators on the My School website are used by schools to assess themselves and in some states, schools are required to publish their NAPLAN results in annual reports. Such standards-driven reform (Klenowski, 2011) uses literacy and numeracy standards to judge student learning and track student progress over time. This student progress data are used as a proxy through which to judge schools’ effectiveness—and potentially teacher effectiveness—with the aim of identifying weaknesses and sharing strengths across the education sector. These benefits have been used by some proponents to justify high-stakes testing programs internationally, but it has yet to be determined if these perceived benefits overcome the limitations of testing.

2.3.2 Limitations of High Stakes Testing

The literature identifies a number of limitations of high stakes testing. The main issues highlighted include the reliability and validity of the tests, questions about sample sizes and problems with using test results in isolation (e.g. Koretz, 2008; Stobart, 2008). There are also strong and well-founded concerns, as discussed in Section 2.3.4, about using test results as a measure of teaching effectiveness (Baker et al., 2010; Wu, 2015). Any discussion regarding high-stakes testing must include a sound understanding of these issues.

There are two key concepts in understanding the uses of the results of tests such as NAPLAN: reliability and validity.

Reliability refers to the degree of the measurement error or fluctuations in scores when students take tests multiple times (Koretz, 2008). Many factors influence test reliability. For example, despite no change in students’ knowledge of content, a change in mood (such as
being tired, unwell or distracted) may alter students’ performance on a test. Other inconsistencies can arise from variations in scoring, particularly in the subjective assessment of portfolios, writing tests and essays. Electronic scoring, training of markers and use of rubrics decrease these variations, but do not eliminate them. The nature and structure of the test will influence the degree to which a test is considered reliable. NAPLAN testing instruments are fairly short, and concerns have been raised about their reliability. For schools with less than 50 students in a year level, the average school performance on NAPLAN could change significantly from one calendar year to another with no difference in underlying achievement (Wu, 2015). The margins of error are large and individual student results can vary widely in NAPLAN (Wu, 2010). If the tests are not reliable, any decisions based on the data are likely to be inaccurate.

In educational measurement, validity is the ability to draw accurate conclusions or inferences from a test score (Koretz, 2008). Validity must be considered when interpreting the NAPLAN data. For example, Wu (2010) notes that NAPLAN tests a small part of the curriculum and does not allow for an accurate indication of a student’s learning across that whole curriculum (Wu, 2010). In addition, much of the variance in educational attainment as measured by standardised tests, such as NAPLAN, can be attributed to non-educational factors. Test results are influenced by many factors outside school. Social factors, such as family and culture, have a profound influence on a student’s performance in tests. Specifically, parent level of education and occupation is highly correlated with NAPLAN scores (Barnes, 2011). Hattie (2008) has also noted that despite teacher quality being one of the strongest influences on test scores, it still only accounts for approximately 15 per cent of the variation in achievement. In sum, a significant part of student attainment is outside the influence of schools and teachers.
NAPLAN only provides an indicative level of student performance. Results are reported on a points interval scale (200–800), with all year levels represented on the same scale. For example, the mean reading score for a Year 3 student (in 2013) was 397, a Year 5 student was 486, a Year 7 student was 542 and a Year 9 student was 583. If a Year 3 student scores 400 on NAPLAN, the 90 per cent confidence interval is 78. That is, NAPLAN assessors are 90 per cent confident the student’s measured level of achievement lies in the range 361–439. The annual growth rate for a Year 3 student is 44 points, so this means the confidence interval is large enough to cover more than an entire year of progress. Thus, the error is too large to pinpoint a student on the scale (Wu, 2015).

A further problem in using NAPLAN to judge school or teacher quality is the nature of the sample size. As mentioned previously, samples from schools with small year level cohorts may have erratically varying averages. This is especially concerning when scores are used to make judgements on the quality of teaching and learning in a school. Using test scores alone, teachers can appear to vary greatly in effectiveness from year to year, without this being an accurate reflection of their true effectiveness. These limitations in measurement can raise important questions about the usefulness of testing programs such as NAPLAN, particularly if we base high-stakes decisions on the results.

A basic assumption of high-stakes testing is that high scores are a result of good teaching and learning. Many researchers have questioned the accuracy of this statement (Amrein & Berliner, 2002; Flores & Clarke, 2003; Koretz, 2008; Alexander, 2010; Baker et al., 2010; Wu, 2010). Over a decade ago, Amrein and Berliner (2002) analysed the negative impacts of high-stakes testing across 18 American states. They determined that, at the time, there was no evidence that high-stakes test results are valid indicators of the broader knowledge domains that they are intended to measure (Amrein & Berliner, 2002).
An Australian submission to the NAPLAN senate inquiry in 2010 supported these views (Wu, 2010). These findings were consistent with other studies involving teachers in Texas, which challenged the notion that high-stakes tests are valid measures of student knowledge or ability to learn new content (Flores & Clark, 2003). Australian teachers were also undecided as to the accuracy of NAPLAN results, with only 50 per cent of teachers believing it to be a useful diagnostic tool to assist teachers (Dulfer et al., 2012). This does not suggest there is no value in testing, but that test results must be used a part of an ongoing conversation regarding educational progress and most importantly, they must be considered one example of multiple pieces of evidence about student learning.

Test results taken in isolation from other student work are generally inadequate indicators of student achievement (Koretz, 2008), and caution should be exercised in the types of conclusions drawn from test results. This view was supported across school sectors in a comprehensive study of the English primary school system, The Cambridge Primary Review (Alexander, 2010). The review used a variety of approaches, such as interviews and data from official and independent organisations, and included feedback from a range of stakeholders. It questioned the validity of using tests for purposes for which they were not designed, in particular the evaluation of teacher effectiveness (Alexander, 2010). Another technical review of high-stakes testing evidence concluded: “There is broad agreement among statisticians, psychometricians and economists that student test scores alone are not sufficient, reliable and valid indications of teacher effectiveness to be used in high stakes personnel decisions.” (Baker et al., 2010, p. 2).

It is understandable that governments require information on the performance of schools. Governments use data to support future reforms and measure the effectiveness of current reforms in education (Australian College of Educators, 2010). The public also seeks access to this information to judge the government’s performance in education policy.
However, it is essential that all information is distributed to the public with the correct supplementary information regarding the limitations of the testing program, including its reliability and validity. This would hold governments accountable for ensuring policy decisions are based on appropriate use of data. Due to the importance of the decisions based on this data, education systems have a responsibility to support both the public and schools to use the data in ways that are justifiable, notwithstanding any perceived benefits of testing.

One use of testing results that is far from justifiable is as a measure of teacher quality. National testing systems lend themselves to the use of results as a means of making judgements about teacher quality, although NAPLAN results have not been used in this way to date. There is the very real danger that the notion of teacher quality may become synonymous with student achievement in large-scale tests, with the assumption that these tests can measure teacher quality. The results can be used to identify and reward supposedly high-performing teachers or remove supposedly ineffective teachers, aiming for an overall improvement of the education sector. Importantly, the tests were never designed to measure teacher performance, and a host of factors beyond the teacher’s control (such as parental education and income, class allocation, and the effectiveness of students’ teachers in the preceding year) can influence students’ achievement.

To acknowledge the variations between schools and what students bring to their schooling, some school districts in the United States (such as City of Chicago School District and the District of Columbia Public Schools) have used a ‘value-added method’ of assessment that measures the improvement of students over a year as a means of using test results to evaluate teachers (Baker et al., 2010). The value-added method has fierce critics. A specific investigation into the use of value-added measures showed large discrepancies between teachers from year to year (Baker et al., 2010). For example, in a study of teachers across five school districts, of the top 20 per cent of ranked teachers, only a third were ranked
in the top 20 per cent in the following year and another third were ranked in the bottom 40 per cent the following year (Baker et al., 2010). Wu (2015) reported that due to the relatively random allocation of students to classes the fluctuations in class averages make it possible for a high-performing teacher to have a low class mean score by chance. The Board on Testing and Assessment of the National Research Council of the National Academy of Sciences, Engineering and Medicine argued that, “value added method estimates of teacher effectiveness should not be used to make operational decisions because such estimates are far too unstable to be considered fair or reliable.” (as cited in Baker et al., 2010, p. 2)

A number of U.S. and British researchers have argued strongly against the use of test scores for teacher evaluation for these reasons (e.g. Alexander, 2010; Baker et al., 2010; Koretz, 2008). Some research has suggested there are better and more comprehensive ways to assess teacher performance. Instead of using test scores, assessors could:

- Use systematic observation protocols with well-developed, research-based criteria to examine teaching, including observations or videotapes of classroom practice, teacher interviews, and artefacts such as lesson plans, assignments, and samples of student work. Quite often, these approaches incorporate several ways of looking at student learning over time in relation to the teacher’s instruction.

(Baker et al., 2010, p. 4)

Unfortunately, these types of teacher evaluation are generally much more time-consuming and expensive than using existing student tests. To date, they have not been widely implemented in education.

2.3.3 Negative impacts of high stakes testing programs

High stakes testing programs, as noted above, can have significant limitations in terms of their reliability and validity. While there are some purported benefits to these programs,
the research literature outlines a range of negative impacts from high stakes testing programs on schools, teachers and students. These are now discussed in terms of impacts on curriculum, teaching and student wellbeing.

In Australia, some teachers view NAPLAN as an instrument of control. For instance, many teachers consider NAPLAN to be primarily an accountability tool, rather than a program that supports student learning. In one large study of teachers, the common perception of NAPLAN was that it was either a school ranking tool or a policing tool (Dulfer et al., 2012). This highlights teacher scepticism regarding the purposes of NAPLAN. Despite ACARA stating the purpose of NAPLAN is to drive improvements in student outcomes (ACARA, 2018), teachers do not share this vision. In line with the view of many teachers, critics of NAPLAN have openly attacked the program, for example Belcastro and Boon (2012, p. 3) claim:

“The move towards NAPLAN testing in Australia can be viewed as the implementation of accountability measures beneath the cloak of providing assessable and equitable education for all: a combination of both the British and American experience of national standardised testing.”

Regardless of the intentions, there is a consensus that the implementation of a national testing program has influenced the actions of educators (Alexander, 2010; Lobascher, 2011; Luke & Woods, 2008; Polesel et al., 2012). Some of these are discussed below.

2.4 Impact on Curriculum

Research indicates that high-stakes testing impacts on the curriculum in four ways: a narrowing of curriculum, a decrease in time on non-testable subjects, a focus on basic content and closer alignment between what is taught and state and national curriculum guidelines.
2.4.1 Narrowing of curriculum

Research has identified a narrowing of curriculum in Australia due to NAPLAN (Klenowski, 2015), with an increased focus on core subjects (i.e., literacy and numeracy). The findings from the large, independent, nationwide review into primary education in England, the Cambridge Primary Review (Alexander, 2010), align with the Australian findings. The introduction of the British national testing program resulted in an increased focus on literacy and numeracy in schools. This change was particularly evident in those schools identified as being low-performing (David, 2011).

Where there are high stakes associated with assessments, teachers will tend to focus on the curriculum areas in which students are tested (Baker et al., 2010; Polesel et al., 2012). The monitoring of results puts pressure on leaders to ensure improvement (Klenowski, 2015). Literacy and numeracy are the focus of NAPLAN and as a result, teachers report that they are devoting resources to these areas (Klenowski, 2011; Polesel et al., 2012). This aligns with international research (Baker, 2012). A metasynthesis of the United States research into testing found that the curriculum was found to have narrowed to testable material in 70 per cent of studies (Au, 2007). Similarly, a national survey of 4,200 teachers in the United States found that 76 per cent of the responding teachers indicated that they had increased the amount of time spent on testable material in response to testing programs (Abrams, 2004).

A small number of studies that have found that high-stakes testing has not narrowed the curriculum. Clarke et al. (2003), for example, interviewed 360 American teachers and concluded that testing regimes did not alter the relationship between curriculum standards and classroom practice. However, these findings are rare and there is a growing body of research demonstrating that testing regimes tend to narrow the teaching focus to the tested material, particularly in secondary education (Au, 2007; Klenowski, 2011; Lingard, 2010; Lobascher, 2011; Moon et al., 2007).
However, those in favour of testing programs have argued that ‘there is nothing that narrows a pupil’s experience of the curriculum so quickly as a poor preparation for the level of literacy... that the subject demands’ (The Children, Schools and Families Committee, 2008, p. 50). As stated by Luce-Kapler and Klinger (2005, p. 171), ‘High-stakes testing must be accompanied by explicit efforts to ensure the tests either support relevant educational goals, or at least, do not limit the educational domains being taught’. As curriculum content narrows towards tested subjects, it occurs to the detriment of non-tested subjects.

2.4.2 Decreased time in other subjects

The focus on literacy and numeracy assessments increases the class time spent on these areas (Brown, Galassi & Akos, 2004), and consequently reduces time spent on other areas of the curriculum. In the United States, researchers have identified reductions in the amount of time spent on sciences, arts and physical education following introduction of high-stakes testing programs (Jones, Jones & Hargrove, 2003; Lobascher, 2011). Au (2007) identified that the most prominent way in which ‘teaching to the test’ has impacted the curriculum is through a reduction in time spent on non-assessed subjects (Au, 2007). In one study in Australia, over 80 per cent of teachers reported that NAPLAN added to an already crowded curriculum and just over two thirds of teachers believed that NAPLAN had led to a timetable reduction for other subjects in their schools (Polesel et al., 2014). At the secondary level, Athanasou (2010) found a reduction of time allocated to languages, arts and social sciences due to the pressures of NAPLAN.

Tested subjects are deemed important and therefore, valuable and these subjects are given increased time in the curriculum (Gregory & Clarke, 2003; Renter et al., 2006). There are common findings regarding the demotion of arts, music and physical education programs in schools following the implementation of testing regimes in which literacy and numeracy are foregrounded (Ewing, 2010; Sabol, 2010). This is particularly concerning as these
subjects—arts, music and physical education—have been found to have a positive influence on increasing engagement among disengaged students (Ewing, 2010). Madaus, Russell and Higgins (2009) also found evidence of structural changes to the school day, with recess and lunch periods being shortened to prepare for tests. Lingard (2010) has also suggested that schools with weaker results (usually those in lower socio-economic areas), would experience these flow-on effects more dramatically than schools in higher socio-economic areas, leading to equity issues. These findings are consistent with research in Australia, as narrow assessment and reporting strategies have been found to limit students’ curriculum choice (Wyn, 2009). Australian teachers in one study overwhelmingly supported this notion, with 75 per cent of teachers believing NAPLAN influences the way school communities view different curriculum areas (Dulfer et al., 2012).

The effects on curriculum do not solely impact students. Teachers have reported feeling they cannot practise authentic pedagogies or assessment as they are forced to implement a standardised and reductionist curriculum (Ball, 2006). Teachers have also expressed concern over an inadequate amount of time to cover content, leaving them unable to investigate advanced topics or pursue areas of interest relevant to particular students or classes (Moon et al., 2007). Other findings have indicated that teachers feel forced to rush content, regardless of whether students understand, and that school leaders demand that teachers focus directly on test preparation at the expense of a holistic curriculum (Dulfer et al., 2012; Moon et al., 2007). It could be argued that there are both positive and negative impacts on teachers. At one end of the continuum, the focus on a limited and standardised curriculum may stifle effective and innovative teachers; however, it may also force ineffective teachers to meet standard curriculum goals.
2.4.3 Potential positive curriculum changes

There are also potential positive outcomes from the implementation of a national testing regime. Some researchers report that the tests encourage teachers to align curriculum with testable material, resulting in more consistent content being taught across schools (Gregory & Clarke, 2003). Students and parents may benefit from this consistency between schools as they have a clearer understanding of the performance standard expected at each stage of education (Hargreaves, 2002). This may potentially benefit students who transfer between schools, as there is no mismatch between schools in what has been taught. Testing regimes may also provide teachers with more explicit knowledge of the content they are required to teach (Clarke et al., 2003; Jones et al., 2003). There is a perception that the policy focus in the early 2000’s was more about ensuring consistency in the curriculum across the states. Since then, testing policies have focussed on student and school performance on these standardised tests. Additionally, a small number of studies in Singapore and England have found that testing had increased student-centred teaching approaches and broad content knowledge in a small number of schools (Au, 2007).

In Australia, recent research has found cases of contrasting perceptions of the impact on curriculum. Rogers, Barblett and Robinson (2018) found mixed attitudes to NAPLAN. The sample of 40 teachers reported experiencing a relatively minor impact of the testing on their curriculum. This is consistent with other students that suggest that teachers from independent schools may not experience the same impact as teachers in government or low SES schools (Thompson & Harbaugh, 2013).

However, Dulfer et al. (2012) found that a large percentage of Australian teachers spent class time practising for NAPLAN (Dulfer et al., 2012). In the two weeks prior to NAPLAN, 30 per cent of teachers reported practising three to five times per week; an additional 30 per cent practised six or more times per week (Dulfer et al., 2012). Teachers
expressed concern regarding this large amount of practice time, as it may decrease their students’ engagement. As the great majority of this practice is reported to be occurring during class time, parts of the curriculum must be omitted to facilitate NAPLAN preparation (Dulfer et al., 2012).

These findings must be viewed with caution, as ‘practice’ for NAPLAN must be carefully defined. ACARA (2011) suggested the best way to practise for NAPLAN was for students to improve their literacy and numeracy skills through their school curriculum, as the tests contain questions like those students would encounter in regular classroom assessments.

In summary, national testing programs would appear to have a large impact on the curriculum. These impacts include a narrowing of curriculum towards the testable material, decreased time on non-assessable subjects, a testing and teaching of basic content and an increased alignment between what is taught and curriculum standards (Klenowski, 2011; Polesel et al., 2012). These findings have been supported in the Australian context (Au, 2007; Polesel et al., 2012). At present, there is limited research on the direct impact of these changes on students, as much of the current data has been collected from teachers.

2.5 Impact on Teaching

As schools are held accountable for student results in high-stakes tests, the accountability pressure has flowed onto teachers. High-stakes tests can allow for the placement of schools in league tables and the judgement of schools and teachers by their results (Klenowski & Wyatt-Smith, 2012). The consequences of these published results have had a large impact on pedagogy and what constitutes effective teaching.

The impacts appear to vary between teachers and schools, with some teachers developing innovative and engaging teaching strategies to teach testable material, but others reverting to a focus on lower order thinking skills with less creative and engaging pedagogies (Jones et al., 2003). Lingard (2010) argues that this is concerning. With the move to a
knowledge economy, it is essential for education systems to continue to teach in-demand higher order thinking; further, the relationship between higher order thinking skills and improved social and economic outcomes is well-established.

Education systems must balance government-driven accountability measures and the provision of teacher autonomy. Australia appears to be following the accountability pathway of England and the United States. After nearly two decades of strict testing and assessment in England, the House of Commons Children, Schools and Families Committee (2008) inquiry into *Testing and Assessment* identified a conflict between accountability measures and high-quality teaching. This inquiry concluded that it is teaching that raises standards, not tests. The reason for this tension is a loss of pedagogical autonomy, as teachers have less freedom to individualise teaching strategies and thus, may fail to engage students (Jones et al., 2003).

A survey of teachers in Florida also found that high-stakes testing decreased creativity and imposed a standard teaching approach (Anagnostopoulos, 2003; Au, 2007; Jones & Egley, 2004; Polesel et al., 2012). For example, Au (2007) found a relationship between high-stakes testing and the teaching of content in small and isolated ‘test-size pieces’, known as fragmentation (p. 262). These pieces of knowledge are related to what is expected in high-stakes tests, rather than being related to other areas of the curriculum and the practical application of knowledge in society. The largest and most recent Australian study of nearly 1,000 teachers explored the impact of NAPLAN on pedagogy (Thompson, 2015). It found that NAPLAN had many of the same unintended consequences as high-testing programs in England and the United States, including a shift to more direct instruction and teaching to the test.

2.5.1 **Negative impacts on teaching**

Research into the effects of national testing programs on pedagogy has identified negative impacts on teaching. For instance, the implementation of high-stakes testing regimes
and the limitations of many testing instruments may lead to the teaching of basic skills, which are more easily and accurately assessed (Anagnostopoulos, 2003; Lobascher, 2011). This is a worrying trend in education. If students are learning primarily for the test—and not for the world outside school—there is a risk that students will not be prepared for society or will become disengaged because of a socially irrelevant learning experience. Testing instruments should match the school curriculum, and both the instruments and curriculum need to match the skills required of students when they finish their formal education.

According to some research, parents have observed that national testing programs have led to the adoption of pedagogical methods that ‘stifle natural creativity and emotional intelligence’ (Alexander, 2010, p. 33) in their children. This is consistent with findings from Darling-Hammond (2010), who found that most of the pedagogical changes that have resulted from the introduction of high-stakes testing and accountability measures ultimately limit the development of higher order thinking. This is because success in the tests mostly does not require the cognitive approaches that are necessary for higher order understanding (Darling-Hammond, 2010). In the United States, similar pedagogical changes have been observed. For example, research suggests that there has been an increase in teacher-centred instruction and the direct transmission of test-related facts (Au, 2007). This is similar to Australian research where 59 percent of teachers believed that NAPLAN narrowed the range of teaching strategies they used (Dulfer et al., 2012).

In addition to the perception that higher order thinking was being de-emphasised, other criticisms suggest that national testing programs have led to less inclusive classrooms for students from different backgrounds (Thompson, 2015). Two other studies from the United States corroborate this finding (Au 2007; Koretz, 2008).

There is a general negative response from teachers, with 67 per cent of teachers in a recent study claiming that NAPLAN did not contribute to learning (Thompson, 2015). The
the same study also found that there exists a range of opinions about the effects of NAPLAN on students (Thompson, 2015).

2.5.2 Intended benefits for teaching

Despite the negative findings, there are claims that current tests may have benefits for teachers. Some teachers in Thompson’s (2015) study found NAPLAN useful because it outlined what was expected for both teachers and students. Further, student results from NAPLAN were also perceived to be valuable if they were used correctly (Thompson, 2015).

In addition to the limited examples of increased student-centred approaches reported by Au (2007), an old, but intriguing study by Firestone, Mayrowitz and Fairman (1998) found that confident teachers were able to adjust for high-stakes testing policies with intellectually challenging approaches. This suggests that it may be possible to combine high-stakes testing with engaging pedagogy. Further, a common argument for national testing programs is that they provide achievement data for teachers and this data can be used to develop more individualised teaching approaches (Au, 2007).

In Australia, one aim of NAPLAN is to give teachers data regarding student performance (ACARA, 2011). It is intended that teachers use this data to better inform pedagogy and therefore, student learning. However, up until the phasing in of online testing at the beginning of 2018, NAPLAN results were received by schools months after the completion of the test, at which point student achievement had shifted and the results were of limited value in informing teaching (Dulfer et al., 2012). This finding is not limited to teachers in Australia. In Germany, Maier (2009) found that only a small minority of teachers altered their teaching in response to test results. However, progress could be possible if the test results were accompanied by professional development aimed at crafting specific pedagogical responses to improve the teaching of weak content areas (Lobascher, 2011). There is some evidence of this occurring with NAPLAN data increasingly being used to
Section 2.4 discussed the impact of national testing programs on curriculum, while Section 2.5 outlined the impacts on pedagogy. Section 2.6 explores the impact of national testing programs on students.

### 2.6 Impact on Students

Student wellbeing has been a small—yet important—research area regarding the effects of high-stakes testing (Polesel et al., 2012). Over the past two decades research has identified the negative effects on students at two levels. At the group level, minorities (such as those with a disability or non-native language speakers) may be disadvantaged due to the language and assumed cultural knowledge in the tests and this could lead to equity issues (Darling-Hammond, 2010). At the individual level, high-stakes testing programs have been linked to lowered self-esteem among some students (Klenowski, 2015; Perrone, 1991), increased pressure on students (Stiggins, 1999), increased dropout rates (Lewis, 2000; Reardon & Raudenbush, 2009) and increased student anxiety and stress (Brown et al., 2004).

The effects of testing may not be limited to lower achieving or more disadvantaged students. It may be that upper middle class students are experiencing stress from the pressures associated with the high expectations of academic performance, possibly leading to undesirable behaviours. For example, there are increasing reports in the United States of upper-middle class students who are at risk of drug and alcohol use as well as internalising and externalising behaviour problems (Luthar, Barkin & Crossman, 2013). It is interesting that these students, who attend schools with high standardised test scores reported higher rates of substance use and delinquency (Luthar & Ansary, 2005). These behaviours may be a
manifestation of the increased pressure on students to perform, and one of the most visible measures of their performance is their results in national testing programs.

2.6.1 Potential equity issues

In addition to the impact on student wellbeing, high stakes testing has the potential to impact on equity. Historically, testing has had a sorting function, most commonly for university selection. Exit examinations have been used extensively to determine access to post-school pathways and they have the power to increase inequality across students and society (Dee & Jacob, 2006). High-stakes testing may be a means through which academic success is transferred into social power among culturally advantaged families (Teese, 2000). Specifically, Dee and Jacob (2006) found exit examinations reduced the probability of completing secondary school, particularly for minority students, and increased the dropout rate in urban and high poverty school districts.

Where high stakes are attached to test results, some low-performing students may be excluded from completing them. An example of this occurred in the United States. Due to the performance pressures under the No Child Left Behind Act, poor-performing students were either temporarily excluded from the tests or encouraged to leave the school (Koretz, 2008). This was designed improve the school’s results as school funding was tied to performance in some American states (Koretz, 2008). Minority groups are more likely to perform poorly in standardised assessments and are therefore, more likely to be excluded from education. This is particularly concerning, as the tests may not accurately measure students’ knowledge or skills due to the limitations of the testing instrument (Athanasou, 2010).

Sorting and exclusion of particular groups of students through high-stakes testing may also increase economic inequality. Students without educational support at home are more likely to be blocked from progressing to the next level of education, which ultimately harms those most vulnerable, and increases the risk of their disengagement with society (Dee &
Another study in Ireland argued that ‘it is generally accepted that there is a disproportionately negative effect of these tests on lower income and minority students’ (Smyth & Banks, 2012, p. 284).

In Australia, participation rates in NAPLAN are very high (ACARA, 2011) as it is a population testing program and parents must request exemption from the test should they not want their children to participate. However; it is not uncommon for a student to be removed by their family and prevented from participating (Dulfer et al., 2012). The most common reported reason for non-participation was the negative impact on the child’s confidence (Dulfer et al., 2012). The negative impact on minorities has been discussed in Australia, where NAPLAN has been criticised for being too Eurocentric and placing Indigenous Australian students at a disadvantage (Klenowski, 2015).

2.6.2 The intrapersonal level

There are many consequences of high-stakes testing programs at the individual student level, including a student’s wellbeing. The reported negative effects of high-stakes tests include anxiety and stress (Lewis, 2000; Stiggins, 1999); negatively affected classroom behaviours and interactions between students, teachers and parents (Brown et al., 2004; Schroeder, 2006); and lowered self-esteem and an increase in student suspensions (Brown et al., 2004). Findings from Singapore identified other issues, including long hours of study and a sense of hopelessness, and more extreme consequences, such as physical punishment and ‘significant numbers’ of children requiring psychiatric treatment (Gregory & Clarke, 2003). The literature also highlights other effects, including distress, inability to sleep, frustration, vomiting (Flores & Clark, 2002), exhaustion, crying and stress (Madaus et al., 2009). The responses to testing may vary due to the expectations students have about their performance. A study in Western Australia indicated that students with higher expectations of academic performance were more likely to experience test anxiety (Thompson, 2003). There are many
factors that may be linked to student responses to assessment including the attitude of their parents and personality of the student.

The value students place on the test and their expectations of success may also be factors. The constructs expectancy and value were presented in a theory of motivation by Eccles (1983), where an individual’s motivation is a product of the value they place on doing well on a task and their expectancy that they are able to do well with effort. Expectancies are a combination of ability beliefs and expectancies for success. Ability beliefs focus on present ability, while expectancies focus on the future (Eccles 1983). Value is defined as attainment value as the importance of doing well on a given task (Eccles, 1983). These two constructs are measured in section 4.1.1 and 4.1.3. There may be other factors, outside parents and academic performance that determine student responses to assessment such as personality.

Personality can be viewed through the lens of the Big Five personality framework, one of the most widely accepted models of personality (John & Srivastava, 1999). The framework measures five major personality types: extraversion, conscientiousness, agreeableness, emotional stability and openness. Students who report negative responses to assessment may also have negative responses to other types of external assessments in life, possibly indicating a tendency to respond poorly to new or stressful situations.

At present, there is limited data gathered from large cohorts of students to investigate the effects of high-stakes tests. Despite the large amount of research listing negative effects on students, these tests could be justified if they resulted in an increase in student learning. Much of this research is from the United States and it is not clear how these findings compare to the Australian context (Rice et al., 2015).

However, it is currently not clear if high-stakes testing improves student learning. As Amrein and Berliner (2002) stated, the success of a high-stakes testing policy is determined by how it affects student learning, not whether it can increase student scores on a particular
test. In a study of schools across 18 American states, it was found that student learning was either indeterminate, remained at the same level it was before the policy was implemented, or decreased when the high-stakes testing policies were instituted. There appear to be consistent negative findings regarding the impact of high-stakes testing on students, without the guarantee of improved student learning.

2.7 The Stakeholders and the Focus of the Literature

Much of the data gathered regarding the emotional and physical effects of testing is based on adult perceptions. That is, the research has sought the perceptions of teachers, parents or counsellors on the impact of the tests. Much less information has been gathered from the primary stakeholders: the children (Polesel et al., 2012). In addition, what research there is into the impacts of national testing programs on students has mostly been conducted outside Australia and research into students’ views is only starting to be investigated in Australia.

However, a large survey of Australian teachers supports many of the international findings (Dulfer et al., 2012). The survey was nationwide, including almost 8,500 educators, and investigated the impact of NAPLAN on pedagogy, curriculum and students’ wellbeing. It is interesting to note that the authors considered the area of student wellbeing as ‘largely ignored’ in the discussion of NAPLAN (Dulfer et al., 2012 p. 4). This study was one of the first in Australia to seek information on students’ emotional responses to NAPLAN.

The survey asked teachers to respond to questions regarding how their students feel about NAPLAN. The results are compelling. Approximately 90 percent of respondents stated that at least some students reported feeling stressed and 60 per cent stated that some students had reported crying (Dulfer et al, 2014). It appeared that NAPLAN is generating stress-related responses in a substantial number of students across Australia (Rice et al., 2015).
Research using teachers as participants identified a variety of responses to NAPLAN. Approximately 50 per cent of teachers reported that ‘some’ students are suffering wellbeing issues due to NAPLAN (Rice et al., 2015); however, 50 per cent of respondents felt that ‘some’ students were looking forward to NAPLAN. This highlights a variation in responses between students. Teachers reported that some students felt stressed, suffered sleeplessness the night before the test, or experienced instances of crying (Rice et al., 2015). The study also gathered information about the perceived impact of NAPLAN on students’ academic self-efficacy—the extent to which a person believes they can execute an action successfully (Bandura, 1997). The data showed a polarisation in teachers’ perceptions, as some students were perceived to be worried about seeming as not as smart as their peers, while others were perceived to have positive responses to the tests (Rice et al., 2015). Generally, teachers reported an increase in student stress responses to NAPLAN until Year 7, with a decrease from Year 7 to Year 9 (Rice et al., 2015). The data also showed that the number of teachers reporting that no students were troubled by stress surrounding NAPLAN increased as students moved through school. The research found that, in order of frequency, student responses to NAPLAN included avoidance, internalising behaviour, perceived pressure to perform, externalising behaviour, mental health issues (e.g., depressed or sad), physical effects and some positive responses (Rice et al., 2015). The review into the impact of NAPLAN in the Australian state of Queensland found that many parents reported that their child was positive and/or highly motivated towards NAPLAN. A majority of these parents also thought their child experienced anxiety about the test (Cumming et al., 2018). These findings were consistent with international studies on student responses to national testing programs, in which stress was the most common negative response reported (Lewis, 2000; Stiggins, 1999). Some teachers also reported that students were concerned about the test affecting their school grades or class allocation (Rice et al., 2015). These findings show that
there are varying responses to NAPLAN; however, the specific number of students with these experiences are unknown and the research taps only teacher perceptions.

A deeper investigation into the impact of NAPLAN is required to gain more detail about how students are impacted by the tests and investigate why they may be experiencing any effects. As Rice et al. (2015) noted, NAPLAN appears to be generating stress-related responses among a substantial number of students across Australia and there is a need for further research to elucidate the reasons behind this. It may be that students are reacting to contextual cues from their parents or teachers (Rice et al., 2015). Messages from the press could also influence their responses. While some data has been reported by teachers, the missing voice in Australian research is data collected directly from students. The inclusion of the students’ perspective is necessary to ensure students are included in the discussion concerning their education.

2.7.1 The students’ perspective

A small number of studies have examined student perceptions of NAPLAN. There are four key studies in this area: Belcastro and Boon (2012), who studied motivation for NAPLAN; Howell (2015) who used a qualitative study to investigate themes in primary school students’ responses to the assessment; Wyn et al. (2014), who conducted interviews with students in 16 schools; and, Rogers et al., (2016), who surveyed primary schools students in Western Australian schools about the stress they felt during NAPLAN. The 2018 review of NAPLAN also included surveys of 2896 students (Cumming et al., 2018). Overall, the majority of the research is focused on students’ negative responses to NAPLAN.

The studies that have drawn on students’ perspectives to date, both here and internationally, have had interesting findings. A two-phase study was conducted in a regional secondary school in northern Queensland and used focus groups and interviews to collect data regarding student motivation for NAPLAN (Belcastro & Boon, 2012). It found that
students were motivated by performance achievement goals over mastery achievement goals. That is, students’ aims were to perform well on the test, rather than master the content. This may be related to the fragmentation of content and scripted pedagogies mentioned by Au (2007; see Section 2.4.3).

In another study that focused on students’ perspectives, Wheelock, Bebell and Haney (2002) explored the pictures that American middle and high school students drew in response to testing. The researchers interpreted the drawings as being associated with anxiety, boredom and anger. This study demonstrated that the students’ perspective could provide insight into the impact of national testing.

These findings about the impact of national testing programs are consistent with research in Australia. A study using 105 responses from primary school students, again using drawings of their experiences, found that NAPLAN had a negative impact on wellbeing of primary school children (Howell, 2015). This appeared to be largely the result of the lack of clear and consistent information about NAPLAN. There was also dissonance between students’ everyday experience of school and the NAPLAN test. This disconnect focused on two themes: the shift in classroom focus from a variety of tasks to a set of skills for a written test and the isolation students felt from the change in their teacher’s role from that of a supporter to that of a supervisor (Howell, 2015). There were also concerns regarding the media discourse surrounding the tests (Howell, 2015; Mockler, 2015). Howell’s (2015) article indicates that primary students experience some negative effects of NAPLAN including feeling nervous, stressed and crying.

One of the most pertinent studies into students’ perceptions of NAPLAN was conducted in Western Australia (Rogers et al., 2016). The study explored Year 3 and 5 students’ perceptions of NAPLAN. The researchers, having reviewed the literature on national testing programs, expected some negative impacts (Rogers et al., 2016). They
investigated emotional distress in parents, teachers and students before and after the NAPLAN tests. They found a general increase in students’ stress levels during NAPLAN, but the average did not exceed ‘a little bit’ (p. 336) when measured by Depression Anxiety and Stress Scale (DASS-21). This level of detail is important as the perception that students are stressed by NAPLAN may not be nuanced enough to capture the complexities within the debate. Interestingly, there is evidence that some students felt prepared before the test, but students’ feelings of confidence dropped significantly after completing the test (Belcastro & Boon, 2012). Other interesting results were that Year 3 students reported less stress than Year 5 students (Rogers et al., 2016 p. 336). This is consistent with the work of Rice et al. (2015), who found that the number of teachers reporting that ‘more than 10 students per class’ were stressed was higher in Year 5 than Year 3. Furthermore, the majority of students who participated in Queensland’s review of NAPLAN indicated that the program was not having a negative impact on their wellbeing (Cumming et al., 2018). The review also identified a considerable proportion of students who reported negative feelings about NAPLAN (Cumming et al., 2018).

Some studies have presented arguments that differ from the previous research. Specifically, data from primary school principals suggested that Year 3 students are more impacted by NAPLAN than Year 5 students (Australian Primary Principals Association, 2013). This discrepancy in the literature highlights the need for data from multiple voices, including students, to provide a comprehensive picture of the impact of NAPLAN. Rogers et al. (2016) found NAPLAN produced a severe negative response, such as crying or feeling sick, in approximately 3 per cent of students in the sample. This is a smaller percentage of students than has been reported by teachers (Dulfer, 2012). It may be that NAPLAN does elevate distress, but not to severe levels for most students. This study aims to clarify some of these discrepancies by collecting data directly from the students themselves.
It is likely that parents have a strong influence on student responses to the tests. Rogers et al. (2016) reported a significant correlation between parent and student self-reported emotional distress during NAPLAN. However, no relationship was found between levels of stress reported by students and their teachers (Rogers et al., 2016). Further research is required to investigate the relationships between parent, teacher and student emotional responses (Rogers et al., 2016).

Rogers et al. (2016) reported that the experience of NAPLAN was not overly stressful for most parents and teachers. Additionally, NAPLAN was not reported to have a severe impact on a large number of primary school students. The study found that a teacher or parent experiencing higher levels of distress during NAPLAN typically also assumed that NAPLAN caused more distress in others. This is consistent with the ‘false consensus effect’ (Rogers et al., 2016 p. 341). According to Rogers et al. (2016), this may be influencing reports about NAPLAN from other stakeholders and obscuring the true impact.

A most recent study using student participants in years 3, 5 and 7 found that attitudes to NAPLAN may vary during the year. Students reported feeling scared, nervous or worried in the preparation for NAPLAN, with more positive feeling after completing the test (Swain & Pendergast, 2018). The case study of two schools showed how different school approaches to NAPLAN may also impact student responses (Swain & Pendergast, 2018). It may be that there are other factors that impact student responses to assessment. For example, research into developmental and educational psychology identified that theories focused on expectancy, value, motivation and cognition can help explain how a student responds to assessment (Eccles & Wigfield, 2002).

Collecting data from students will deepen our understanding of how assessment impacts students and also demonstrate to students that their voice is valuable. It is argued that the student voice should be central to any educational reform agenda (Quaglia, 2014). We
must hear the perspectives of all stakeholders, so we have a clear picture of the current reality in education. Quaglia (2014) believed that the student voice was the missing variable in the educational reform agenda. As such, more research needs to be conducted to listen to the student voice about students’ own perceptions of the impact of NAPLAN.

Finally, while there have been a number of negative responses to testing programs documented in the literature (Brown et al., 2004; Flores & Clark, 2003; Madaus et al., 2009; Rice et al., 2016; Swain et al., 2018; Wyn et al. 2014) it is not known whether these responses are due to the national testing regimes themselves (such as NAPLAN), or whether they are student responses to testing situations more generally, and are equally prevalent when students take internal tests. This study sought to compare student responses to NAPLAN to their responses to their internal tests to understand the degree to which negative reactions are more general responses to testing.

2.8 Conclusion

NAPLAN is a high-stakes test in Australian education (Dulfer et al., 2012; Lingard, 2010; Lobascher, 2011). It has the potential benefits of high-stakes testing, such as increased diagnostic information for students, teachers and schools, the ability to measure the effectiveness of pedagogical and policy interventions, and increased alignment of classroom teaching curriculum with national standards (Polesel et al., 2012). Despite the benefits, high-stakes testing has several limitations. For example, the results of high-stakes tests must be used with caution, as they are susceptible to both measurement error and sampling error and may be affected by outright cheating (Koretz, 2008). In addition, the validity of test scores as an indicator of student learning has been questioned (Baker et al., 2010; Koretz, 2008). Using test scores or increases in test scores as a measure of school effectiveness is concerning for
these reasons. Further, using the scores as a measurement of teacher effectiveness is highly problematic.

International research has demonstrated that high-stakes testing also influences curriculum. High-stakes testing has been found to lead to a narrowing of the curriculum towards the testable material, decreased time on non-assessable subjects, testing and teaching of basic content and an increased alignment towards curriculum standards (Klenowski, 2011; Polesel et al., 2012). There is a risk that non-tested subjects—such as arts, physical education and the humanities—will be marginalised in the curriculum.

There is also evidence that high-stakes testing programs are associated with changes in pedagogy. The impacts include an increased focus on rote learning, drills and practice tests at the expense of deeper learning tasks (Lobascher, 2011; Polesel et al., 2012). The reporting of these changes in pedagogy are primarily from the teacher’s perspective.

Other impacts on students from high-stakes testing regimes have been identified. In some circumstances, high-stakes testing may increase dropout rates (Lewis, 2000), and have negative impacts on disadvantaged students, partly due to their use as a sorting mechanism for educational advancement (Moon et al., 2007). The pressure to perform in high-stakes tests can also contribute to a lack of student engagement due to repetitive material and may lead to increased student anxiety and stress (Darling-Hammond, 2010). Much of the research documenting these negative impacts on students are reported on behalf of the students, usually through the teachers. Therefore, there is a gap in the research regarding data gathered directly from students (Polesel et al. 2012; Rogers et al., 2016).

Of the limited data collected directly from students, there have been interesting findings. In a regional school in Queensland, Australia, Belcastro and Boon (2012) found that students with higher self-efficacy were more likely to feel prepared for the NAPLAN test. It was also highlighted that students who were disengaged from the test lacked the motivation
to master literacy and numeracy skills. In Ireland, a study that investigated changes to student attitudes and pedagogy due to high-stakes testing found that middle-class students showed a preference for narrow examination preparation and considered this good teaching (Smyth & Banks, 2012).

In the light of these findings, it is reasonable to conclude that NAPLAN is a powerful force in Australian education. To understand its impact on the education sector, we must consult all stakeholders. Teachers and school leaders are the sources of much of the current data regarding high-stakes testing. The students’ perspective is yet to be heard. Interestingly, the most recent research into student perceptions of NAPLAN is inconsistent with the findings from educators and parents (Rogers et al., 2016; Cumming et al., 2018) and demonstrates the need for a deeper investigation. This study seeks to fill a gap in the literature by highlighting the students’ perspective and comparing responses to NAPLAN to internal tests. To comprehensively understand the impact of education policy, policymakers need to investigate attitudes to NAPLAN from the students’ perspective.

The research questions in this study aim to understand the impact of assessment from the student perspective. At present, most of the reports about student responses are from secondary sources, such as teachers or parents. Through the analysis of direct student responses, we can add valuable information to the current discourse on the impact and value of NAPLAN in Australian education. These findings can be compared to other studies to develop a more accurate understanding about how students are experiencing assessment in education. This can then be used to improve assessment practices in schools and allow for an improved educational experience for Australian students.
Chapter 3: Methodology and Sampling

3.1 Introduction

The previous chapter outlined the need to seek further knowledge about student attitudes to testing and perceptions of the impact of testing on their experiences of school. This chapter outlines the background in the Australian school system, sampling techniques, construction of the instrument to measure students’ responses to the 11 constructs, procedures in the collection process, data analysis and limitations. This study used a cross-sectional survey methodology to assess students’ attitudes and responses to NAPLAN and internal testing.

3.1.1 Background on the Australian School System

The context for this study was Victoria, a state in the south-east of Australia. The general structure of Australian schooling is that students spend seven or eight years in primary school (from pre–Year 1 to Year 6 or 7), followed by a period of five to six years in secondary school (from Year 7 or 8 to Year 12). Primary school students are aged between 5 and 12 years; secondary school students are aged from 11 to 18 years.

There are two sectors in Australian school education: government and non-government. Non-government sector schools can be further categorised as Catholic or independent, with the Catholic schools forming a system in each state bound by common policies and headed by a branch of the National Catholic Education Commission. Independent schools, as the name suggests, operate independently. Governance of each of these schools is the responsibility of a school council or board of governors.

The responsibility for the funding of schools is shared between State/Territory Governments and the Australian Government. For government schools, the states and
territories are the majority public funder and the Australian Government is a minority public funder. Conversely, for non-government schools, the Australian Government is the majority public funder and the State governments are the minority public funder.

In Australian secondary schools, 66 per cent of students attend government schools, 20 per cent attend Catholic schools and 14 per cent attend independent schools (Australian Bureau of Statistics [ABS], 2018). The overall social mix, based on family income of students, differs significantly between the three types of schools (Preston, 2018). In 2011, the Australian government school sector had almost twice the number of secondary students from low-income families relative to the number from high-income families. By contrast, in the non-government sector (i.e., Catholic and independent schools), the number of secondary students from low-income families was less than half the number from high-income families (Preston, 2018). Overall, on average, independent school students are the most socio-economically advantaged and government school students are the most disadvantaged, with students in Catholic schools falling between (Preston, 2018). In the last 20 years, there has been a general shift of enrolments from government to non-government schools with increasing socio-economic stratification across Australian schools (Watson & Ryan, 2010).

In line with many other systems internationally, Australia has adopted what could be described as market-based reforms in its education system. This includes the implementation of NAPLAN as a national program of external standardised testing, with school comparisons provided through the My School website. Data from this program feeds into a national reporting program that provides information at the individual, school and system level.

There are multiple levels of feedback from the NAPLAN tests. The aggregated year level results of individual schools are made publicly available on the My School website. Users can search for a school, view the school averages in all domains from 2008 onwards and compare the school to similar schools and to national average levels of achievement.
There are also annual national NAPLAN reports, which show the performance across the nation of students in each state and territory. At the school level, schools receive information about how their students performed at the aggregate and at the individual level; meanwhile, parents receive an individual report that compares their child to national averages and to national minimum standards (Lingard et al., 2016). In addition, state departments of education receive each school’s student achievement data. These data are often used to target lower-performing schools for improvement measures.

These testing and accountability reforms are similar to those adopted in the United States, Canada and the United Kingdom. NAPLAN allows for comparisons of performance over time and is used to make comparisons between schools. This has promoted market-like conditions, as parents are encouraged to choose a school for their child, based on NAPLAN data on the My School website (Lingard et al., 2016). These changes in assessment may also be used to drive changes in teaching and learning practices. This study aims to identify how these changes in assessment are influencing the stakeholders who, arguably, are most affected: the students.

3.2 Sampling: Using Students as the Critical Reference Group

Students are the most important stakeholders in education and they hold a unique position in educational research. Previous research with students has often positioned them passively as objects, rather than the subjects of research (Veale, 2005). Investigations into the impact of national testing programs have often drawn on the views of teachers (e.g., Lingard, 2010; Moon et al., 2007), school administrators (e.g., Athanasou, 2010) and to a lesser degree, parents (Wyn et al., 2014). At the heart of this thesis is the belief that students are not passengers in education, but active and capable participants who can proactively participate for the greater good of learning (Quaglia, 2014). Therefore, this study uses students as the
primary source of data to understand their responses to NAPLAN, and how it might be associated with their expectancy, value and motivation for the test.

Of the many theories for motivation, this study used the Eccles and Wigfield (2002) Expectancy-Value model to understand the students’ perspective. This model assumes that expectancies and values are influenced by social cognitive variables including, perceptions of the difficulty of tasks, perceptions of competence and an individuals’ goals (Eccles & Wigfield, 2002). These variables are also influenced by individuals’ perceptions of other peoples’ attitudes and expectations for them (Eccles & Wigfield, 2002). I anticipated that a student’s expectancy, value and motivation for a task would be associated with the attitudes and expectations of their teachers and parents. In turn, this would impact their responses to an assessment task. For example, students with high value and low expectations for a task may be more likely to experience a stressful response. Conversely, students with low value and high expectations may be more confident and less likely to experience negative responses. Identifying these relationships could allow for identification of students who are likely to respond negatively to assessments and provide opportunities to support them.

A maximum variation sampling strategy was adopted, with the aim of capturing participants from a range of different settings. The sampling frame targeted schools from various sectors and geographical locations. The aim was to include students from a cross-section of socio-economic and geographical backgrounds and school sectors. The incorporation of students from a variety of schools was essential for creating a broad information base in this study, from which the researcher could identify general attitudes and differences between groups. The study aimed for a sample of 50 students from each participating school. This was to allow for the identification of common patterns across a range of different groups and identification of variations between groups.
3.2.1 Participant recruitment and administration of the survey

The process for collecting data consisted of several steps. Initially, a letter was sent to principals of schools inviting them to participate in the study. This was followed by a phone call to ascertain their interest in participation. If the principal agreed to participate, a teacher usually became the main contact point. In Australian educational research, principals must give consent before teachers or students in their school can participate in an educational study. After consent from the principal, an information package in which the study was described was sent to the teacher. It included Plain Language Statements for parents and teachers, permission forms for teachers and parents to complete (on behalf of their students), the link to the survey and incentives for the students. Students were offered the incentive of a small piece of confectionary for returning the completed permission form (regardless of whether or not their parents had agreed to their participation). Students were also required to give consent before completing the survey.

Once the parents of a full class (or majority of a full class) of students had returned the permission forms, the students would take the online survey as a class. The collection of the data occurred at a time convenient to the teachers within the school. Surveys were completed in class time, usually in the last ten minutes of class, on school computers or on personal mobile devices.

3.2.2 Sample

The study was conducted in Victoria, a state that is home to approximately 25 per cent of the Australian population (ABS, 2018). A total of 15 schools were approached, of which four agreed to participate in the study. The schools were all secondary schools offering places in Years 7–12 (catering for children who were 11–18 years old). Table 1 shows the characteristics and response rates of the participating schools.
Table 1

<table>
<thead>
<tr>
<th>Sector</th>
<th>Metro, rural or regional</th>
<th>Boys, girls or co-ed</th>
<th>Approximate total students in school</th>
<th>Number of students invited to participate</th>
<th>Number of participants</th>
<th>Response rate (%)</th>
<th>ICSEA¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Catholic</td>
<td>Metro</td>
<td>Boys</td>
<td>1,500</td>
<td>75</td>
<td>53</td>
<td>70%</td>
<td>1,072</td>
</tr>
<tr>
<td>B Government</td>
<td>Metro</td>
<td>Co-ed</td>
<td>1,000</td>
<td>65</td>
<td>65</td>
<td>87%</td>
<td>1,011</td>
</tr>
<tr>
<td>C Catholic</td>
<td>Regional</td>
<td>Co-ed</td>
<td>1,000</td>
<td>50</td>
<td>16</td>
<td>32%</td>
<td>1,002</td>
</tr>
<tr>
<td>D Government</td>
<td>Rural</td>
<td>Co-ed</td>
<td>300</td>
<td>100</td>
<td>72</td>
<td>70%</td>
<td>955</td>
</tr>
</tbody>
</table>

Note. ¹ The ICSEA was created by ACARA specifically to enable meaningful comparisons of students’ NAPLAN test achievement scores in schools across Australia. Key factors in students’ family backgrounds (parents’ occupation, school education and non-school education) have an influence on students’ educational outcomes at school. In addition to these student-level factors, research has shown that school-level factors (e.g. a school’s geographical location and the proportion of Indigenous Australian students for which a school caters) need to be considered when summarising educational advantage or disadvantage at the school level. ICSEA provides a scale that numerically represents the relative magnitude of this influence and is constructed to account for both student- and school-level factors. ICSEA is set at an average of 1,000. The lower the ICSEA value, the lower the level of educational advantage for students who go to this school; the higher the ICSEA value, the higher the level of educational advantage for students who go to this school (ACARA, 2018).
The selection of participant schools was not random. The participant schools each had a member of staff with a previous relationship to the researcher and this contact was used to facilitate the research. The students targeted for participation in the study were those who had most recently completed NAPLAN (i.e., Year 7 and 9 students), as it was felt that these students were likely to have the most recent and accurate recollections about their participation. NAPLAN tests are held in May of each year; data were collected in May and June, to ensure that memories of participation were relatively fresh. All students participating completed the NAPLAN ‘pen and paper’ test, compared to the online NAPLAN test piloted in 2017.

The justification for choosing secondary school students rather than primary students was due to the students’ experience of testing. Having sat the NAPLAN tests on multiple occasions (every two years since Year 3), the students would have experienced the testing several times. Additionally, students in secondary school would have had considerable experience in formal internal tests. The questionnaire was only open to students at the four schools invited to participate in the study. Limiting the survey to selected students is preferable to an uncontrolled web survey, as the latter is more likely to produce a very biased sample (Dillman, 2000).

The gender of students participating in the study is outlined in Table 2.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage of sample</th>
<th>State percentage (2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>135</td>
<td>66%</td>
<td>51.9%</td>
</tr>
<tr>
<td>Female</td>
<td>71</td>
<td>34%</td>
<td>48.1%</td>
</tr>
<tr>
<td>Total</td>
<td>206</td>
<td>100%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
Compared with the data on gender composition of the Victorian student population, there is a greater percentage of males in the sample compared to the general population. In Victorian schools, 52 per cent of the student population is male (Victorian Department of Education and Training, 2017) compared to 66 per cent in this study’s sample. This difference exists because there was one boys’ school in the sample. This limitation needs to be borne in mind when interpreting results from the study.

Table 3 describes the sample by year level.

Table 3
*Student Participation by Year Level*

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
<th>Percentage of sample</th>
<th>Total in state (2015)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>131</td>
<td>64%</td>
<td>67,014</td>
</tr>
<tr>
<td>Year 9</td>
<td>75</td>
<td>34%</td>
<td>67,146</td>
</tr>
</tbody>
</table>

In Victorian schools, there are very similar numbers of students in Year 7 and Year 9. In this sample, there is an over-representation of students in Year 7. There are two reasons for this. Some participating schools had smaller classes in Year 9 than in Year 7. Additionally, a greater number of Year 7 students returned their permission forms.

The description of the sample by sector can be found in Table 4.

Table 4
*Student Participation by Sector*

<table>
<thead>
<tr>
<th>Sector</th>
<th>Number</th>
<th>Percentage of sample</th>
<th>State percentage (2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government</td>
<td>137</td>
<td>67%</td>
<td>61%</td>
</tr>
<tr>
<td>Catholic</td>
<td>69</td>
<td>33%</td>
<td>22%</td>
</tr>
<tr>
<td>Independent</td>
<td>0</td>
<td>0%</td>
<td>17%</td>
</tr>
</tbody>
</table>
The sample has a slightly higher percentage of students from government schools than the state percentage of Victorian students attending government schools. In Victoria, 64 per cent of the secondary student population attends government schools (Victorian Department of Education and Training, 2017). There is also a considerable over-representation of Catholic students in this study compared to the state’s 17 per cent. The sample does not include independent schools, as the independent schools contacted declined to participate in the study.

Table 5 outlines the sample by language background. The sample has a slightly lower percentage of students with a LBOTE than the state average. This is shown by 12 per cent of the sample population identifying as LBOTE, compared to 16 per cent of the Victorian school population.

Table 5

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage of sample</th>
<th>State percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native English speaker</td>
<td>182</td>
<td>88%</td>
<td>84%</td>
</tr>
<tr>
<td>Language background other</td>
<td>24</td>
<td>12%</td>
<td>16%</td>
</tr>
</tbody>
</table>

It must be noted that the sample was never intended to be fully random or fully representative. Random sampling for postgraduate education research students is difficult due to the limited access to schools. Further, it can be difficult to ensure students return their signed permission forms to participate in a study. Also, as only four schools were represented in the study, it does not have the breadth to give a true indication of the impact of testing on students across the state. All conclusions and policy implications must consider the limitations of the relatively small sample form a particular set of schools. The sample did not
include independent schools and further research should aim to include a broader range of schools and students.

Despite some limitations in the nature of the sample, the 206 students from Victoria who participated can give us valuable knowledge about the impact of national testing programs on their experience of school.

3.2.3 Ethics

Ethics approval was obtained from the Victorian Department of Education and Training, the Catholic Education Office (Melbourne), the Catholic Education Office (Sale) and the University of Melbourne Human Research Ethics Committee. The ethics approvals for the study can be found in Appendix A. The Plain Language Statement and consent forms can be found in Appendices B and C.

Participation in the student survey was anonymous and individual student responses cannot be identified. Participating schools have been given pseudonyms in the thesis and this will be the case for any publications from the research.

3.3 The Study Design

Surveying the students from a variety of school settings captured a broad range of views from students with different school and life experiences. This is consistent with the primary goal of the research, which was to identify students’ responses to and perceptions of testing. The survey was the primary method of data collection used in this research. The survey was chosen as the primary instrument for its efficiency in many areas. The research was conducted in school settings—environments in which time is a valuable resource—and the survey was an efficient way of measuring many variables. The cost was minimal and it allowed for data collection from a broad spectrum of individuals within a reasonable period (Alasuutari, Bickman & Brannen, 2008). In the context of this study, it was also the most
appropriate means available for developing a picture of the attitudes and characteristics of a small sample of the greater student population (Alasuutari et al., 2008).

This study aimed to investigate the relationships between variables of gender and year level, and attitudes and emotional responses to assessment. Therefore, a quantitative approach, using a survey, allowed for these types of analyses.

The interface used to collect the information was an online survey tool, ‘typeform’. The choice of an online interface was due to the ease of data aggregation, student familiarity with online tools and the efficiency of distribution of the survey.

There are multiple reasons for choosing surveys when working with children. An electronic survey provides anonymity for participants. Students were also made aware of the purpose of their participation via an introduction page to the survey (in addition to the Plain Language Statement) that gave a simple explanation of the purpose of the survey and the ways in which the data would be used.

3.3.1 The instrument

A 36-item questionnaire was developed, which gathered information from students on a range of topics. Specifically, the survey collected data on the following:

- demographic information
- student attitudes to assessment
- responses to tests
- student perceptions of teacher attitudes to assessment
- student perceptions of parent attitudes to assessment
- student expectations of performance in assessment
- student levels of motivation for assessment
- student experience of feedback from the tasks
- student experience of preparation for the tasks
• student self-efficacy in the assessed subjects
• personality factors.

The survey instrument used in this study was built upon existing instruments. Some of the constructs investigated have been measured in previous research on the impact of NAPLAN (Polesel et al., 2012; Rice et al., 2015; Wyn et al., 2014). Several limitations to this approach need to be acknowledged. Gotch & Hall (2004) identify weaknesses when using questionnaires to obtain data from children. Some weaknesses include the difficulty of examining complex opinions through surveys and it may not be known if participants understand the questions asked (Gotch & Hall, 2004). Also, literary skills at year 7 could limit students’ ability to comprehend information, to alleviate this; the classroom teacher was available to interpret questions for students who needed support.

There may be a response bias within the sample, where students who did not return permission forms may have reasons for not wanting to discuss NAPLAN including negative experiences in the past. Additionally, students who did not return their permission forms may be part of a forgetful or lower achieving group. The study assumes that students were be honest in their answers and that students who did not participate did not differ significantly from those who did.

The full survey instrument can be found in Appendix D; however, a summary of the construct and rationale for questionnaire items is included in Section 3.5.1.1.

3.3.1.1 The construct and rationale for questionnaire items

Demographic information was collected through items 1a, 1b, 1c, 1d and 1e. Participants were asked to indicate their year level, gender, postcode, school type (i.e., government, independent or Catholic) and the language spoken at home.
The instrument allowed for data on student attitudes to be collected. This included responses to NAPLAN and internal tests, perceptions of value, expectancy, motivation, feedback, impact on learning at school, academic self-concept and personality traits.

Data on student emotional and physical responses to assessment tests were collected according to three categories: positive (confident, excited and enthusiastic), negative emotional (nervous, stressed and bored) or negative physical (sleeplessness, freezing, feeling sick and crying). The stem for these items was, ‘Did you experience any of the following before or during NAPLAN?’ The students responded with either ‘Yes’ or ‘No’ to the items.

Data on student perceptions of the value placed on assessments by parents and teachers were collected through items 3, 14 and 15. For example, students were asked, ‘How important is it to your teachers/parents, for you to do well in NAPLAN this year?’ They were also asked if their teachers or parents had spoken to them about their NAPLAN results. The survey asked matched questions about internal tests.

Data on student expectations of performance in each of the NAPLAN tests were collected in items 6, 7 and 8 and expectations for internal tests in items 9 and 10. On a five-point Likert scale, students were asked to respond to statements such as, ‘I expect to do well in NAPLAN English this year’. The possible responses ranged from (1) ‘strongly disagree’ to (5) ‘strongly agree’.

Student levels of motivation for NAPLAN and internal tests were measured through items 27, 28, 29 and 30. On a Likert scale, students were asked ‘How hard did you try in the NAPLAN English/mathematics test?’ The possible responses ranged from (1) ‘I didn’t try at all’ through to (5) ‘I put in my best effort’.

The level of feedback from the assessment tests was measured through items 11, 12, 13, 14 and 15. A sample item was ‘How are NAPLAN results used at your school?’ The
possible responses were ‘I don’t know’, ‘To place students in classes’, ‘To decide who can enrol in particular classes’ and ‘They aren’t used at all’.

The impact on learning at school was measured through items 18, 19, 20, 21, 22 and 31. Students were asked about their perceptions of the emphasis the school placed on NAPLAN and how many practice tests they had completed at home or at school. A sample item was ‘Have you completed any practice tests for NAPLAN at school?’ Possible responses were ‘no, never’, ‘1–2 tests’, ‘3–5 tests’ or ‘6 or more tests’.

Academic self-concept was measured in items 23 and 24. Academic self-concept it is an individual’s perceptions about themselves in achievement situations (Bong & Skaalvik, 2003). It consists of fairly stable perceptions of the self that are past-oriented and are used to predict outcomes including motivation, emotion and performance (Bong & Skaalvik, 2003). Assessing academic self-concept relies heavily on social comparative information and the most commonly used method for measuring academic self-concept is self-reports (Marsh, 1999). In this study, students were asked about their perceptions of their relative achievement in each subject using questions based on the Academic Self-Description Questionnaire (Marsh, 1999). A sample item was ‘Which sentence best describes how you are going in mathematics?’ Possible responses were ‘I am doing really well in this subject’, ‘I am not one of the top students, but I am above average’, ‘I’m about average in this subject’ or ‘I find this subject difficult’.

Personality was measured in item 2. Students completed the Ten Item Personality Inventory (TIPI) questions (Gosling, Rentfrow & Swann, 2003). The TIPI was included to explore the relationships between students’ personalities and their response to testing. The TIPI is constructed using the Big Five personality framework, one of the most widely accepted models of personality (John & Srivastava, 1999). The most comprehensive measure of the Big Five personality traits is Costa and McCrae’s (1992) NEO personality inventory,
but with 240 items and an estimated 45-minute completion time it was deemed impractical for this study. Evidence of the appropriateness of the TIPI has been previously identified by Gosling et al. (2003), who found that as a brief measure of personality, the TIPI reached adequate levels in each of the criteria against which it was evaluated: convergent and discriminant validity, test–retest reliability and patterns of external correlates. The TIPI measures introversion/extroversion, agreeableness, conscientiousness, openness to new experiences and emotional stability. Participants were asked to ‘rate the extent to which you agree or disagree with statements e.g. ‘I am outgoing, enthusiastic’. Responses were on a 7-point Likert scale and range from disagree strongly to agree strongly.

A copy of the full survey can be found in Appendix D.

3.3.2 Pilot testing

The survey was pilot tested in three stages to ensure coverage of the key concepts and clarity. Initially, the survey was piloted with researchers and professional colleagues to determine if it covered the concepts in question. In the second stage, it was reviewed, edited and tested again with researchers and professional colleagues. Once accepted, the third stage involved piloting the study with 10 students in the target age group. Adaptations and clarifications on key terms were made before deeming the survey appropriate for the research participants.

3.4 Data Analysis

In a sample size of 206 students, a probability level of 0.5 was set and \( p \) values were reported. The open-ended responses were analysed using Miles and Huberman’s (1994) style thematic analysis grids.

The data analysis required the use of appropriate statistical tests: Goodman and Kruskal’s gamma, McNemar’s test and Fisher’s exact test. Primarily, this study used Goodman and Kruskal’s gamma as a measure of ordinal by ordinal association. This study
used the following conventions to describe the strength of the relationship when assessing findings (Laerd Statistics, 2016):

- 0.00–0.20 denotes no association
- 0.20–0.40 denotes weak relationship
- 0.40–0.60 denotes moderate relationship
- 0.60–1.00 denotes strong relationship.

This study used McNemar’s test to look for significant differences in paired binary data, in which the same student was reporting their responses to both NAPLAN and internal tests.

For associations between binary variables, Fisher’s exact test (rather than the chi-square test) was used, due to its superior properties for smaller sample sizes. Fisher’s exact test was used when the sample was split into two groups (e.g., participants with high academic self-concept and those with low academic self-concept) to explore the differences in participants’ responses (e.g., those reporting being stressed or not stressed by assessments).

3.4.1 Sample limitations

The study consisted of 206 students from four Victorian schools. There was a high response rate for those invited to the study (71%). The schools in the sample were mostly near the median of the ICSEA value and this allowed for data gathered from middle class students, but meant that student responses in very high and low SES schools are not included. To address possible comprehension issues for students with low literacy skills, the survey was piloted with students and the classroom teacher was available to interpret questions for students who needed support. The study assumed that students were honest in their answers and that those who did not participate did not differ significantly from those who did.

Overall, this study will give educators, policymakers and parents a valuable insight into how students experience NAPLAN and internal tests. By students directly reporting their own
experience of NAPLAN, this study created a detailed exploration of how assessment impacts various groups of students.
Chapter 4: Student Perceptions of Assessment: NAPLAN and Internal Tests

4.1 Student Perceptions of NAPLAN

This chapter reports the student perceptions of NAPLAN and internal tests. It quantifies the reporting of responses to give a snapshot of how a group of students experience NAPLAN. It also determined whether student responses to NAPLAN were similar to their responses to internal testing. Section 4.1 reports students’ emotional and physical responses to NAPLAN. Sections 4.1.2 to 4.1.4 show the student expectancy, value and motivation for NAPLAN. Section 4.2 reports the students’ emotional and physical responses to internal testing and 4.2.1-4.2.4 show the student expectancy, value and motivation for internal testing.

4.1.1 Student responses to NAPLAN

Previous research has explored student’s emotional and physical responses to national testing programs. Findings indicate that some students experience increased anxiety and stress (Brown et al., 2004; Lewis, 2000) and negative physical responses, including freezing during the test (Paris & McEvoy, 2000), an inability to sleep the night before (Flores & Clark, 2003) and crying (Madaus et al., 2009).

More recently, research in the Australian context suggests a large proportion of educators perceived that at least some of their students are suffering wellbeing issues because of NAPLAN (Rice et al., 2015). While studies have stated these responses exist, how prevalent they are within the student population has yet to be quantified. There appear to be no studies investigating whether there are positive student responses to testing.

To bridge this gap, this study sought to gain insight into both positive and negative emotional and physical student responses to assessment. To achieve this I adapted previous
research by Dulfer et al. (2012) who identified teacher perceptions of students’ negative emotional and physical responses to NAPLAN and extended their work to include positive responses (e.g., feeling excited, confident or enthusiastic).

Students were asked, ‘Thinking back to the last time you did NAPLAN, did you experience any of following?’ and were required to answer ‘Yes’ or ‘No’ to each response. Students were able to select ‘yes’ to more than one answer. Responses were grouped into three categories: positive, negative emotional and negative physical. These included the following descriptors:

- positive emotional responses—feeling confident, enthusiastic or excited
- negative emotional responses—feeling nervous, bored or stressed
- negative physical responses—sleeplessness, freezing, feeling sick and crying.

Table 6 shows student reported responses to NAPLAN.
Table 6

*Student Reported Responses to NAPLAN*

<table>
<thead>
<tr>
<th></th>
<th>Number of ‘yes’ responses (out of 206)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bored</td>
<td>108</td>
<td>52%</td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>99</td>
<td>48%</td>
</tr>
<tr>
<td>Confident</td>
<td>59</td>
<td>29%</td>
</tr>
<tr>
<td>Stressed</td>
<td>55</td>
<td>27%</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>36</td>
<td>18%</td>
</tr>
<tr>
<td>Excited</td>
<td>34</td>
<td>17%</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>26</td>
<td>13%</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>22</td>
<td>11%</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>20</td>
<td>10%</td>
</tr>
<tr>
<td>Crying</td>
<td>6</td>
<td>3%</td>
</tr>
</tbody>
</table>

*Note: Students could select ‘yes’ to more than one response so percentages sum to more than 100.*

The most common responses were negative emotional responses, followed by positive emotional responses. The least commonly reported responses were negative physical responses.

4.1.1.1 *Negative emotional responses: Feeling bored, nervous or stressed*

The most frequently cited negative emotional response was boredom, with over half of the participants (i.e., 52 per cent) reporting that they felt bored during the tests. Boredom is associated with tedious activity that students find meaningless (Mora, 2011). It may be that students do not see the meaning of these tests and do not find them stimulating. Additionally, many participants reported either feeling nervous (48 per cent) or stressed (27 per cent). The
variety of negative responses suggests that testing is having a negative impact on some students.

4.1.1.2 Positive emotional responses: Confident, enthusiastic and excited

Despite negative emotional responses many students in this sample also reported positive responses to testing. Approximately one in four students reported at least one positive response, while eight per cent of students reported all three positive responses—feeling confident, enthusiastic and excited. Furthermore, four per cent reported being both confident and enthusiastic, whereas three per cent reported feeling enthusiastic and excited. This suggests that for some students, testing may be a positive experience.

4.1.1.3 Negative physical responses

A small number of students reported negative physical responses to testing. The most common were sleeplessness the night before NAPLAN (13 per cent) and freezing during the test (13 per cent)—this was slightly above the percentage of students who reported feeling sick (12 percent) and substantially above the percentage reporting crying (3 per cent). It must be noted that these small percentages refer to a small number of students. For example, the 3 per cent of students who reporting crying refers to six students.

4.1.1.4 Open-ended responses

In addition to the structured survey questions, students were given the opportunity to respond more broadly with an open-ended question. Students were asked ‘Did you have any other responses to the NAPLAN test?’ Students were able to type a sentence in a text box. Data from the open-ended responses were analysed using Miles and Huberman (1994) grids to identify themes. Almost one in three (29 per cent) participants wrote an open-ended response. The comments were analysed and grouped into seven subgroups: internalising behaviour, avoidance, concerns about performance, externalising behaviour, negative
emotional responses, positive responses and neutral. The key themes are reported in Table 7 below.

Table 7

**Thematic Categorisation of the Responses to the Open-Ended Section of the Survey**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Sample responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internalising behaviour</td>
<td>‘Sometimes nervous’ and ‘weird’</td>
</tr>
<tr>
<td>Avoidance</td>
<td>‘Didn’t want to come to school’</td>
</tr>
<tr>
<td>Concerns about performance</td>
<td>‘Worried’ and ‘didn’t think I would do that good’</td>
</tr>
<tr>
<td>Externalising</td>
<td>‘I got into an argument with one of my friends’</td>
</tr>
<tr>
<td>Negative emotional responses</td>
<td>‘Depressed’ and ‘sad’</td>
</tr>
<tr>
<td>Positive</td>
<td>‘Fun’, ‘calm’ and ‘ready’</td>
</tr>
<tr>
<td>Neutral</td>
<td>‘Didn’t care’, ‘normal’ and ‘hungry’</td>
</tr>
</tbody>
</table>

Of those who gave an opened ended response, the three largest subgroups included responses that suggest concerns about performance (30 per cent; e.g., “nervous”, “worried”, “pressured” or “dumb”), neutral/apathetic responses (26 per cent; e.g., “normal”, “bored”, “same as any other lesson” or “hungry”) and negative emotional responses (16 per cent; e.g., “depressed”, “scared” or “sad”). These findings suggest that NAPLAN elicits different responses depending on the individual student. NAPLAN causes some students to stress but may be exciting for others.

An important factor impacting students’ responses to NAPLAN is the feedback they receive. To contextualise the student response it is important to explore how students perceive the feedback they gain from NAPLAN.
Individual differences in responses to assessment can be explored through the value and expectancy students place on the tasks. For the purpose of this study expectancy is defined as the domain-specific perceived ability to perform well on a given task (Eccles & Wigfield, 2002) (e.g. ability to perform well on the NAPLAN mathematics test) and the value as the perceived importance of doing well on a task (Eccles & Wigfield, 2002).

4.1.2 Student expectations of performance in NAPLAN

This section investigated how students expected to perform in NAPLAN (i.e., their expectancy). Expectancy is the perceived ability of an individual to perform a task (Eccles & Wigfield, 2002). To measure expectancy, participants were asked to what extent they agreed or disagreed with the statement, ‘I expect to do well on NAPLAN this year’ for English and mathematics. They responded on a five-point Likert scale from (1) ‘Strongly disagree’ to (5) ‘Strongly agree’.

Table 8 shows the student perceptions of expectancy regarding their performance on NAPLAN for English and mathematics.

Table 8

Student expectations for performing well in NAPLAN

<table>
<thead>
<tr>
<th></th>
<th>1: Strongly disagree</th>
<th>2: Disagree</th>
<th>3: Neither agree or disagree</th>
<th>4: Agree</th>
<th>5: Strongly agree</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>4%</td>
<td>6%</td>
<td>27%</td>
<td>42%</td>
<td>22%</td>
<td>3.68</td>
<td>1.02</td>
</tr>
<tr>
<td>Mathmatics</td>
<td>5%</td>
<td>7%</td>
<td>23%</td>
<td>34%</td>
<td>31%</td>
<td>3.76</td>
<td>1.11</td>
</tr>
</tbody>
</table>

First, a paired sample t test indicated that there were no statistically significant differences between student expectations to do well in NAPLAN English and to do well in NAPLAN Mathematics. Most students had positive expectations about their performance in
NAPLAN. A high percentage of students—64 per cent in English and 65 per cent in mathematics—agreed or strongly agreed with the statements. Open-ended comments following this question included:

- ‘I felt calm’ (Year 7 student)
- ‘give my best shot’ (Year 7 student)
- ‘focussed and fun’ (Year 9 student).

Despite the general response showing that most students expect to do well, the data showed approximately a quarter (23 and 27 per cent) of students were somewhat ambivalent in their expectations, selecting ‘Neither agree nor disagree’. There was also a small group of students (4–5 per cent) who strongly disagree with the statements. Some of these students in the open-ended questions reported they felt a variety of emotions leading up to the test including:

- ‘low self-confidence’ (Year 7 student)
- ‘didn’t think I would be that good’ (Year 7 student)
- ‘dumb and worried’ (Year 7 student).

In sum, students have a variety of expectations regarding NAPLAN. Most students expected to perform well; however, a small group of students held negative expectations.

4.1.3 Value of doing well in NAPLAN

This study sought to examine the relationship between value, expectancy, and motivation for NAPLAN among students. Firstly, we sought to measure student perceptions of the value placed on doing well in NAPLAN according to three stakeholder groups: the students, their parents and their teachers. Three survey items, one per stakeholder group, were used to measure the ‘value placed on NAPLAN’ construct:
1. How important is it you to do well in NAPLAN? (i.e., student value)

2. How important is it to your parents that you do well in NAPLAN? (i.e., perceived parental value)

3. How important is it to your teachers that you do well in NAPLAN? (i.e., perceived teacher value).

The participating students indicated the level of importance on a five-point Likert type scale from (1) ‘Not important at all’ to (5) ‘Very important’. The primary outcome variable was the overall sample mean for students’ perceptions of value to self, parents and teachers. Table 9 shows student perceptions of the value placed on NAPLAN by each stakeholder group.

Table 9

<table>
<thead>
<tr>
<th></th>
<th>1: Not important at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers</td>
<td>4%</td>
<td>8%</td>
<td>21%</td>
<td>37%</td>
<td>31%</td>
<td>3.72</td>
<td>1.16</td>
</tr>
<tr>
<td>Parents</td>
<td>7%</td>
<td>13%</td>
<td>16%</td>
<td>29%</td>
<td>34%</td>
<td>3.63</td>
<td>1.31</td>
</tr>
<tr>
<td>Students</td>
<td>12%</td>
<td>10%</td>
<td>23%</td>
<td>31%</td>
<td>24%</td>
<td>3.32</td>
<td>1.37</td>
</tr>
</tbody>
</table>

A paired samples t test indicated that there was a significant difference between the importance teachers were perceived to place on NAPLAN and the importance placed on NAPLAN by students. Students placed a significantly lower value on doing well in NAPLAN ($M = 3.32, SD = 1.37$) than their teachers were perceived to place on doing well ($M = 3.72, SD = 1.16, t(205) = 4.21, p < 0.001, d = 0.29$). Further, a paired samples t test indicated that there was a significant difference between the perceived importance parents and students placed on NAPLAN, with parents ($M = 3.63, SD = 1.31$) perceived to place a
significantly higher value compared to students, \( [M = 32, \text{SD} = 1.37, t(205) = 2.34, p < .001, \text{d} = 0.231] \). There was no significant difference between students’ perceptions of value between teachers and parents. Overall, students perceived that NAPLAN scores were more important to their teachers and parents than themselves.

A small group of students placed a low value (i.e., ‘Not at all important’ or ‘Not important’) on doing well in NAPLAN. Qualitative data from this minority, collected in the open-ended section of the survey highlighted their apathy towards NAPLAN:

- ‘NAPLAN doesn’t reflect our learning’ (Year 7 student)
- ‘didn’t care, we just did it’ (Year 9 student)
- ‘wasting hours of my life for a pointless test’ (Year 9 student).

The relationships between student perceptions of teacher and parent value and responses to NAPLAN will be discussed in chapter 6.

### 4.1.4 Motivation for NAPLAN tests

This section assesses students’ motivation for NAPLAN, with motivation defined as persistence and vigour in completing tasks (Eccles & Wigfield, 2002). To measure student motivation in NAPLAN, students were asked, ‘How hard did you try in NAPLAN?’ across each of the four test domains: mathematics, spelling, language conventions and writing. The students responded using a Likert scale, ranging from (1) ‘I didn’t try at all’ to (5) ‘I put in my best effort’. Table 10 shows student reported motivation for NAPLAN.
Table 10

Student motivation for NAPLAN

<table>
<thead>
<tr>
<th></th>
<th>1: I didn’t try at all effort</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>5%</td>
<td>5%</td>
<td>12%</td>
<td>32%</td>
<td>47%</td>
<td>4.05</td>
<td>1.13</td>
</tr>
<tr>
<td>Spelling</td>
<td>4%</td>
<td>4%</td>
<td>17%</td>
<td>27%</td>
<td>49%</td>
<td>4.05</td>
<td>1.12</td>
</tr>
<tr>
<td>Language</td>
<td>3%</td>
<td>4%</td>
<td>15%</td>
<td>30%</td>
<td>49%</td>
<td>4.09</td>
<td>1.06</td>
</tr>
<tr>
<td>Conventions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>4%</td>
<td>5%</td>
<td>17%</td>
<td>26%</td>
<td>49%</td>
<td>4.03</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Most students reported putting in their best effort in every domain. Almost half of the sample reported putting in their best effort across all test domains and the comments in the open-ended section suggested some students enjoyed the experience:

- ‘Ready and prepared’ (Year 9 student)
- ‘felt good’ (Year 7 student)

A paired sample t test indicated that there were no significant differences in student motivation between each test domain of NAPLAN. Furthermore, seven per cent of students reported ‘I didn’t try at all’ across all domains, suggesting that low motivation for one domain often coincides with low motivation in other domains. Some of the students who reported that they did not try at all indicated in the comments section that they wanted to avoid the test or did not see the relevance of the testing program:

- ‘I didn’t want to come to school’ (Year 9 student)
- ‘Wasting hours of my life for a pointless test’ (Year 9 student)
- ‘NAPLAN doesn’t reflect our learning’ (Year 7 student).
Most students are motivated for NAPLAN and try hard to do well, however a small group of students reported low levels of motivation for NAPLAN and put little effort into the test. The relationships between expectancy, value and motivation are explored in Chapter 6.

4.1.5 Feedback from NAPLAN

NAPLAN provides Australian schools and parents with feedback about student achievement in literacy and numeracy. The feedback most relevant to this study is the individual student report, given to the parents. It contains the ‘band’ or level the student achieved and compares it to the school and national average for the tested domains of mathematics, reading, writing, and language conventions (spelling, grammar and punctuation) (Lingard et al., 2016). Parents receive the report approximately five months after the test is completed and schools receive student data approximately four months after testing.

This section investigates the impact of the feedback from NAPLAN, as perceived by the students. Students were asked: ‘Have you ever seen your NAPLAN results in the past?’ There were three possible responses: ‘I’ve never seen my results’, ‘My parents showed me’ and ‘My teacher showed me’. Table 11 shows students’ responses to this survey item.

Table 11

<table>
<thead>
<tr>
<th>Student perceptions of the source of NAPLAN feedback</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>My parents showed me</td>
<td>161</td>
<td>78%</td>
</tr>
<tr>
<td>I’ve never seen my results</td>
<td>25</td>
<td>12%</td>
</tr>
<tr>
<td>My teacher showed me</td>
<td>20</td>
<td>10%</td>
</tr>
</tbody>
</table>
Most students reported that they have seen their NAPLAN results; 78% had been shown by their parents and 10% by their teachers. Interestingly, a small percentage of students have not seen their NAPLAN results. The next logical step was to investigate how the students, parents and teachers use the results.

4.1.5.1 Uses of feedback

Students were asked, ‘Have your parents ever spoken to you about your NAPLAN results?’ The possible responses included ‘No’, ‘Yes, to point out areas to improve’, ‘Yes, to congratulate me’, ‘Yes, to highlight the importance of the test’ and ‘Yes, to suggest I work harder’. Table 12 shows the results of this survey item.

Table 12

<table>
<thead>
<tr>
<th>Student perceptions of the feedback about NAPLAN from Parents. Students were able to select more than one answer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>To point out areas improve</td>
</tr>
<tr>
<td>To congratulate</td>
</tr>
<tr>
<td>Highlight importance</td>
</tr>
<tr>
<td>Suggest I work harder</td>
</tr>
</tbody>
</table>

While the majority of students had been shown their NAPLAN results by their parents, most students (59 per cent) had not discussed their results with parents. Of those who did discuss their NAPLAN results at home, 16 per cent reported that the results were used to identify areas of improvement. 14 per cent reported that their parents congratulated them and six percent reported that their parents highlighted the importance of NAPLAN. Five per cent
of students reported that their parents used their NAPLAN results to suggest they work harder. In spite of the fierce public debates around NAPLAN, there appears to be relatively little discussion of the results between parents and students.

Despite 88 per cent of students seeing their NAPLAN results, 59 per cent of students had not spoken with their parents about them. This suggests that NAPLAN gives feedback to parents, but this feedback is not reaching students at the individual level, from their parents at least. Students were also asked: ‘Have you ever spoken to your teachers about your NAPLAN results?’ Table 13 shows the responses to this survey item.

Table 13 Student perceptions of the feedback about NAPLAN from Teachers. Students were able to select more than one answer.

<table>
<thead>
<tr>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>142</td>
</tr>
<tr>
<td>To point out areas improve</td>
<td>38</td>
</tr>
<tr>
<td>To congratulate</td>
<td>13</td>
</tr>
<tr>
<td>Highlight importance</td>
<td>3</td>
</tr>
<tr>
<td>Suggest I work harder</td>
<td>10</td>
</tr>
</tbody>
</table>

Most students (69%) had not spoken to their teachers about their NAPLAN results. Of those who had, the most common response was to point out areas of improvement. This is similar to the feedback from parents, with feedback from the test not reaching students at the individual level.

4.1.6 Conclusions about student responses to NAPLAN

Most students perceived that their doing well in NAPLAN was more important to their teachers and their parents than themselves. In exploring how students respond to
NAPLAN, the responses varied from positive to strongly negative. The most common responses were boredom (52 per cent) and feeling nervous (48 per cent), followed by several positive responses, including feeling confident (29 per cent) and enthusiastic (19 per cent). There was a small group of students who reported negative physical responses, including feeling sick, sleeplessness or crying.

Regarding feedback, most students have seen their NAPLAN results, primarily from their parents but few discuss the results. The limited stakes and feedback associated with the test may have unintended consequences.

Many of the studies examining the impact of high stakes testing regimes do not identify if the observed effects are due to the specific testing regime or are more generally associated with all forms of assessment. This study sought further insights around this by exploring student responses to internal school tests.

4.2 Student perceptions of internal testing

This study aimed to differentiate between student responses to NAPLAN and internal tests. For this study, internal tests were defined as any type of assessment task created and completed within the school (e.g., mathematics tests, science tests and English assessments). Amongst the uses of such assessments are giving feedback to students and determining grade allocations.

We hypothesise that at the individual level, internal tests may have higher stakes for students than NAPLAN. Section 4.2.1–4.2.3 discusses students’ perceptions of expectancy, value and motivation, and their emotional responses, regarding internal tests. A direct comparison between responses to the two types of assessment can be found in Chapter 6.

4.2.1 Student responses to internal testing

Previous research has found evidence of students having negative emotional and physical responses to high-stakes testing such as NAPLAN (Dulfer et al., 2012; Thompson,
This section aims to identify if internal testing also causes these negative responses, given that there appears to be little in the literature that explores student responses to internal tests. Participants were asked a question about internal tests that matched the question about NAPLAN, ‘Did you experience any of the following before your last internal test?’ For each of the responses in the list, participants could choose ‘Yes’ or ‘No’. It was possible to choose ‘Yes’ for more than one of the listed responses. Responses were grouped into three categories: positive, negative emotional and negative physical. Table 14 shows student reported responses to internal tests.

**Table 14**

**Student responses to internal tests**

<table>
<thead>
<tr>
<th>Response</th>
<th>Number (out of 206)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeling nervous</td>
<td>118</td>
<td>57%</td>
</tr>
<tr>
<td>Confident</td>
<td>88</td>
<td>43%</td>
</tr>
<tr>
<td>Stressed</td>
<td>71</td>
<td>34%</td>
</tr>
<tr>
<td>Bored</td>
<td>51</td>
<td>25%</td>
</tr>
<tr>
<td>Excited</td>
<td>52</td>
<td>25%</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>44</td>
<td>21%</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>21</td>
<td>10%</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>17</td>
<td>8%</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>15</td>
<td>7%</td>
</tr>
<tr>
<td>Crying</td>
<td>3</td>
<td>2%</td>
</tr>
</tbody>
</table>

*Note.* Students could answer yes to more than one response, therefore percentages sum to more than 100%.
The most commonly reported response was feeling nervous (57 per cent), followed by feeling confident (43 per cent). It is understandable that students feel nervous before tests, but it was interesting to note that many students also felt confident going into the assessment. There were similar numbers of students who reported being either bored (25 per cent) or excited (25 per cent) by the assessment. The total number of endorsements of negative emotional responses was 250. This compares to 184 positive emotional responses. Endorsements of negative physical responses totalled 56. This suggests there are large numbers of students who are experiencing both negative and positive responses to internal testing.

4.2.1.1 Open-ended responses

In addition to the structured survey questions, participants were given the opportunity to respond to an open-ended question: ‘Did you have any other responses to your internal test?’ Fifty students provided an open-ended response. Of these, the majority (30 per cent) were neutral (e.g., ‘Tired’, ‘calm’ or ‘normal, just wanted to get it done’).

The second most common response was internalising behaviour (28 per cent) and included students reporting feeling ‘nervous, ‘anxious’ or ‘irritated’. The third most common response were students who reported negative emotions including feeling ‘sad’, ‘low self-confidence’, ‘scared’ or ‘forgetful’.

4.2.2 Student expectations of performance in internal tests

This section explored how students expected to perform in internal tests (i.e., their expectancy). Students were asked; to what extent they agreed or disagreed with the statement, ‘I expect to do well in my internal [English/Mathematics] tests this year’. Students responded using a Likert scale, ranging from (1) ‘Strongly disagree’ to (5) ‘Strongly agree’. Table 15 shows student expectancy regarding their performance on internal tests for English and mathematics.
Table 15

*Student Expectations of Performance on Internal Mathematics and English Tests*

<table>
<thead>
<tr>
<th></th>
<th>1: Strongly disagree</th>
<th>2: Disagree</th>
<th>3: Neither agree or disagree</th>
<th>4: Agree</th>
<th>5: Strongly agree</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>2%</td>
<td>4%</td>
<td>19%</td>
<td>43%</td>
<td>32%</td>
<td>3.68</td>
<td>0.94</td>
</tr>
<tr>
<td>Mathematics</td>
<td>4%</td>
<td>5%</td>
<td>19%</td>
<td>31%</td>
<td>43%</td>
<td>3.78</td>
<td>1.04</td>
</tr>
</tbody>
</table>

A paired sample *t* test indicated there were no statistically significant differences between student expectations in mathematics and English (*t*(205) = -0.69, *p* = .24). Students generally reported high levels of expectancy in internal tests, suggesting that students were confident in their ability to perform well on these assessments. Three quarters of students agreed or strongly agreed with the statement, ‘I expect to do well in English this year’ and a similar proportion agreed for mathematics. Conversely, 19% per cent of students neither agreed nor disagreed which suggests they may be uncertain about their performance. There was a small group of students (6–9 per cent) who either selected ‘Strongly disagree’ or ‘Disagree’ and thus, reported a low expectancy in their ability to succeed in internal tests. The expectancy–value theory predicts that students with high levels of expectancy and value would have higher levels of motivation.

4.2.3 Value placed on doing well in internal tests

We also sought to measure student perceptions of the value placed on doing well in internal tests according to three stakeholder groups: the students, their parents and their teachers, using the following 3 questions:

1. How important is it to *you* to do well on internal tests?
2. How important is it to your *parents* that you do well in internal tests?
3. How important is it to you teachers that you do well in internal tests?

Participating students were asked to rate the perceived importance of internal tests for stakeholders on a Likert scale, ranging from (1) ‘Not important at all’ to (5) ‘Very important’. The primary outcome variable was the overall sample means for students’ perceptions of value to self, parents and teachers.

Table 16 shows the perceptions of the value placed on internal tests by each stakeholder group.

Table 16

*How important is it to you(r) [teachers/parents/students] for you to do well in your internal tests?*

<table>
<thead>
<tr>
<th></th>
<th>1: Not important at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers</td>
<td>1%</td>
<td>1%</td>
<td>11%</td>
<td>34%</td>
<td>53%</td>
<td>4.39</td>
<td>0.82</td>
</tr>
<tr>
<td>Parents</td>
<td>1%</td>
<td>1%</td>
<td>7%</td>
<td>35%</td>
<td>56%</td>
<td>4.49</td>
<td>0.74</td>
</tr>
<tr>
<td>Students</td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>28%</td>
<td>60%</td>
<td>4.46</td>
<td>0.85</td>
</tr>
</tbody>
</table>

The overwhelming majority of students perceived that they, their parents and their teachers see it as ‘important’ or ‘very important’ to perform well in internal tests. Very few students reported that it was ‘not at all important’ to them, their teachers or their parents to perform well in such tests. A paired samples *t* test indicated that there were no statistically significant differences between the perceived value placed on internal tests by teachers, parents and students. This demonstrates a consistent view by students that all three stakeholders place a high value on doing well in internal tests. Very few students (1-2%), 2-4
students in total, reported that doing well on internal tests was not important or not important at all.

4.2.4 Student motivation for internal tests

This section reports on students’ motivation for internal tests. To measure motivation, students were asked, ‘How hard did you try in internal [mathematics/English] tests?’ They responded on a Likert scale, ranging from (1) ‘I didn’t try at all’ to (5) ‘I put in my best effort’. Table 17 shows student reported motivation for internal tests.

Table 17

<table>
<thead>
<tr>
<th>Student motivation for internal English/Mathematics Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: I didn’t try at all</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>Maths</td>
</tr>
</tbody>
</table>

Consistent with expectancy–value theory, students overwhelmingly reported that they applied their best effort to their internal tests. There was a very small percentage of students (1–3 per cent) who reported not being motivated for internal tests.

Section 4.2.6 discusses students’ perception of the feedback from internal tests.

4.2.5 Feedback from internal tests

Student were asked; ‘Have your parents ever spoken to you about your results in internal tests?’ The possible responses included ‘No’, ‘Yes, to point out areas to improve’, ‘Yes, to congratulate me’, ‘Yes, to highlight the importance of the test’ and ‘Yes, to suggest I work harder’. Table 18 shows student perceptions of the feedback on internal tests from parents.
Table 18

*Student Perceptions of Feedback on Internal Tests from Parents*

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>13</td>
<td>6%</td>
</tr>
<tr>
<td>To point out areas improve</td>
<td>55</td>
<td>27%</td>
</tr>
<tr>
<td>To congratulate</td>
<td>78</td>
<td>38%</td>
</tr>
<tr>
<td>Highlight importance</td>
<td>19</td>
<td>9%</td>
</tr>
<tr>
<td>Suggest I work harder</td>
<td>41</td>
<td>20%</td>
</tr>
</tbody>
</table>

As can be seen in Table 18, almost all students (94%) had parents speak to them about their internal tests. A very small percentage of students had never been spoken to by their parents about internal tests. The most common feedback from parents was to congratulate students (38 per cent), followed by pointing out areas of improvement (27 per cent). The reasons for the increased feedback on internal tests (as compared to NAPLAN) may be due to the shorter time between completing an internal test and receiving feedback. In addition, the larger number of internal tests may increase the likelihood that parents will discuss results with students.

Students were also asked ‘Have your teachers ever spoken to you about your internal tests? Table 19 shows the student perceptions of the feedback on internal tests by teachers.
Table 19

Student Perceptions of Feedback on Internal Tests from Teachers

<table>
<thead>
<tr>
<th>Feedback Type</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>38</td>
<td>18%</td>
</tr>
<tr>
<td>To point out areas improve</td>
<td>107</td>
<td>52%</td>
</tr>
<tr>
<td>To congratulate</td>
<td>36</td>
<td>18%</td>
</tr>
<tr>
<td>Highlight importance</td>
<td>6</td>
<td>3%</td>
</tr>
<tr>
<td>Suggest I work harder</td>
<td>19</td>
<td>9%</td>
</tr>
</tbody>
</table>

The great majority of students (82%) reported that they had received feedback from teachers about their internal tests. The most commonly reported use of these tests by teachers was to point out areas of improvement. That is, the internal tests are being used as a learning tool for students and my findings indicate that most students are receiving feedback from internal tests.

We have identified student perceptions of expectancy and motivation towards internal tests and the perceived value various stakeholders place on the students doing well in internal tests. Further, we have identified students’ emotional responses to internal tests.

4.3 NAPLAN and Its Impact on School Experiences

The general impact of national testing programs on students’ experiences of school identified in the literature includes a narrowing of curriculum (Au, 2007; Klenowski, 2011), a decrease in time on non-testable subjects (Jones et al., 2003; Lobascher, 2011; Polesel et al., 2012), a focus on basic content (Au, 2007) and closer alignment with state curriculum guidelines (Polesel et al., 2012).

Students were asked two questions to determine the influence of NAPLAN on the curriculum. The first question sought to examine the perceived value the school placed on
NAPLAN, while the second sought to examine the influence of NAPLAN on how time was spent in school by seeking information on the amount of NAPLAN practice students had been required to undergo.

Their responses to the first question, ‘How much emphasis does your school place on doing well in NAPLAN?’ are shown in Table 20.

Table 20

*Student perceptions of the value their school places on doing well in NAPLAN*

<table>
<thead>
<tr>
<th>Student response</th>
<th>1: Not important at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student response</td>
<td>4%</td>
<td>7%</td>
<td>26%</td>
<td>46%</td>
<td>17%</td>
<td>3.65</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Most students believed that their schools placed a moderate amount of emphasis on NAPLAN. Only four per cent of students reported the school placing ‘Very little’ emphasis on doing well in NAPLAN. Another way to identify the emphasis schools place on NAPLAN is to identify how many practice tests students have completed.

Students were asked if they had completed any practice tests for NAPLAN at school or at home. There were four possible responses ‘No, 1–2 tests’, ‘3–5 tests’ or ‘6 or more’. Table 21 shows the number of practice tests students reported that they had completed at school.
Of the students in this sample, 50 per cent reported having completed 1-2 practice tests. Only 6 per cent of students completed six or more tests and 25 per cent of students reported that they had never completed a practice test at school.

In addition to the practice tests completed at school, there is an emerging business of NAPLAN practice tests being available for sale to parents. To explore the prevalence of students being involved in NAPLAN practice outside school, students were asked ‘Have you completed any NAPLAN workbooks at home?’ Table 22 presents students’ responses to this survey item.

Table 22

<table>
<thead>
<tr>
<th>Number of Students who have Completed NAPLAN Workbooks at Home</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
</tbody>
</table>

The majority of students reported that they did not complete NAPLAN workbooks at home, with only one in five students surveyed reporting having completed practice tests at home.
The impact on the curriculum has also been found internationally through the practice of schools cancelling some classes to practise for national tests, as has been the case in the United States (Sabol, 2010). Students were asked, ‘Has your school ever cancelled classes to give you class time to practise for NAPLAN?’ The results for this survey item are presented in Table 23.

Table 23

<table>
<thead>
<tr>
<th>Number of classes cancelled to practise for NAPLAN</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>130</td>
<td>63%</td>
</tr>
<tr>
<td>Once or twice</td>
<td>65</td>
<td>32%</td>
</tr>
<tr>
<td>Three or more times</td>
<td>11</td>
<td>5%</td>
</tr>
</tbody>
</table>

Within the sample, 63 per cent of students reported that they have never had a class cancelled to practise for NAPLAN. Almost one in three students (32 per cent) say they have had classes cancelled once or twice and a small group (5 per cent) of students stated they have had classes cancelled three or more times. The variability across the group (given only four schools participated in the study) suggests that students are making different judgements as to whether classes have been cancelled, or that some classes spend more time practising for the tests than others.

Students were also asked, ‘How are NAPLAN results used at your school?’ The results of this survey item are presented in Table 24.
Table 24

*Student understanding of the uses of NAPLAN in their school*

<table>
<thead>
<tr>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don’t know</td>
<td>160</td>
</tr>
<tr>
<td>They aren’t used at all</td>
<td>30</td>
</tr>
<tr>
<td>To decide who can enrol in subjects</td>
<td>10</td>
</tr>
<tr>
<td>To place students in classes</td>
<td>6</td>
</tr>
</tbody>
</table>

The data showed that the majority (78%) of students do not understand how schools use NAPLAN data. A further 15% believe NAPLAN results are not used at all. The large percentage of students who are unsure how the NAPLAN test results are used or who believe that they are not used could be linked to the high levels of boredom students report for the tests.

This study was designed to compare students’ perceptions of the use of NAPLAN and internal results. Table 25 highlights students’ response to the question ‘How are results from your internal tests used in your school?’

Table 25

*Student Perceptions of the Use of Internal Test Results*

<table>
<thead>
<tr>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don’t know</td>
<td>123</td>
</tr>
<tr>
<td>They aren’t used</td>
<td>11</td>
</tr>
<tr>
<td>To decide who can enrol in subjects</td>
<td>32</td>
</tr>
<tr>
<td>To place students in classes</td>
<td>40</td>
</tr>
</tbody>
</table>
Surprisingly, most students (60%) did not know how internal tests results were used. The use of internal tests to determine whether students can enrol in particular subjects was perceived by 15 per cent of students and 20 per cent of students thought tests were used to place students in classes.

After exploring attitudes to both types of assessment, students were asked, ‘Which statement is most correct for you?’ There were three options: ‘Internal tests are more important to me than NAPLAN, NAPLAN tests are more important to me than internal tests, both are equally important to me’. Results of this item are shown in Table 26.

Table 26

*Student perceptions of the importance of assessment tests.*

<table>
<thead>
<tr>
<th>Statement</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal tests are more important to me than NAPLAN</td>
<td>121</td>
<td>59%</td>
</tr>
<tr>
<td>Both are equally important to me</td>
<td>75</td>
<td>36%</td>
</tr>
<tr>
<td>NAPLAN tests are more important to me than internal tests</td>
<td>10</td>
<td>5%</td>
</tr>
</tbody>
</table>

Most students (59 per cent) identified that internal tests were more important to them than NAPLAN. Only five per cent of students considered NAPLAN to be the more important assessment activity at school. The remaining students (37 per cent) perceived both forms of assessment to be equally important. Students’ perception of the importance of internal tests may be due to the frequency of internal tests, the increased stakes applied to the tests and the relevance of the testable material to their students’ day-to-day work in schools. A Year 9 student made the following summary regarding assessments, “NAPLAN sucks. School-based tests are better, they have more purpose”.
4.4 Summary

This chapter investigated student attitudes to two forms of assessment: NAPLAN and internal tests. Students reported that they perceived their teachers and parents placed higher value on doing well in NAPLAN than they did themselves. Students on average had a moderate expectancy about performing well on the NAPLAN tests. The motivation to apply their best effort to NAPLAN was relatively high.

Students were also asked about their response to NAPLAN. The most commonly reported responses were boredom, feeling nervous, stressed and confident. Feedback from students about NAPLAN suggests that 80 per cent of students see their results, but most students had not been spoken to by their parents or their teachers about their NAPLAN results. Parents and teachers were much more likely to have spoken with students about their results on internal tests than NAPLAN.

Students placed a high value on doing well in internal tests and perceived that teachers and parents also placed a high value on doing well in them. There was no significant difference between students’ expectancies in English and mathematics internal tests. Reported motivation for internal tests was high. The most common responses to the internal tests were nervousness, confidence and stress. Most students had spoken to their parents about their results in internal tests.

Regarding NAPLAN’s impact on their learning experiences, 80 per cent of students stated they had completed some practice tests at school and 20 per cent of students had completed practice tests at home. Most students (63 per cent) had not had classes cancelled to practise for NAPLAN.

The data suggest that students are unaware of the purpose of NAPLAN, as 80 per cent of students reported not knowing how NAPLAN data are used.
While this chapter has explored student perceptions of NAPLAN and internal tests separately, Chapter 5 investigates the similarities and differences between student attitudes and responses to these two forms of assessment.
Chapter 5: Comparing Responses to NAPLAN and Internal Tests

Chapter 4 identified that responses to internal tests vary between students. This section investigates the difference in participants’ reactions to, and perceptions of, national testing and internal tests. Other studies have drawn strong negative conclusions about student responses to NAPLAN and other forms of national testing (e.g. Howell, 2012; Wyn et al., 2014). However, it may be that these responses are related to the testing experience, rather than high stakes testing per se. This chapter sought to determine if this is the case by considering differences in students’ perceptions and reactions to NAPLAN and internal tests. Hence, the types of assessments being compared in this study are limited to test formats and as such, the responses cannot be generalised to all modes of internal assessment tasks. This study limits its comparison to that between the NAPLAN tests and internal tests.

5.1 A Comparison of Student Responses to NAPLAN and Internal Tests

While many studies have found negative responses to high-stakes testing programs (Brown et al., 2004; Flores & Clark, 2003; Lewis, 2000; Madaus et al., 2009; Paris & McEvoy, 2000; Rice et al., 2015) it is not clear whether these are students’ responses to a national testing program, or to testing more generally. Therefore, this study compares student responses to internal tests and NAPLAN to ascertain if there is a difference between how students respond to internal tests and high-stakes assessment, and whether negative responses as just as prevalent in internal testing situations.

Students were asked: ‘Think back to the last time you completed a [NAPLAN/Internal] test, did you experience any of the following?’ Table 27 shows the results to these survey items. Probability estimates are rounded to three decimal places.
Table 27

Students’ Reported Responses to NAPLAN and Internal Tests

<table>
<thead>
<tr>
<th></th>
<th>NAPLAN</th>
<th>Internal tests</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeling nervous</td>
<td>48%</td>
<td>57%</td>
<td>.025*</td>
</tr>
<tr>
<td>Confident</td>
<td>29%</td>
<td>43%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Stressed</td>
<td>27%</td>
<td>34%</td>
<td>.044*</td>
</tr>
<tr>
<td>Bored</td>
<td>52%</td>
<td>25%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>17%</td>
<td>25%</td>
<td>.005**</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>18%</td>
<td>21%</td>
<td>.243</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>13%</td>
<td>10%</td>
<td>.424</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>11%</td>
<td>8%</td>
<td>.701</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>10%</td>
<td>7%</td>
<td>.332</td>
</tr>
<tr>
<td>Crying</td>
<td>3%</td>
<td>2%</td>
<td>.250</td>
</tr>
</tbody>
</table>

Note. Students were able to select more than one answer so percentages sum to more than 100%

An exact McNemar’s test determined that there was a statistically significant difference in students’ responses to NAPLAN and internal tests for feeling nervous, confident, stressed, bored and excited.

Almost half (43 per cent) of students in the sample reported feeling confident before their internal tests, compared to 29 per cent reporting confidence prior to NAPLAN. There were also larger numbers of students reporting excitement for internal tests than for NAPLAN. There were no statistically significant differences between students being enthusiastic.

In negative responses, there was a small but statistically significant difference between students feeling nervous in NAPLAN and internal tests, with fewer students reporting feeling nervous for NAPLAN than internal tests. Students were also more likely to
report feeling stressed during internal testing situations than for NAPLAN. It would appear that internal testing may cause more negative responses than NAPLAN.

The greatest difference in the reported results was that secondary school students were much more likely to report being bored with NAPLAN, with almost double the number of students reporting that they were bored during NAPLAN compared to internal tests ($p < .001$). This may be due to the fact that fewer students understand the purpose of NAPLAN and the stakes attached to doing well appear to be low for the students (although not necessarily for their teachers or schools).

A small percentage of students reported negative physical responses to both types of assessment. In this study, there were no significant differences in the negative physical responses reported regarding NAPLAN and internal tests. This also suggests that NAPLAN itself might not cause negative physical responses; it may be assessment more generally. Further, if the type of assessment does not change the likelihood of the negative physical response, perhaps there are individual characteristics that make the student more likely to experience these responses.

It may be that for approximately four per cent of students, there are negative physical responses to testing situations, whether this is for NAPLAN or internal tests.

5.1.1 Differences in open-ended responses

Students were also asked the open-ended question, ‘Are there any other feelings you had before the [NAPLAN/Internal Tests]?’ The percentage of students who gave an open-ended response was 24% for the item about NAPLAN and 28% for the item about internal tests. Responses were analysed thematically. Table 28 shows the categories of responses to the open-ended questions for NAPLAN and internal tests.
Table 28

*Differences in Open-Ended Responses to NAPLAN and Internal tests. Percentages of the entire sample.*

<table>
<thead>
<tr>
<th></th>
<th>NAPLAN</th>
<th>Internal tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apathy</td>
<td>7%</td>
<td>6%</td>
</tr>
<tr>
<td>Internalising behaviour</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>Negative responses</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>Concerns around performance</td>
<td>5%</td>
<td>8%</td>
</tr>
<tr>
<td>Avoidance</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>Positive</td>
<td>2%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Table 28 shows that students reported similar concerns about performance for both tests. An independent samples *t*-test was not run as the cell sizes were too small for an adequately powered analysis. There were more students reporting negative emotional responses (including feeling depressed, scared or sad) for NAPLAN, compared with internal tests. The other key difference was a larger percentage of students reporting apathy (including feeling bored, tired or ‘didn’t care’) for NAPLAN than internal tests. After identifying students’ responses to the tests, we will now compare students’ experience of the feedback from each type of test (see Section 5.4).

5.2 Comparing Value in NAPLAN and Internal Tests.

Students were asked how they, their parents and their teachers perceived NAPLAN and internal tests. Participants were asked, ‘How important is it to your teachers, for you to do well in [NAPLAN/Internal tests]?’ Table 29 shows the different values that students perceive teachers to possess, regarding NAPLAN and internal tests.
Table 29

Students’ Perception of Teacher Value on NAPLAN Compared to Internal Tests

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all important</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers (NAPLAN)</td>
<td>6%</td>
<td>9%</td>
<td>21%</td>
<td>33%</td>
<td>30%</td>
<td>3.72</td>
<td>1.16</td>
</tr>
<tr>
<td>Teachers (Internal)</td>
<td>2%</td>
<td>1%</td>
<td>10%</td>
<td>31%</td>
<td>56%</td>
<td>4.39</td>
<td>0.83</td>
</tr>
</tbody>
</table>

A paired sample $t$ test found a significant difference between student perceptions of teacher value of NAPLAN ($M = 3.72$, $SD = 1.16$) and internal tests ($M = 4.39$, $SD = 0.83$, $t(205) = -6.73$, $p < 0.001$, $d = 0.66$). Students perceived their teachers placed far more value on internal tests than NAPLAN. Most students (56 per cent) believed that teachers saw it as very important that they perform well on internal tests. Compared to this, only 31 per cent of students perceived that their teachers saw it as very important for them to perform well in NAPLAN. This suggests that the message students receive from teachers is that internal tests hold more value than NAPLAN.

The next section compares how students perceive parent value of the two forms of assessment. Participants were asked, ‘How important is it to your parents, for you to do well in [NAPLAN/Internal tests?]’ Table 30 outlines the differences in student perceptions of parent value regarding NAPLAN and internal tests.
Table 30

*Students’ Perception of Parent Value on NAPLAN Compared to Internal Tests*

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents (NAPLAN)</td>
<td>7%</td>
<td>13%</td>
<td>16%</td>
<td>29%</td>
<td>34%</td>
<td>3.63</td>
<td>1.31</td>
</tr>
<tr>
<td>Parents (Tests)</td>
<td>1%</td>
<td>1%</td>
<td>7%</td>
<td>35%</td>
<td>56%</td>
<td>4.49</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*Note.* The paired sample *t* test showed a significant difference at the *p* < .001 level.

A paired sample *t* test found that there was a significant difference between student perceptions of parent value of NAPLAN (*M* = 3.63, *SD* = 1.31) and internal tests (*M* = 4.49, *SD* = 0.74, *t*(205) = -8.23, *p* < .001, *d* = 0.81). Students perceived that their parents placed more value on doing well internal tests than on NAPLAN. Students were far less likely to indicate that their parents perceived internal tests to be ‘Not important at all’ compared to NAPLAN. Only one per cent of students reported that doing well on internal tests was ‘Not important at all’ to their parents, while seven per cent reported this in relation to NAPLAN. However, overall the majority of students (63%) still believed that their doing well in NAPLAN was important or very important to their parents.

The next section compares the value students themselves placed on doing well in NAPLAN and internal tests. Participants were asked, ‘How important is it to you, for you to perform well on NAPLAN/internal tests?’ Table 31 shows the results of these survey items.
Table 31

Comparison of Student Value of NAPLAN and Internal Tests

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>12%</td>
<td>10%</td>
<td>23%</td>
<td>31%</td>
<td>24%</td>
<td>3.32</td>
<td>1.37</td>
</tr>
<tr>
<td>Student (Tests)</td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>28%</td>
<td>60%</td>
<td>4.46</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Students reported a placing a higher value on internal tests than NAPLAN. A paired t test found that there was a significant difference between student value for doing well in NAPLAN (M = 3.32, SD = 1.37) and student value for doing well in internal tests (M = 4.46, SD = 0.84, t(205) = -10.17, p < 0.001, d = 1.00). Most students (60 per cent) reported that it was ‘Very important’ to them to perform well in internal tests, which was more than double the percentage of students (24 per cent) who identified doing well in NAPLAN as ‘Very important’. One in eight students selected the ‘Not important at all’ option regarding NAPLAN, compared to only three per cent of students choosing the ‘Not important’ option for internal tests. These findings align with students’ perceptions of the value placed on both forms of assessment by their teachers and parents, suggesting that students believe that all stakeholders place a higher value on internal tests than NAPLAN.

Participants in this study perceived that parents and teachers place considerably more value on doing well in internal tests than in NAPLAN. The students also placed a higher value on doing well in internal tests than in NAPLAN. Expectancy-value theory would suggest that, given the lower value placed on doing well in NAPLAN by students, student motivation and effort will be lower for NAPLAN than for internal tests. We turn to consider the differences between students’ expectation for NAPLAN and internal tests.
5.3 Comparing Expectations in NAPLAN and Internal Tests.

In addition to comparing the value placed on tests by stakeholders, we aimed to compare the expectations of the students participating in the tests. Using the expectancy–value theory, students’ motivation will be a function of the value placed on the tests and their expectancies for doing well (Eccles & Wigfield, 2002). Students were asked to what extent they agreed with the following statements:

- I expect to do well in NAPLAN mathematics this year
- I expect to do well in NAPLAN writing this year
- I expect to do well in NAPLAN reading this year
- I expect to do well in NAPLAN language conventions this year

Students responded on a scale from (1) ‘Strongly disagree’ to (5) ‘Strongly agree’ and a paired samples t test was used to calculate statistical significance. To improve comparability, the three NAPLAN English domains were averaged to create a student expectancy in English (composite) variable. This was done as internal English assessments often assess multiple domains including reading, writing and language conventions. Table 32 shows the expectancy in mathematics for NAPLAN and internal tests and Table 33 shows the average expectancy of a student across the three domains in English. The results are presented in Tables 32 and 33.
Table 32

*Student Expectancy in Mathematics for NAPLAN and internal tests*

<table>
<thead>
<tr>
<th></th>
<th>1:</th>
<th>2:</th>
<th>3:</th>
<th>4:</th>
<th>5:</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly disagree</td>
<td>Disagree</td>
<td>Neither Agree or Disagree</td>
<td>Agree</td>
<td>Strongly agree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maths</td>
<td>5%</td>
<td>7%</td>
<td>23%</td>
<td>34%</td>
<td>31%</td>
<td>3.76</td>
<td>1.11</td>
</tr>
<tr>
<td>(NAPLAN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maths (Internal)</td>
<td>4%</td>
<td>4%</td>
<td>19%</td>
<td>31%</td>
<td>43%</td>
<td>4.02</td>
<td>1.04</td>
</tr>
</tbody>
</table>

A paired sample t test found students reported significantly higher levels of expectancy in internal mathematics tests (M = 4.02, SD = 1.04) than in NAPLAN mathematics (M = 3.76, SD = 1.11, t(205) = -2.52, p = 0.006, d = 0.24).

Table 33

*Student Expectancy in English (composite) for NAPLAN and internal tests*

<table>
<thead>
<tr>
<th></th>
<th>1:</th>
<th>2:</th>
<th>3:</th>
<th>4:</th>
<th>5:</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly disagree</td>
<td>Disagree</td>
<td>Neither Agree or Disagree</td>
<td>Agree</td>
<td>Strongly agree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>4%</td>
<td>6%</td>
<td>27%</td>
<td>42%</td>
<td>22%</td>
<td>3.68</td>
<td>1.02</td>
</tr>
<tr>
<td>(NAPLAN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English (Internal)</td>
<td>2%</td>
<td>4%</td>
<td>17%</td>
<td>40%</td>
<td>29%</td>
<td>3.93</td>
<td>0.94</td>
</tr>
</tbody>
</table>
A paired sample $t$ test found students reported significantly higher levels of expectancy in internal English tests ($M = 3.93, SD = 0.94$) than in NAPLAN English ($M = 3.68, SD = 1.02, t(205) = -2.51, p = 0.006, d = 0.25$). Further analysis showed that most students (65 per cent) reported the same expectancy for both tests. Approximately one in four (27 per cent) students expected to perform better on internal tests than NAPLAN. Very few students expected to perform better in NAPLAN than in internal tests (8 per cent). This shows that positive student expectations are skewed towards internal tests.

Expectancy-value theory would predict that if an individual’s expectancy to do well and value placed on doing well are higher, their motivation will be higher. It was therefore hypothesised that student motivation as judged by reported effort would be higher for internal tests than for NAPLAN. Student motivation was compared for the two forms of testing by asking students to respond on a five-point Likert scale ranging from (1) ‘I didn’t try at all’ to (5) ‘I put in my best effort’. The question students were asked to respond to was ‘How hard did you try in [NAPLAN/internal] tests this year?’ The motivation for NAPLAN (English) section reports the average motivation across the three English domains. Table 34 shows the results for mathematics and Table 35 shows the results for English.
Table 34

**Motivation for Mathematics Assessment**

<table>
<thead>
<tr>
<th></th>
<th>1: I didn’t try</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics (NAPLAN)</td>
<td>5%</td>
<td>5%</td>
<td>12%</td>
<td>32%</td>
<td>47%</td>
<td>4.05</td>
<td>1.13</td>
</tr>
<tr>
<td>Mathematics (Internal)</td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>24%</td>
<td>65%</td>
<td>4.52</td>
<td>0.82</td>
</tr>
</tbody>
</table>

A paired sample t-test found students reported significantly higher motivation for internal mathematics tests (M = 4.52, SD = 0.82) than NAPLAN mathematics (M = 4.05, SD = 1.13, t(205) = -4.83, p < 0.001, d = 0.48).

Table 35

**Motivation for English Assessment**

<table>
<thead>
<tr>
<th></th>
<th>1: I didn’t try</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>English (NAPLAN)</td>
<td>4%</td>
<td>4%</td>
<td>17%</td>
<td>27%</td>
<td>49%</td>
<td>4.09</td>
<td>1.12</td>
</tr>
<tr>
<td>English (Internal)</td>
<td>1%</td>
<td>2%</td>
<td>8%</td>
<td>32%</td>
<td>56%</td>
<td>4.39</td>
<td>0.86</td>
</tr>
</tbody>
</table>
A paired sample t-test found a similar results in English, with students reporting significantly higher motivation for internal English tests ($M = 4.39$, $SD = 0.86$) than NAPLAN English ($M = 4.09$, $SD = 1.12$, $t(205) = -3.45$, $p < 0.001$, $d = 0.30$). Students were more likely to have put in their best effort for internal tests than for NAPLAN. In internal mathematics tests, 65 per cent of students reported that they put in their best effort, compared to 47 per cent of students reporting the same for NAPLAN. The pattern is the same in students’ reporting of English, in which 49 per cent of students reported that they put in their best effort for NAPLAN compared to 56 per cent reporting that they put in their best effort for internal English test. These findings were consistent with what would be predicted by the expectancy–value theory. Students place more value on doing well in internal tests than doing well in NAPLAN and therefore, are more motivated for and put in more effort for internal tests.

Students in this study demonstrated higher expectancies to do well, value and motivation for internal tests than NAPLAN. We will now compare the feedback students receive from both types of assessment.

5.4 Comparing Feedback from NAPLAN and Internal Tests

An essential element of any assessment program is the feedback it gives to stakeholders. This section compares the feedback students receive from parents and teachers in relation to the two types of assessment. Students were asked, ‘Have your parents ever spoken to you about your [NAPLAN/internal] test results?’ There were five options, as shown in Table 36. Table 36 below compares student responses regarding the feedback on their results from parents in relation to the two types of assessment.
Table 36

<table>
<thead>
<tr>
<th></th>
<th>NAPLAN</th>
<th>Internal</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>59%</td>
<td>6%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>To point out areas improve</td>
<td>16%</td>
<td>27%</td>
<td>.005**</td>
</tr>
<tr>
<td>To congratulate</td>
<td>14%</td>
<td>38%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Highlight importance</td>
<td>6%</td>
<td>9%</td>
<td>.286</td>
</tr>
<tr>
<td>Suggest I work harder</td>
<td>5%</td>
<td>20%</td>
<td>&lt;.001***</td>
</tr>
</tbody>
</table>

An exact McNemar’s test determined that there were statistically significant differences between the feedback students received from parents regarding NAPLAN and internal tests. Students were significantly more likely to report discussing the results of their internal tests with their parents than their NAPLAN results. Only 31 per cent of students stated they have discussed their NAPLAN results with their parents, compared to 94 per cent of students reporting that they have discussed their results from internal tests.

Students also receive feedback on their assessments from their teachers. To compare feedback from different stakeholders, students were asked, ‘Have your teachers ever spoken to you about your [NAPLAN/Internal test] results?’ and were presented with five responses options, as shown in Table 37. Table 37 compares the feedback students have received from their teachers as a result of the two types of assessment.
Table 37

Student perceptions of the Feedback on NAPLAN and Internal test from Teachers

<table>
<thead>
<tr>
<th></th>
<th>NAPLAN</th>
<th>Internal</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>65%</td>
<td>17%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>To point out areas improve</td>
<td>21%</td>
<td>52%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>To congratulate</td>
<td>7%</td>
<td>20%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Highlight importance</td>
<td>2%</td>
<td>4%</td>
<td>.453</td>
</tr>
<tr>
<td>Suggest I work harder</td>
<td>5%</td>
<td>8%</td>
<td>.049</td>
</tr>
</tbody>
</table>

Note. Students were able to select more than one answer so percentages can sum to more than 100%

An exact McNemar’s test determined that there was a statistically significant difference in the percentage of students reporting receiving feedback from teachers regarding NAPLAN and internal tests. Eighty-three percent of students reported that teachers gave them feedback about internal tests, compared to only 35% receiving feedback about NAPLAN. In addition, students were more likely to report that they received feedback that focused on identifying areas of improvement from internal testing, than NAPLAN.

As noted in previous chapters, students were asked if they knew how NAPLAN and internal results were used in their school. There were five response options, as shown in Table 38. Table 38 compares how students perceive the results from NAPLAN and internal tests are used at their school.
Table 38

*Student understanding of the use of Results from NAPLAN and Internal tests*

<table>
<thead>
<tr>
<th></th>
<th>NAPLAN</th>
<th>Internal</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don’t know</td>
<td>78%</td>
<td>60%</td>
<td>.027**</td>
</tr>
<tr>
<td>To decide who can enrol in particular subjects</td>
<td>5%</td>
<td>16%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>They aren’t used at all</td>
<td>15%</td>
<td>11%</td>
<td>.003**</td>
</tr>
<tr>
<td>To place students in classes</td>
<td>3%</td>
<td>20%</td>
<td>&lt;.001***</td>
</tr>
</tbody>
</table>

*Note.* Students were able to select more than one answer so percentages sum to more than 100%

An exact McNemar’s test determined that there was a statistically significant difference between students’ understanding about how NAPLAN and internal test results were used. We can assume that students would know that internal results were used for determining a student’s grade, but beyond this, students generally did not appear to know a great deal about how the results from their internal tests are used. The responses showed there is a statistically significant difference between students’ knowledge of how NAPLAN and internal test results are used, with students being less likely to know how NAPLAN results are used \((p = .027)\). Approximately 20 per cent of students reported that internal results are used to place students in various classes. Only a small percentage of students reported that NAPLAN had an impact on admission into certain subjects or class placement.

5.5 *Summary*

This chapter investigated the differences between student perceptions of and responses to NAPLAN and internal tests. It found that students perceive that all stakeholders place more value on internal tests than NAPLAN. Students also reported higher levels of expectancy for doing well on internal tests than for NAPLAN. As predicted by the
expectancy–value theory, students reported higher levels of motivation for internal tests compared to NAPLAN. These findings suggest that internal tests are a much more important part of student life than NAPLAN. A further discussion of the comparison between NAPLAN and internal tests can be found in sections 7.1.2.

Chapter 6 will explore the characteristics of students who reported either positive or negative responses to testing.
Chapter 6: The Relationships between Expectancy, Value and Student Responses to Assessment

The findings to this point have shown that students have different responses to assessment. Chapter 4 showed student expectancy, value and responses to NAPLAN and internal tests, while Chapter 5 highlighted the differences between student expectancy, value and responses to the two types of assessment. This chapter aims to investigate the relationships between the value students place on testing, the expectations they have about their performance and their responses to testing. It answers the questions ‘Is student perception of teacher or parent value more closely associated with student value?’ and ‘Does academic self-concept influence student expectancy?’ It also shows how these variables relate to responses to NAPLAN, for example ‘Do students with low academic self-concept experience more negative responses?’ and ‘Do students who place a high value on NAPLAN experience more stress?’ The final section in this chapter describes the differences in responses to NAPLAN between Year 7 and Year 9 students, and the differences in responses to internal tests between Year 7 and Year 9 students. This provides specific information about how different groups and types of students respond to NAPLAN and highlights that the experience of testing will vary depending on the student.

By identifying relationships between these variables, it may be possible to identify students who are at risk of negative responses. Identifying these students is the first step towards the implementation of strategies to lessen the negative effects of testing. Further, this data also makes it possible to identify characteristics of students who flourish in assessment situations.
6.1 Expectancy and Responses to Assessment

This section explores the relationship between student expectancy and student responses to assessment.

6.1.1 Student Expectancy and Responses to NAPLAN

The association between student expectancy and responses to NAPLAN was assessed using Goodman and Kruskal’s gamma. Analysis of positive responses indicated that student expectancy in NAPLAN English had a positive relationship with feeling excited ($G=0.799$, $p<.001$), enthusiastic ($G=0.642$, $p<.001$) and confident ($G=0.463$, $p<.001$) prior to the test. For negative responses, expectancy in NAPLAN had a negative relationship with boredom ($G=−0.523$, $p<.001$). There was no relationship between expectancy and students reporting negative physical responses of sleepless, feeling sick, freezing in the test or crying.

In NAPLAN mathematics, there were similar strong relationships between expectancy and positive responses – feeling excited ($G=0.617$, $p<.001$) confident, ($G=0.430$, $p<.001$) and enthusiastic ($G=0.546$, $p<.001$). Consistent with NAPLAN English, there was a negative relationship between expectancy in NAPLAN mathematics and boredom ($G=−0.529$, $p<.001$). No relationships were found between expectancy and students reporting negative physical responses.

Thus, students who expected to do well in NAPLAN (whether English or mathematics) were more likely to feel excited, enthusiastic and confident, and less likely to feel bored.

6.1.2 Student Expectancy and Responses to Internal Tests

Goodman and Kruskal’s Gamma test showed that for internal tests, higher levels of expectancy in internal English assessments were positively associated with the positive responses of confidence ($G=0.252$, $p=.020$), and enthusiasm ($G=0.469$, $p<.001$) but not excitement. The only negative emotion correlated with expectancy was boredom ($G=−0.425$, $p<.001$).
Similar to NAPLAN, there was no relationship between student expectancy to internal English assessments and negative physical responses to internal tests.

Results were similar for mathematics, as the gamma correlation showed positive, but weaker relationships with the student responses of feeling excited ($G=0.354$, $p=.004$), confident, ($G=0.363$, $p=.001$) and enthusiastic ($G=0.408$, $p=.002$). Expectancy was again negatively correlated with boredom ($G=-0.389$, $p=.002$). There were no significant relationships between expectancy and the negative emotional responses of feeling nervous and stressed or the negative physical responses of freezing during the test, being sleepless the night before, feeling sick or crying.

This would suggest that students who expect to perform well on internal tests are more likely to experience positive response to the tests. Comparatively, in both NAPLAN and internal tests, the relationship between expectancy and responses to assessment was weaker than the relationship between value and responses.

The intra-individual concordance in expectancy across assessment types was analysed. Results revealed that there was a strong positive relationship between expectancies for both types of assessment. There exists a strong relationship ($G=0.707$, $p<.001$) between students’ expectancies in internal English tests and their expectancies in NAPLAN English. A similar strong relationship exists between expectancies to do well in internal mathematics tests and NAPLAN mathematics ($G=0.782$, $p<.002$). As would be anticipated, students who expect to be successful in internal testing also expect to perform well during NAPLAN.

6.2 Student Value and Responses to Assessment

As demonstrated in Section 4.1.1, students differ in both the value they place on, and their emotional responses to assessment. To assess whether or not the value students place on
assessment is associated with their emotional response Goodman and Kruskal’s gamma tests were used.

### 6.2.1 The Relationship between Student Value and Responses to NAPLAN

Analysis of the value students place on doing well in NAPLAN and their emotional responses indicated that student value was positively correlated with positive emotional responses, including feeling excited ($G=0.687, p<.001$), confident ($G=0.564, p<.001$) and enthusiastic ($G=0.523, p<.001$).

Second, with regards to negative emotional responses, there was a strong positive correlation between feeling nervous ($G=0.442, p<.001$), and student value placed on NAPLAN, yet there was no statistically significant relationship between value placed on NAPLAN and feeling stressed. The only significant negative correlation was between student value placed on NAPLAN and boredom ($G=–0.774, p<.001$), indicating that the less value students place on doing well in the NAPLAN tests, the more likely they were to report feeling bored with the assessment.

Finally, looking at negative physical responses (sleeplessness, feeling sick, freezing during the test or crying), no significant correlations were revealed. In sum, these findings suggest that students who placed a higher value on NAPLAN were more likely to experience feeling excited, confident and enthusiastic, but also nervous in response to the tests. They were less likely to report feeling bored than students placing a low value on doing well in NAPLAN.

### 6.2.2 The Relationship between Student Value and Responses to Internal Tests

Analysis of the relationship between student value placed on internal tests and student responses to internal tests revealed that, as for NAPLAN, student value was positively associated with positive responses including feeling excited ($G=0.799, p<.001$), confident ($G=0.463, p<.001$) and enthusiastic ($G=0.642, p<.001$). With regards to negative emotional
responses, there was a negative relationship between student value and boredom \((G=-0.523, p<.001)\) but no relationship between value and nervousness or stress. The decision was made not to examine the relationships between student value and negative physical responses due to the very small number of negative physical responses reported.

Students who placed a high value on internal tests were more likely to experience positive responses and less likely to report being bored during assessment, which aligns with student responses to NAPLAN.

6.2.3 Relationship between Parent, Teacher and Student Value of NAPLAN

As section 6.1.1 indicated, student value is linked to an increased likelihood of experiencing positive responses. As such it is important to investigate what affects students’ valuation of NAPLAN. The adults in any environment can influence children’s perceptions of their world, so we expect student perceptions of both teacher and parent value of assessment tests to be closely related to student value.

For NAPLAN, Goodman and Kruskal’s gamma test found a very strong positive relationship between student perceptions of parent value placed on NAPLAN and student value placed on NAPLAN \((G=0.695, p<.001)\). Student perceptions of teacher value and student value also exhibited a moderate positive relationship \((G=0.431, p<.001)\). This suggests stronger concordance between parent and student value – and could indicate that parents may have a greater influence than teachers on the value their child places on NAPLAN. Further research is required to explore this relationship.

In Section 6.1.4 we investigate if these findings regarding value and NAPLAN are consistent with internal testing.

6.2.4 Relationship between Parent, Teacher and Student Value of Internal Tests

For internal tests, the Goodman and Kruskal’s Gamma test for concordance showed a strong positive correlation between student perceptions of parent value of doing well in
internal tests and students’ value of doing well in internal tests ($G=0.681$, $p<.001$). This was slightly stronger than the correlation between student perceptions of teacher value and student value of internal tests ($G=0.575$, $p<.001$). This is consistent with findings for the NAPLAN assessments.

The data has shown that student perceptions of parent value have the strongest positive correlation with student value for both types of assessment. In addition, higher student value was associated with positive responses, less boredom, and specific to NAPLAN—nervousness.

6.2.5 Academic self-concept and expectancy in NAPLAN

Academic self-concept encompasses an individual’s knowledge and perceptions about themselves in achievement situations (Bong & Skaalvik, 2003). These perceptions rely on social comparison and tend to be fairly stable over time (Bong & Skaalvik, 2003). To assess academic self-concept students were asked to compare themselves to peers when responding to statements, such as ‘Which of the following best describes how you are going in Mathematics?’ Students responded on a four-point scale based on the previous Academic Self-Description Questionnaire (Marsh, 1999): (1) ‘I find this subject difficult’, (2) ‘I’m about average in this subject’, (3) ‘I’m not one of the very top students but I’m above average’ or (4) ‘I’m doing really well in this subject’. To assess the relationship between students’ academic self-concept, expectancy and motivation, a Goodman and Kruskal’s Gamma test was used. As would be predicted, there was a positive correlation between academic self-concept in English and expectancy in NAPLAN English ($G=0.441$, $p<.001$). A positive association was also found between academic self-concept in mathematics and expectancy in NAPLAN mathematics ($G=.529$, $p<.001$). Together these findings indicate that students who believe they are doing well in a subject are more likely to have higher levels of expectancy in NAPLAN.
6.3 Comparing responses between students situated low and high on value, motivation and academic self-concept

6.3.1 Comparing low and high value groups and responses to NAPLAN.

There is evidence of NAPLAN having a negative impact on particular groups of students, such as those with low levels of literacy and numeracy skills (Cumming, Wyatt-Smith & Colbert, 2015). This section aims to investigate whether responses to assessment vary between students with high and low perceptions of teacher/parent or student value. It also compares students with high and low levels of motivation and academic self-concept to determine if specific characteristics of students are associated with responses. For example, do students who find mathematics difficult experience more stress in NAPLAN than students who are doing well in the subject?

To evaluate each of the three stakeholder groups, students were grouped according to those who reported high value (i.e., selected 4 or 5 on the Likert scale) and those who reported low value (i.e., selected 1 to 3 on the Likert scale). Table 39 shows the results of this grouping for perceptions of teacher value.
Table 39

*Responses to NAPLAN by Perceptions of Teacher Value: Perception of High and Low Teacher Value Groups*

<table>
<thead>
<tr>
<th>Perception</th>
<th>Low perceived</th>
<th>High perceived</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher value for NAPLAN</td>
<td>$n=75$</td>
<td>$n=131$</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>38%</td>
<td>54%</td>
<td>.014*</td>
</tr>
<tr>
<td>Confident</td>
<td>17%</td>
<td>35%</td>
<td>.005**</td>
</tr>
<tr>
<td>Stressed</td>
<td>21%</td>
<td>30%</td>
<td>.124</td>
</tr>
<tr>
<td>Bored</td>
<td>76%</td>
<td>39%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>5%</td>
<td>22%</td>
<td>.001**</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>12%</td>
<td>21%</td>
<td>.083</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>13%</td>
<td>12%</td>
<td>.488</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>11%</td>
<td>8%</td>
<td>.053</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>11%</td>
<td>9%</td>
<td>.450</td>
</tr>
<tr>
<td>Crying</td>
<td>3%</td>
<td>3%</td>
<td>.619</td>
</tr>
</tbody>
</table>

Fisher’s exact test was run to determine the association between student perceptions of teacher value for doing well in NAPLAN and responses to testing. The results demonstrated that there were significant differences between students who perceived high levels of teacher value in NAPLAN and reporting ‘feeling nervous’ (54 per cent) and students who perceived low levels of teacher value and reported ‘feeling nervous’ (38 per cent, $p=.014$). Students who were in the high perception of teacher value group also reported...
significantly higher rates of two other positive responses, feeling confident (35 per cent to 17%, $p = 0.005$) and excited (22 per cent to 5 per cent, $p = 0.001$). The most significant results were that the students in the high perceptions of teacher value for NAPLAN reported significantly less boredom during the tests than students in the low perception of teacher value groups (39 per cent to 76 per cent, $p < 0.001$).

Student perceptions of teacher value placed on doing well in NAPLAN showed no significant relationships with any negative physical responses. There was evidence that increased student perceptions of teacher value placed on NAPLAN were associated with increased likelihood of feeling nervous, confident and excited and decreased instances of boredom. This suggests that impressions teachers pass on to their students about NAPLAN may well influence student responses to the tests.

A similar analysis with student perceptions of parent value was conducted. Table 40 shows the perceptions of parent value and student responses to NAPLAN. In each of these, students were grouped according to those who reported high value (i.e., selected 4 or 5 on the Likert scale) and those who reported lower value (i.e., selected 1 to 3 on the Likert scale). Table 40 shows the results of this grouping for perceptions of parent value.
Table 40

Responses to NAPLAN by Student Perceptions of Parent Value: Perception of High and Low Parent Value Groups

<table>
<thead>
<tr>
<th>Perception</th>
<th>Low parent value of NAPLAN n=81</th>
<th>High parent value of NAPLAN n=125</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeling nervous</td>
<td>36%</td>
<td>56%</td>
<td>.003**</td>
</tr>
<tr>
<td>Confident</td>
<td>19%</td>
<td>35%</td>
<td>.007**</td>
</tr>
<tr>
<td>Stressed</td>
<td>21%</td>
<td>30%</td>
<td>.091</td>
</tr>
<tr>
<td>Bored</td>
<td>80%</td>
<td>34%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>10%</td>
<td>21%</td>
<td>.028*</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>6%</td>
<td>25%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>12%</td>
<td>13%</td>
<td>.552</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>12%</td>
<td>10%</td>
<td>.344</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>10%</td>
<td>10%</td>
<td>.564</td>
</tr>
<tr>
<td>Crying</td>
<td>4%</td>
<td>2%</td>
<td>.167</td>
</tr>
</tbody>
</table>

Fisher’s exact test was run to determine the relationships between student perceptions of the value parents placed on the student doing well in NAPLAN and student responses to NAPLAN. Students who perceived their parents placed a high value on NAPLAN reported statistically significant higher percentages of feeling nervous (56 per cent to 36 per cent, \( p =.003 \)), excited (21 per cent to 10 per cent, \( p = 0.28 \)), confident (35 per cent to 19 per cent, \( p = .007 \)) and enthusiastic (25 per cent to 10 per cent, \( p < .001 \)) and reported significantly lower rates of boredom (34 per cent to 80 per cent, \( p < .001 \)). There were no statistically significant differences between student perceptions of parent value of NAPLAN and negative
physical responses of sleeplessness, feeling sick, freezing, crying or feeling stressed.

According to this study, students with high perceptions of parent value are less likely to report boredom, and more likely to report feeling nervous or report positive responses, such as feeling excited, confident and enthusiastic. As was the case with teachers, it appears that the impressions parents give to their children about the importance of NAPLAN are likely to influence student responses, but do not appear to be linked to some of the negative physical responses reported in research and the media.

I also explored the differences in responses to testing according to the value students themselves placed on doing well in NAPLAN (see Table 41). In each of these groupings, students were grouped according to students who reported high value (i.e., selected 4 or 5 on the Likert scale) and those who reported lower value (i.e., selected 1 to 3 on the Likert scale). Table 41 shows the results of this grouping for perceptions of student value.
Table 41

Responses to NAPLAN by Student Value: High and Low Student Value Groups

<table>
<thead>
<tr>
<th></th>
<th>Low student value NAPLAN</th>
<th>High student value NAPLAN</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=101</td>
<td>n=105</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>34%</td>
<td>62%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Confident</td>
<td>15%</td>
<td>42%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Stressed</td>
<td>23%</td>
<td>31%</td>
<td>.135</td>
</tr>
<tr>
<td>Bored</td>
<td>83%</td>
<td>24%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>4%</td>
<td>29%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>7%</td>
<td>28%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>17%</td>
<td>9%</td>
<td>.047*</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>14%</td>
<td>8%</td>
<td>.089</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>13%</td>
<td>7%</td>
<td>.085</td>
</tr>
<tr>
<td>Crying</td>
<td>4%</td>
<td>2%</td>
<td>.219</td>
</tr>
</tbody>
</table>

Student responses to assessment showed the most significant differences between high value and low value groups. Fisher’s exact test was conducted to determine the differences between the responses of students who placed a high value on NAPLAN and students who placed a low value on NAPLAN. Table 41 shows that students who placed a high value on doing well in NAPLAN were more likely than students in the low value group to experience feeling nervous (62 per cent), but also experienced higher levels of confidence (42 per cent to 15 per cent, \( p < .001 \)), excitement (29 per cent to 4 per cent, \( p < .001 \)) and enthusiasm (28 per cent to 7 per cent, \( p < .001 \)). Students who placed high value on doing well in NAPLAN were also much less likely to report boredom (24 per cent to 83 per cent, \( p \)
<0.001). The only statistically significant difference in negative physical responses was that students who placed a low value on doing well in NAPLAN were more likely to report sleeplessness the night before (17 per cent to 9 per cent, \( p = 0.047 \)). There were no significant differences between other negative physical responses to testing, possibly due to the fact that there were few reported cases of each response.

This suggests that students who believe that doing well in NAPLAN is important are more likely to experience positive responses to test, although they were also more likely to report being nervous than the students who placed a low value on doing well. Interestingly, value did not show significant differences in the reported negative physical responses.

6.3.2 Differences between groups: Internal Tests.

This section aims to determine whether student responses to internal tests differ between students with high and low perceptions of teacher value, parent value and the value the individual student placed on doing well in internal tests. Table 42 depicts students’ response to internal tests, according to high and low reported value on internal tests. In each of these groupings, students were grouped according to students who reported high value (i.e., 4 or 5 on the Likert scale) and those who reported lower value (i.e., selected 1 to 3 on the Likert scale). Table 42 shows the results of this grouping for perceptions of teacher value.
Table 42

Responses to Internal Tests by Student Value: High and Low Student Value Groups

<table>
<thead>
<tr>
<th></th>
<th>Low student value internal</th>
<th>High student value internal</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=25</td>
<td>n=181</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>39%</td>
<td>59%</td>
<td>.072</td>
</tr>
<tr>
<td>Confident</td>
<td>9%</td>
<td>46%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Stressed</td>
<td>17%</td>
<td>36%</td>
<td>.044*</td>
</tr>
<tr>
<td>Bored</td>
<td>70%</td>
<td>19%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>4%</td>
<td>28%</td>
<td>.006*</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>9%</td>
<td>23%</td>
<td>.116</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>4%</td>
<td>11%</td>
<td>.481</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>4%</td>
<td>9%</td>
<td>.700</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>13%</td>
<td>6%</td>
<td>.401</td>
</tr>
<tr>
<td>Crying</td>
<td>0%</td>
<td>2%</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Fisher’s exact test showed that students who placed a high value doing well in internal tests were more likely to report feeling confident or excited and less likely to report feeling bored, when compared to the students in the low value group. Students in the high value group were also more likely to report feeling stressed. This suggests first, that the value students place on doing well in assessment tests may be associated with mostly positive responses to assessment tests and second, that the variation in students’ responses is greater for NAPLAN than for internal tests. There were fewer significant differences in responses to internal tests between the students in the low value and high value groups compared to responses to NAPLAN. It may be that responses to NAPLAN are similar to students’
previous responses to other forms of assessment. For example, students who were bored for internal tests may be more likely to be bored for NAPLAN and those who experienced positive responses to internal tests will be more likely to report positive responses to NAPLAN. These findings are evidence that responses to national testing programs will vary according to the value students place on assessment tests in general.

6.4 Motivation

6.4.1 Motivation and responses to NAPLAN

This study measured motivation using the proxy of self-reported effort in NAPLAN and internal tests. This section aims to explore differences in responses to assessment between groups of students with high motivation and groups of students with low motivation for NAPLAN. It was expected that students with higher levels of motivation would experience more positive responses to the assessment than students with lower motivation.

Students were asked, ‘How hard did you try in NAPLAN this year?’ Possible responses ranged from ‘1: I didn’t try at all’ to 5: ‘I put in my best effort’. Those who reported a three or lower on the Likert scale were placed in the low motivation group and those who reported a four or higher were placed in the high motivation group. Table 43 shows students’ responses to NAPLAN, according to high and low motivation groupings.
Table 43

*Student Responses to NAPLAN by Motivation: High and Low Student Motivation Groups*

<table>
<thead>
<tr>
<th></th>
<th>Low Motivation</th>
<th>High Motivation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAPLAN</td>
<td>n=54</td>
<td>n=152</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>26%</td>
<td>56%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Confident</td>
<td>11%</td>
<td>35%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Stressed</td>
<td>16%</td>
<td>30%</td>
<td>.036*</td>
</tr>
<tr>
<td>Bored</td>
<td>78%</td>
<td>43%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>0%</td>
<td>34%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>9%</td>
<td>20%</td>
<td>.045*</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>13%</td>
<td>13%</td>
<td>.548</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>9%</td>
<td>11%</td>
<td>.459</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>11%</td>
<td>9%</td>
<td>.432</td>
</tr>
<tr>
<td>Crying</td>
<td>2%</td>
<td>3%</td>
<td>.504</td>
</tr>
</tbody>
</table>

Fisher’s exact test showed significant differences in responses between the high motivation and low motivation groups in NAPLAN. The significant findings were that students with low motivation were less likely to report the negative emotional responses of feeling nervous (26 per cent to 56 per cent, \( p < .001 \)) or stressed (16 per cent to 30 per cent, \( p = .036 \)). It is noteworthy that the high motivation students were more likely to experience stress. For positive responses, the low motivation were less likely to report confidence (11 per cent to 35 per cent, \( p < .001 \)) and enthusiasm (9 per cent to 20 per cent, \( p = .045 \)). A large difference was in students reporting excitement, with no students in the low motivation group
being excited for the test. The low motivation group are significantly more likely to report boredom (78 per cent to 43 per cent, \( p < 0.001 \)). Again, there were no statistically significant differences in negative physical responses to the testing between the high and low motivation groups. These results suggest that increased motivation for the test increases the likelihood of experiencing positive responses to the test, but also increases the likelihood of experiencing feeling nervous or stressed.

6.4.2 Motivation and responses to internal testing

This section investigated whether the association between motivation and responses to internal tests are similar to the relationships between motivation and responses to NAPLAN. As with NAPLAN, students were grouped by the average of their responses to two questions, ‘How hard did you try in Mathematics tests this year?’ and ‘How hard did you try in English tests this year?’ Those who reported an average of three or lower on the Likert scale were placed in the low motivation group and those who reported an average four or higher were placed in the high motivation group. Table 44 shows students’ responses to internal tests, according to high and low motivation groupings.
Table 44

*Student Responses to Internal Tests and Motivation: High and Low student motivation groups*

<table>
<thead>
<tr>
<th></th>
<th>Low motivation internal</th>
<th>High motivation internal</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=21</td>
<td>n=185</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>25%</td>
<td>59%</td>
<td>.031*</td>
</tr>
<tr>
<td>Confident</td>
<td>8%</td>
<td>45%</td>
<td>.015*</td>
</tr>
<tr>
<td>Stressed</td>
<td>25%</td>
<td>35%</td>
<td>.550</td>
</tr>
<tr>
<td>Bored</td>
<td>75%</td>
<td>22%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>8%</td>
<td>26%</td>
<td>.302</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>0%</td>
<td>23%</td>
<td>.074</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>8%</td>
<td>10%</td>
<td>1.00</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>0%</td>
<td>9%</td>
<td>.604</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>8%</td>
<td>7%</td>
<td>1.00</td>
</tr>
<tr>
<td>Crying</td>
<td>0%</td>
<td>2%</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The differences in motivation and responses to internal test are similar to the differences in responses to NAPLAN. Fisher’s exact test showed that students who have a high level of motivation for internal tests are more likely to experience positive emotional responses, such as confidence (45 per cent to 8 per cent, p = .015). Interestingly, students in the high motivation group are also more likely to experience feeling nervous about internal tests (59 per cent as opposed to 25 per cent for the low motivation group). The low value group are more likely to report feeling bored in internal tests (75 per cent to 22 per cent, p < .001).
In this chapter, I demonstrated that differences in expectancy, value and motivation are associated with student responses to both types of assessment. The results suggest that high levels of expectancy, value and motivation are associated mostly with more positive emotional responses to assessment. Students who report low levels of each of these variables are more likely to report being bored during internal tests. One notable result was that negative physical responses to testing were not associated with student expectancy, value or motivation for tests.

6.5 Academic Self-Concept

6.5.1 High and Low Academic self-concept groups and responses to NAPLAN

This section explores whether students with high or low academic self-concept differ in their responses to assessment. Students who reported, ‘I find this subject difficult’ or ‘I’m about average in this subject’ were categorised into the low academic self-concept group. The high academic self-concept group consisted of those students who reported ‘I’m not one of the best students but I’m above average’ or ‘I am doing really well in this subject’. Table 45 presents students’ responses to NAPLAN, according to whether they were high or low in academic self-concept in English.
Table 45

Responses to NAPLAN by High and Low Academic Self-Concept in English

<table>
<thead>
<tr>
<th></th>
<th>Low academic self-concept</th>
<th>High academic self-concept</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=73</td>
<td>n=103</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>43%</td>
<td>51%</td>
<td>.385</td>
</tr>
<tr>
<td>Confident</td>
<td>20%</td>
<td>33%</td>
<td>.076</td>
</tr>
<tr>
<td>Stressed</td>
<td>34%</td>
<td>23%</td>
<td>.732</td>
</tr>
<tr>
<td>Bored</td>
<td>55%</td>
<td>51%</td>
<td>.486</td>
</tr>
<tr>
<td>Excited</td>
<td>8%</td>
<td>21%</td>
<td>.019*</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>9%</td>
<td>22%</td>
<td>.034*</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>18%</td>
<td>10%</td>
<td>.124</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>11%</td>
<td>11%</td>
<td>1.000</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>9%</td>
<td>10%</td>
<td>1.000</td>
</tr>
<tr>
<td>Crying</td>
<td>4%</td>
<td>2%</td>
<td>.668</td>
</tr>
</tbody>
</table>

Fisher’s exact test showed that students who reported a higher level of academic self-concept in English were more likely to feel excited (21 per cent to 8 per cent, \( p = 0.019 \)) and enthusiastic (22 per cent to 9 per cent, \( p = 0.034 \)) prior to NAPLAN tests. No significant differences between low and high academic self-concept were found across the negative responses of feeling stressed, nervous, sleeplessness the night before the test or crying. Table 46 present student responses to NAPLAN, according to high and low academic self-concept in Mathematics.
Table 46

*Student Responses to NAPLAN by High and Low Academic Self-Concept in Mathematics*

<table>
<thead>
<tr>
<th></th>
<th>Low academic self-concept</th>
<th>High academic self-concept</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=73</td>
<td>n=133</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>45%</td>
<td>50%</td>
<td>.563</td>
</tr>
<tr>
<td>Confident</td>
<td>20%</td>
<td>32%</td>
<td>.077</td>
</tr>
<tr>
<td>Stressed</td>
<td>32%</td>
<td>23%</td>
<td>.143</td>
</tr>
<tr>
<td>Bored</td>
<td>59%</td>
<td>47%</td>
<td>.082</td>
</tr>
<tr>
<td>Excited</td>
<td>8%</td>
<td>21%</td>
<td>.019*</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>12%</td>
<td>20%</td>
<td>.181</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>15%</td>
<td>11%</td>
<td>.512</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>14%</td>
<td>9%</td>
<td>.335</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>15%</td>
<td>7%</td>
<td>.083</td>
</tr>
<tr>
<td>Crying</td>
<td>5%</td>
<td>2%</td>
<td>.188</td>
</tr>
</tbody>
</table>

Fisher’s exact showed that students who reported a higher academic self-concept in Mathematics were likely to report feeling excited for the NAPLAN. Again, there were no significant differences in negative or negative physical responses between the two groups.

These results indicate that students who have higher levels of academic self-concept in English were more likely to report positive responses of feeling enthusiastic and excited in response to NAPLAN.
6.5.2 **High and Low Academic self-concept and responses to internal tests.**

This section explores the difference in responses to internal testing, according to low and high levels of academic self-concept. Students who reported, ‘I find this subject difficult’ or ‘I’m about average in this subject’ were categorised into the low academic self-concept group. The high academic self-concept group consisted of those students who reported ‘I’m not one of the best students but I’m above average’ or ‘I am doing really well in this subject’. Table 47 below shows the percentage of students according to low and high academic self-concept for each response.

**Table 47**

*Responses to Internal Tests by High and Low Academic Self-Concept in English*

<table>
<thead>
<tr>
<th></th>
<th>Low academic self-concept</th>
<th>High academic self-concept</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=73</td>
<td>n=103</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>52%</td>
<td>60%</td>
<td>.303</td>
</tr>
<tr>
<td>Confident</td>
<td>32%</td>
<td>49%</td>
<td>.019*</td>
</tr>
<tr>
<td>Stressed</td>
<td>37%</td>
<td>33%</td>
<td>.646</td>
</tr>
<tr>
<td>Bored</td>
<td>37%</td>
<td>18%</td>
<td>.004**</td>
</tr>
<tr>
<td>Excited</td>
<td>8%</td>
<td>29%</td>
<td>.182</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>18%</td>
<td>23%</td>
<td>.381</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>7%</td>
<td>12%</td>
<td>.336</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>7%</td>
<td>9%</td>
<td>.792</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>5%</td>
<td>8%</td>
<td>.581</td>
</tr>
<tr>
<td>Crying</td>
<td>1%</td>
<td>2%</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Fisher’s exact test indicated that students who reported a higher academic self-concept in English were more likely to report feeling confident in internal tests (49 per cent to 32 per cent, \( p = 0.019 \)). This same group of students were less likely to report feeling bored during internal tests (18 per cent to 37 per cent, \( p = 0.004 \)). Table 48 below shows the reported responses to internal tests by academic self-concept in Mathematics.

Table 48

*Student Responses to Internal Tests by High and Low Academic Self-Concept in Mathematics*

<table>
<thead>
<tr>
<th></th>
<th>Low academic self-concept</th>
<th>High academic self-concept</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n=73 )</td>
<td>( n=133 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>56%</td>
<td>58%</td>
<td>.883</td>
</tr>
<tr>
<td>Confident</td>
<td>27%</td>
<td>51%</td>
<td>.001***</td>
</tr>
<tr>
<td>Stressed</td>
<td>41%</td>
<td>31%</td>
<td>.168</td>
</tr>
<tr>
<td>Bored</td>
<td>30%</td>
<td>22%</td>
<td>.238</td>
</tr>
<tr>
<td>Excited</td>
<td>19%</td>
<td>29%</td>
<td>.180</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>12%</td>
<td>26%</td>
<td>.021*</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>14%</td>
<td>8%</td>
<td>.235</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>7%</td>
<td>9%</td>
<td>.792</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>14%</td>
<td>4%</td>
<td>.012*</td>
</tr>
<tr>
<td>Crying</td>
<td>3%</td>
<td>1%</td>
<td>.287</td>
</tr>
</tbody>
</table>

Fisher’s exact test showed that students who reported higher levels of academic self-concept in Mathematics were more likely to report feeling confident (51 per cent compared to 27 per cent, \( p = .001 \)) and enthusiastic (26 per cent compared to 12 per cent, \( p = .013 \)) before
internal tests. Students in the low academic self-concept group were much more likely to report feeling sick before internal tests (14 per cent to 4 per cent \( p = 0.012 \)). This shows that the differences in responses to internal tests according to high and low academic self-concept in Mathematics were show that the high academic self-concept group were more likely to experience positive responses including feeling confident and enthusiastic, whereas the low academic self-concept in mathematics group were more like to report feeling bored or sick before the test.

Based on these findings, a high or positive academic self-concept is linked to positive responses towards assessment in general. The higher students perceive their own ability, the more likely they are to feel confident, enthusiastic and excited about testing. Conversely, there also a propensity for students who reported low levels of academic self-concept to report feeling sick before both some tests. These findings are intuitive, as assessment is most likely to benefit those who succeed. Students who are successful in assessments can look forward to their results as a confirmation of their ability but for those with a record of poor performances it may simply be a confirmation of their beliefs about their academic capacity.

6.6 **Gender**

This section explores whether students’ responses to assessment differ by gender. Responses to NAPLAN by male and female are shown in Table 49.
Table 49

*Responses to NAPLAN by Gender*

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th><strong>p value</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>n=135</em></td>
<td><em>n=71</em></td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>42%</td>
<td>59%</td>
<td>.028*</td>
</tr>
<tr>
<td>Confident</td>
<td>18%</td>
<td>34%</td>
<td>.023*</td>
</tr>
<tr>
<td>Stressed</td>
<td>19%</td>
<td>41%</td>
<td>.002**</td>
</tr>
<tr>
<td>Bored</td>
<td>51%</td>
<td>53%</td>
<td>.063</td>
</tr>
<tr>
<td>Excited</td>
<td>17%</td>
<td>15%</td>
<td>.855</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>10%</td>
<td>21%</td>
<td>.052</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>9%</td>
<td>20%</td>
<td>.045*</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>7%</td>
<td>18%</td>
<td>.016*</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>7%</td>
<td>15%</td>
<td>.050</td>
</tr>
<tr>
<td>Crying</td>
<td>1%</td>
<td>6%</td>
<td>.119</td>
</tr>
</tbody>
</table>

Fisher’s exact test showed that there were differences in responses to NAPLAN by gender. Female participants were more likely to report being stressed (41% to 19%, *p* = .002) and feeling nervous (59% to 42%, *p* = .028) but also confident (34% to 18%, *p* = .023) than males. Females were also more likely than males to report freezing during the test (18% to 7%, *p* = .016).

Student responses to internal tests by gender are shown in Table 50.
Fisher’s exact test indicated there were three significant differences between males and females responses to internal tests. Males were more likely to report feeling confident (49% to 28%, \( p = .005 \)) and enthusiastic (27% to 7%, \( p < .001 \)) than females prior to internal tests. Females were more likely than males to report being stressed before internal tests (46% to 28%, \( p = 0.013 \)).

### 6.7 Year Level Differences

Student responses to testing have been investigated in both primary and secondary schools (Howell, 2015; Rogers et al., 2016; Wyn et al., 2014). Recent evidence has shown complex relationships between age and responses to testing, in particular with national
programs such as NAPLAN. Rice et al. (2015) found the percentage of teachers who reported that students experienced stress from NAPLAN increased from Year 3 to 7 then decreased from Year 7 to 9. Rogers et al. (2016) found that reported stress in students undertaking NAPLAN tests increased from Year 3 to Year 5. It is likely that as students move through school their understandings and perceptions around tests change, and with it, their responses. Therefore, this section will investigate if responses to assessment differ between Year 7 and Year 9 students.

6.7.1 Year level and differences in value of NAPLAN

This section explores the difference in value that Year 7 and 9 students place on NAPLAN, starting with the value that students perceive their teachers place on NAPLAN. Table 51 depicts students’ perceptions of teacher value of NAPLAN by year level.
Table 51

*Student Perceptions of Teacher Value of NAPLAN by Year Level*

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year 7 (NAPLAN)</strong></td>
<td>3%</td>
<td>8%</td>
<td>15%</td>
<td>34%</td>
<td>39%</td>
<td>3.98</td>
<td>1.07</td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Year 9 (NAPLAN)</strong></td>
<td>12%</td>
<td>9%</td>
<td>32%</td>
<td>32%</td>
<td>14%</td>
<td>3.28</td>
<td>1.18</td>
</tr>
<tr>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t-test was run to determine the association between year level and student perceptions of teacher value of NAPLAN. Year 9 students had a significantly lower perception of teacher value for doing well in NAPLAN (M = 3.28, SD = 1.18), than Year 7 students (M = 3.98, SD = 1.07, t(204) = 4.30, p < .001, d = 0.62). This means that Year 7 students are more likely to perceive that teachers place a higher value on doing well in NAPLAN than the students in Year 9. This relationship was also expected to extend to student perceptions of parent value. Table 52 shows the student perceptions of parent value of NAPLAN by year level.
Table 52

**Student Perceptions of Parent Value of NAPLAN by Year Level**

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>5%</td>
<td>8%</td>
<td>16%</td>
<td>30%</td>
<td>42%</td>
<td>3.97</td>
<td>1.14</td>
</tr>
<tr>
<td>(NAPLAN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 9</td>
<td>17%</td>
<td>23%</td>
<td>19%</td>
<td>23%</td>
<td>19%</td>
<td>3.03</td>
<td>1.38</td>
</tr>
<tr>
<td>(NAPLAN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t-test found that Year 9 students perceived that their parents placed less value on doing well in NAPLAN (M = 3.03, SD = 1.38) than Year 7 students (M = 3.97, SD = 1.14, t(204) = 5.27, p < .001, d = 0.74. Given the lower student perceptions of parent and teacher value among Year 9 students, it is likely that Year 9 students themselves place a lower value on NAPLAN than Year 7 students. Table 53 shows the value placed on NAPLAN by students, according to year level.
Table 53

Student Value for NAPLAN by Year Level

<table>
<thead>
<tr>
<th></th>
<th>1: Not all important</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7 (NAPLAN)</td>
<td>10%</td>
<td>8%</td>
<td>19%</td>
<td>31%</td>
<td>31%</td>
<td>3.66</td>
<td>1.27</td>
</tr>
<tr>
<td>Year 9 (NAPLAN)</td>
<td>28%</td>
<td>12%</td>
<td>29%</td>
<td>21%</td>
<td>9%</td>
<td>2.72</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Consistent with the findings from parents and teachers, an independent samples $t$-test found Year 9 students placed a significantly lower value on doing well in NAPLAN (M=2.72, SD = 1.32) compared to students in Year 7 (M = 2.72, SD = 1.32, $t(204) = 4.99, p < .001, d = 0.72$). Overall, this suggests that by Year 9, student perceive that teachers, parents and the students themselves all place a lower value on NAPLAN than year 7 stakeholders.

6.7.2 Year level and differences in value of internal tests

Section 6.6.1 identified statistically significant differences in the student perceptions of the value of NAPLAN between Year 7 and 9. This section investigates if these differences can also be found for internal testing. Table 54 shows the difference in value students placed on doing well in internal tests between Year 7 and Year 9.
Table 5

**Student Value for Internal Tests by Year Level**

<table>
<thead>
<tr>
<th>Year Level (Internal)</th>
<th>1: Not at all important</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>29%</td>
<td>60%</td>
<td>4.43</td>
<td>0.86</td>
</tr>
<tr>
<td>Year 9</td>
<td>1%</td>
<td>1%</td>
<td>10%</td>
<td>21%</td>
<td>66%</td>
<td>4.51</td>
<td>0.82</td>
</tr>
</tbody>
</table>

An independent samples t-test found no significant differences between year level and value for internal tests. Furthermore, an independent samples t-test for independent samples found no statistically significant difference between the value students place on internal testing from Year 7 (M = 4.43, SD = 0.86) to Year 9 (M = 4.51, SD = 0.82, t(204) = -0.64, p = .260, d = 0.10). This indicates that the value of internal testing remains relatively constant from year 7 to 9; in contrast, the reported values for doing well in NAPLAN are lower for Year 9 students than Year 7 students.

6.7.3 **Year level and differences in Expectancy for NAPLAN and Internal Tests**

The next section investigated differences in expectations of performance between Year 7 and Year 9. Table 55 shows the responses to ‘I expect to do well in NAPLAN English this year’ according to year level.
Table 5

Student Expectancies for NAPLAN English by Year Level

<table>
<thead>
<tr>
<th></th>
<th>1: Strongly Disagree</th>
<th>2: Disagree</th>
<th>3: Neither Agree or Disagree</th>
<th>4: Agree</th>
<th>5: Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>4%</td>
<td>7%</td>
<td>21%</td>
<td>40%</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td>3.81</td>
<td>1.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 9</td>
<td>5%</td>
<td>7%</td>
<td>34%</td>
<td>43%</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>3.47</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t test indicated that Year 7 student expectancies in NAPLAN English (M=3.81, SD=1.03) were significantly higher than those of Year 9 students (M=3.45, SD=0.963, t(204) = 2.337, p = .010, d = 0.38).

Table 56 shows student expectancies for NAPLAN mathematics, according to year level.

Table 56

Student Expectancies for NAPLAN Mathematics by Year Level

<table>
<thead>
<tr>
<th></th>
<th>1: Strongly Disagree</th>
<th>2: Disagree</th>
<th>3: Neither Agree or Disagree</th>
<th>4: Agree</th>
<th>5: Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>3%</td>
<td>5%</td>
<td>24%</td>
<td>32%</td>
<td>37%</td>
</tr>
<tr>
<td></td>
<td>3.95</td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 9</td>
<td>9%</td>
<td>8%</td>
<td>32%</td>
<td>32%</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>3.40</td>
<td>1.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
An independent samples t test indicated that Year 7 students reported significantly higher levels of expectancy in NAPLAN Mathematics (M=3.95, SD=1.02) than Year 9 students (M=3.40, SD=1.15, t(204) = 3.32, p < .001, d = 0.51).

Overall, expectancies regarding performance in NAPLAN are lower for Year 9 students than Year 7 students.

6.7.4 Year level and expectancy for internal tests

This section compares students’ expectancies for internal testing by year level. Students were asked the extent to which they agreed with the statement, ‘I expect to do well in [English/Mathematics] at school this year’ and Table 57 shows students’ expectancy for internal English assessment, while Table 58 shows students’ expectancy for internal mathematics assessment.

Table 57

<table>
<thead>
<tr>
<th>Student Expectancies for Internal English Tests by Year Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Strongly Disagree</td>
</tr>
<tr>
<td>Strongly Disagree</td>
</tr>
<tr>
<td>Year 7</td>
</tr>
<tr>
<td>n=131</td>
</tr>
<tr>
<td>Year 9</td>
</tr>
<tr>
<td>n=75</td>
</tr>
</tbody>
</table>
An independent samples t-test found no significant differences between the Year 7’s expectations to do well in internal English tests (M = 3.95, SD = 0.96) and the expectations of Year 9 students (M = 3.89, SD = 0.90, t(204) = .389, p = .349, d = .064).

Table 58

<table>
<thead>
<tr>
<th>Student Expectations for Internal Mathematics Test by Year Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Strongly Disagree</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Year 7</td>
</tr>
<tr>
<td>Year 9</td>
</tr>
</tbody>
</table>

An independent sample t-test showed there were no significant differences in the expectations for internal mathematics tests between Year 7 (M = 4.06, SD = 0.98) and Year 9 students (M = 3.96, SD = 1.12, t(204) = .672, p = .503, d = .095).

The data discussed in Section 6.6.4 would suggest that the expectation to perform well in NAPLAN is lower for Year 9 students than Year 7 students; in contrast, expectations for internal tests are similar for both year levels.

6.7.5 Year level and motivation for NAPLAN

Student expectancy and value for NAPLAN was shown to decrease from Year 7 to Year 9. Drawing on expectancy-value theory, we would expect a lower level of motivation for students in Year 9, than students in Year 7. This is important as lower levels of motivation may
lead to poor student results and a possibly inaccurate measure of students’ literacy and numeracy skills in Year 9. Students were asked; ‘How hard did you try in the NAPLAN mathematics test this year?’ Table 59 shows student motivation for NAPLAN mathematics by year level.

Table 59

Student Motivation for NAPLAN Mathematics by Year Level

<table>
<thead>
<tr>
<th>Year 7</th>
<th>1: I didn’t try at all</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4%</td>
<td>3%</td>
<td>8%</td>
<td>29%</td>
<td>55%</td>
<td>4.30</td>
<td>1.01</td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 9</td>
<td>8%</td>
<td>11%</td>
<td>19%</td>
<td>37%</td>
<td>25%</td>
<td>3.61</td>
<td>1.20</td>
</tr>
<tr>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t test indicated there was a significant difference between Year 7 motivation for NAPLAN mathematics (M=4.30, SD=1.01) and Year 9 motivation for NAPLAN mathematics (M=3.61, SD=1.20, t(204) = 4.347, p <.001, d = .662). Year 7 students reported higher average levels of effort than students in Year 9.

Table 60 presents the data for students’ motivation for NAPLAN English, according to year level.
Table 60

*Student Motivation for NAPLAN English by Year Level*

<table>
<thead>
<tr>
<th></th>
<th>1: I didn’t try at all</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>3%</td>
<td>1.5%</td>
<td>13%</td>
<td>29%</td>
<td>53%</td>
<td>4.24</td>
<td>1.05</td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 9</td>
<td>5.3%</td>
<td>9%</td>
<td>20%</td>
<td>35%</td>
<td>30%</td>
<td>3.72</td>
<td>1.15</td>
</tr>
<tr>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples *t* test indicated there was a significant difference between Year 7 motivation for NAPLAN English (*M*=4.28, *SD*=1.05) and Year 9 motivation for NAPLAN English (*M*=3.76, *SD*=1.16, *t*(204) = 3.32, *p* < .001, *d* = .472), with lower levels of motivation among Year 9 students. These differences support the idea that student motivation for NAPLAN is lower for Year 9 students than for Year 7 students.

6.7.6 Year level and motivation for internal tests

Previous sections have identified that there is a significant difference in motivation for NAPLAN between Year 7 and Year 9 students, suggesting that the value and expectancies vary between those year levels. This section aims to determine if the difference in motivation applies to assessment tests generally or is limited to NAPLAN. Table 61 shows the motivation for internal tests by year level.
Table 61

_Student Motivation for Internal English Tests by Year Level_

<table>
<thead>
<tr>
<th></th>
<th>1: I didn’t try at all</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year 7</strong></td>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>31%</td>
<td>57%</td>
<td>4.40</td>
<td>0.85</td>
</tr>
<tr>
<td><strong>Year 9</strong></td>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3%</td>
<td>1%</td>
<td>5%</td>
<td>37%</td>
<td>53%</td>
<td>4.37</td>
<td>0.98</td>
</tr>
</tbody>
</table>

An independent samples t test indicated there was no significant difference between Year 7 motivation for internal English assessment (M = 4.40, SD = 0.85) and Year 9 motivation for internal English assessment (M = 4.37, SD = 0.98, t(204) = -.251, p = .802, d = .032).

Table 62 compares students’ motivation for internal mathematics, according to year level.
Table 62

Student Motivation for Internal Mathematics by Year Level

<table>
<thead>
<tr>
<th></th>
<th>1: I didn’t try at all</th>
<th>2: I put in a small effort</th>
<th>3: I put in an average effort</th>
<th>4: I put in a large effort</th>
<th>5: I put in my best effort</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>1%</td>
<td>0%</td>
<td>8%</td>
<td>24%</td>
<td>67%</td>
<td>4.57</td>
<td>0.70</td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 9</td>
<td>4%</td>
<td>1%</td>
<td>8%</td>
<td>21%</td>
<td>65%</td>
<td>4.43</td>
<td>0.99</td>
</tr>
<tr>
<td>n=75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t test indicated there was no significant difference between Year 7 (M = 4.57, SD = 0.70) and Year 9 students’ motivation for internal mathematics assessments (M = 4.43, SD = 0.99, t(204) = 1.232, p = .110, d = .163)

By comparing the data in Tables 57–60, we observe another key difference between internal testing and NAPLAN. There were no significant differences between student motivations for internal tests by year level. For NAPLAN, as student value and expectancy decrease from Year 7 to Year 9, motivation also decreases. This is what would be predicted by the expectancy-value theory. NAPLAN appears to be seen as less important as students move through secondary schools, with the result that student motivation for doing well in the tests decreases.

6.7.7 Differences in Value for NAPLAN and Internal Testing within Year Levels

The section examined the differences in value place on doing well in NAPLAN and in internal testing within year levels. Students were grouped by year level and the survey items used to measure this data were ‘How important is it for you to do well in NAPLAN this
year?’ and ‘How important is it to you to do well in internal tests this year?’ Table 63 depicts this data for Year 7 students.

Table 63

Comparing Student Value of NAPLAN and Internal Tests in Year 7 Students

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 7</td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>29%</td>
<td>60%</td>
<td>4.43</td>
<td>0.86</td>
</tr>
<tr>
<td>Internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 7</td>
<td>10%</td>
<td>8%</td>
<td>19%</td>
<td>31%</td>
<td>31%</td>
<td>3.66</td>
<td>1.27</td>
</tr>
<tr>
<td>NAPLAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=131</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A paired samples t test indicated that there was a significant difference between the value placed on doing well in NAPLAN ($M=3.66$, $SD=1.27$) and value placed on doing well in internal tests by students in Year 7 ($M=4.43$, $SD=0.86$, $t(130) = -7.20$, $p < .001$, $d = 0.71$). This supports this study’s hypothesis that students place a higher value on internal tests than NAPLAN. The same data was examined for Year 9 students. Table 64 shows data concerning the value of internal tests and NAPLAN for Year 9 students.
Table 64

*Comparing Student Value of NAPLAN and Internal Tests in Year 9 Students*

<table>
<thead>
<tr>
<th></th>
<th>1: Not at all</th>
<th>2: Not important</th>
<th>3: A bit important</th>
<th>4: Important</th>
<th>5: Very important</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 9 Internal</td>
<td>1%</td>
<td>1%</td>
<td>9%</td>
<td>21%</td>
<td>67%</td>
<td>4.51</td>
<td>0.82</td>
</tr>
<tr>
<td>Year 9 NAPLAN</td>
<td>28%</td>
<td>12%</td>
<td>29%</td>
<td>21%</td>
<td>9%</td>
<td>2.72</td>
<td>1.32</td>
</tr>
</tbody>
</table>

A paired samples $t$ test indicated that there was a significant difference between the value Year 9 students placed on doing well in internal tests ($M=4.55$, $SD=0.83$) compared to the value they place on doing well in NAPLAN ($M=2.72$, $SD=1.32$), $t(74)=9.87$, $p<0.001$, $d=1.63$). Both Year 7 and Year 9 students placed a significantly higher value on doing well in internal tests than in NAPLAN.

### 6.7.8 Year level and responses

After finding that expectancy, value and motivation for assessment differ between year levels, the study also investigated how student responses to assessment vary from Year 7 to Year 9. It is hypothesised that year 9 students would experience fewer negative responses to NAPLAN as the value placed on doing well is low for the students. Students were asked;
‘Did you experience any of the following before NAPLAN?’ The responses are presented in Table 65.

Table 65

**Student Responses to NAPLAN by Year Level**

<table>
<thead>
<tr>
<th></th>
<th>Year 7 NAPLAN</th>
<th>Year 9 NAPLAN</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeling nervous</td>
<td>54% n=131</td>
<td>37% n=75</td>
<td>.025*</td>
</tr>
<tr>
<td>Confident</td>
<td>35%</td>
<td>17%</td>
<td>.001**</td>
</tr>
<tr>
<td>Stressed</td>
<td>25%</td>
<td>29%</td>
<td>.044*</td>
</tr>
<tr>
<td>Bored</td>
<td>42%</td>
<td>71%</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>Excited</td>
<td>22%</td>
<td>7%</td>
<td>.005**</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>20%</td>
<td>13%</td>
<td>.243</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>8%</td>
<td>20%</td>
<td>.424</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>12%</td>
<td>9%</td>
<td>.701</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>9%</td>
<td>7%</td>
<td>.332</td>
</tr>
<tr>
<td>Crying</td>
<td>0%</td>
<td>8%</td>
<td>.250</td>
</tr>
</tbody>
</table>

An exact McNemar’s test determined that there were statistically significant differences reported in emotional responses between Year 7s and Year 9s. For positive responses, Year 7s were more likely to report feeling confident (35 per cent to 17 per cent, \( p = .001 \)) and excited (22 per cent to 7 per cent, \( p = .005 \)) than students in Year 9. There were two small differences for the negative responses, with Year 9s more likely to report feeling stressed than students in Year 7 (29 per cent compared to 24 per cent \( p = .044 \)), but Year 7 students more likely to feel nervous than those in Year 9 (54 per cent compared to 37 per cent, \( p = .025 \)). The largest difference was the proportion of students reporting boredom:
Year 9 students were much more likely to report feeling bored (71 per cent) than Year 7 students (42 per cent) \((p < .001)\). There were no statistically significant differences for negative physical responses between the year levels.

After exploring the differences between Year 7 and Year 9 students’ responses to NAPLAN, I investigated if these findings were specific to NAPLAN or if they existed in internal tests. Students were asked; ‘Did you experience any of the following before internal assessment?’ Table 66 showcases the results to this question, according to year level.

### Table 66

**Student Responses to Internal Tests by Year Level**

<table>
<thead>
<tr>
<th></th>
<th>Year 7 internal</th>
<th>Year 9 internal</th>
<th>(p) value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=131)</td>
<td>(n=75)</td>
<td></td>
</tr>
<tr>
<td>Feeling nervous</td>
<td>52%</td>
<td>65%</td>
<td>.052</td>
</tr>
<tr>
<td>Confident</td>
<td>41%</td>
<td>47%</td>
<td>.235</td>
</tr>
<tr>
<td>Stressed</td>
<td>30%</td>
<td>43%</td>
<td>.043*</td>
</tr>
<tr>
<td>Bored</td>
<td>25%</td>
<td>25%</td>
<td>.494</td>
</tr>
<tr>
<td>Excited</td>
<td>28%</td>
<td>20%</td>
<td>.126</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>21%</td>
<td>21%</td>
<td>.571</td>
</tr>
<tr>
<td>Sleepless the night before</td>
<td>8%</td>
<td>13%</td>
<td>.187</td>
</tr>
<tr>
<td>Freezing during the test</td>
<td>9%</td>
<td>7%</td>
<td>.366</td>
</tr>
<tr>
<td>Feeling sick before the test</td>
<td>7%</td>
<td>8%</td>
<td>.483</td>
</tr>
<tr>
<td>Crying</td>
<td>0%</td>
<td>3%</td>
<td>.057</td>
</tr>
</tbody>
</table>

An exact McNemar’s test determined that the only statistically significant difference was that Year 9 students more likely to report feeling stressed for internal tests (43 per cent)
than Year 7s (30 per cent, \( p = .043 \)). There were no statistically significant differences in the other responses to internal tests between Year 7 and Year 9.

It may be that as most students reach Year 9, they develop an apathy towards NAPLAN, as demonstrated by increased reporting of boredom and fewer positive responses. Conversely, there are minimal differences in Year 7 and Year 9 students’ responses to internal tests. Importantly, there appears to be a small group of students who are experiencing negative physical responses to both types of assessments.

6.8 Personality

To investigate responses to assessment further, students were asked about their personality to determine if there were general traits that were associated with more positive or negative responses to NAPLAN.

TIPI was used to investigate if personality traits were associated with student responses to NAPLAN. The TIPI measures five major personality traits: extraversion, agreeableness, conscientious, emotional stability and openness to experience. Goodman and Kruskal’s Gamma statistic was used to identify concordance with each emotional response to NAPLAN. This section explores the relationships between the five major personality traits and responses to NAPLAN.

First, Extraversion was associated with the positive responses of feeling excited \((G=0.310, p=.034)\) and confident \((G=0.233, p=.050)\) and was inversely related to feeling stressed \((G=−0.249, p=.040)\). There were no positive correlations with Agreeableness, but strong negative correlations with reported excitement \((G=−0.633, p=.003)\) and freezing during the test \((G=−0.694, p=.006)\). Conscientiousness correlated strongly with reported confidence \((G=0.390, p=.001)\) and negatively correlated with reported boredom \((G=−0.338, p=.002)\). Openness had a weak negative correlation with students reporting feeling bored \((G=−0.216, p=.012)\).
Finally, the personality type that had the strongest correlations with responses to NAPLAN was Emotional Stability. Higher levels of Emotional Stability were positively correlated with positive responses to NAPLAN, including enthusiasm ($G=0.568, p=.002$) and confidence ($G=0.362, p=.008$). Higher levels of Emotional Stability were also negatively correlated with sleepless the night before ($G=-0.491, p=.019$), stressed ($G=-0.445, p=.001$) and excitement ($G=-0.569, p=.001$). Overall, it appears students who have low levels of Emotional Stability are more likely to experience negative responses to NAPLAN.

6.9 Summary

This chapter aimed to determine the variables that are associated with student value for NAPLAN and internal tests. For NAPLAN, the strongest relationship was between student perceptions of parent value and student value. There was also a moderate relationship between student perceptions of the value their teachers place on doing well and student value. This may indicate that the strongest influence on student attitude to NAPLAN is from their parents. To a lesser extent, students learn of the importance of NAPLAN from their teachers through direct or indirect messages. For internal tests, there is a similar, strong relationship between student perceptions of parent value and student value. Student perceptions of teacher value also have a positive, yet slightly weaker relationship with student value for internal tests. This sample showed that student perceptions of parent value has the strongest relationship with student value in both types of assessment.

This chapter also discussed differences in responses to NAPLAN and internal tests by different groups of students. The data showed that students who reported higher value, expectancy and motivation were more likely to experience positive responses to testing including excitement, enthusiasm and confidence. Students who reported lower value, expectancy and motivation for assessments were more likely to experience boredom. There
were no significant differences between the high and low value groups and negative physical responses to testing.

Responses to assessment also differed by year level. Year 9 students were more likely to report lower value and motivation for NAPLAN than Year 7 students. Year 9s were also more likely to report being bored during NAPLAN. This finding suggests that once students reach Year 9, the low stakes associated with NAPLAN may lower their motivation to perform.

Another variable associated with responses to testing was an individual’s level of emotional stability. Students with higher levels of emotional stability were more likely to experience positive responses to assessment.

There was also a group of students who reported ‘finding the subject difficult’. These students were more likely to report negative responses to assessment. Therefore, it may be that the negative responses to testing are more closely related to the value students place on the test, their personality and academic self-concept, rather than the type of assessment.

Overall, students with higher levels of expectancy or academic self-concept are more likely to experience positive responses to testing and students with lower reported levels of these constructs are more likely to be bored.

Chapter 7 considers the impact of these findings within the context of current educational policy and explores how this increased understanding can lead to an improved educational experience for students.
Chapter 7: Discussion of Findings, Implications for Policy, and Further Research

7.1 Responses to Assessment in the Australian Context

This chapter outlines the purpose and research questions, describes the methodology and summarises the major findings. The findings are discussed against the available literature and the limitations of the study and recommendations for future research and policy implications are offered.

This study investigated the issue of national testing in Australian schools. Three and a half million school children and over 350,000 teaching staff are affected by NAPLAN. These literacy and numeracy tests are justified by two perceived benefits; (1) to help drive improvements in student outcomes and; (2) to provide increased accountability to the community (ACARA, 2018).

Using tests to produce change is not a new concept in education. There is a history of aptitude, intelligence and achievement tests being used to change education and influence stakeholders. However, the current shift to national testing has been referred to as a ‘global phenomenon’ (Lingard et al., 2016, p. 1). It can be argued that these tests are driven by neoliberal ideologies and the increasing influence of government in education (Lingard et al., 2016). Neoliberal apologists for the testing could argue that the increased transparency of results will result in free-market forces driving competition and therefore improvement in education. There is some controversy around the drivers behind NAPLAN. The political drivers may include accountability, transparency and return on investment. It could also be viewed as a feedback tool used by schools and teachers to improve the quality of their service.
to students. Other perspectives suggest that NAPLAN operates as a marker of school quality, in which teacher and school effectiveness is conceptualised in terms of value for money, supposedly protecting the nation’s human capital investment (Howell, 2015).

Education can be considered one of the core mechanisms for the development of human capital within a nation; as such, its importance is closely related to economic policies within nations (Lingard et al., 2016). This view has resulted in national testing programs that attempt to improve educational practices through measurement and evaluation (Lingard et al., 2016). These testing programs affect stakeholders—students, parents and teachers—and the Australian case shows that national testing is a complex and multifaceted entity (Lingard et al., 2016). The Australian school system is changing as a result of a data-driven feedback loop. Student performance data are used to guide policies and school-level practice and the impact of these changes is measured by student performance data. The focus of data has been described in both positive and negative ways. For example, teachers are more aware of the importance of data driven decision making with an aim to develop high quality, reliable assessments which will be part of the feedback loop. Conversely, there have also been concerns that the data are not aligned with the Australian Curriculum or integrated with the latest theories of learning (Cumming et al., 2018). In addition schools, the government, media and public also use information from NAPLAN to form conclusions about school performance and the success of the education system (Mockler, 2015).

The provision of NAPLAN data for the general population through the My School website is done with the aim of improving the transparency and accountability of schools (Lingard et al., 2016). While the media’s publication of league tables has increased public interest, it has often focused on the problems in education (Mockler, 2015). One study found that at the school level, 95 per cent of educators perceived that weaker than expected NAPLAN results would have a negative or very negative impact on the school’s reputation.
(Dulfer et al., 2012), indicating that the test is high stakes for schools. Testing advocates would argue that parents are encouraged to exercise choice in relation to their child’s school enrolment based on NAPLAN results, thereby contributing to school improvement (by removing their children from poor-performing schools).

This study aimed to understand how NAPLAN affects one key stakeholder in education: the students. It intended to give students a voice in describing how they experience NAPLAN testing. This research also compared student perceptions around NAPLAN to those related to internal tests and explored differences between groups of students. As such, this study aimed to answer four research questions:

1. How do students experience and respond to NAPLAN?
2. What differences are there in students’ responses to internal tests and to NAPLAN?
3. To what extent do responses to these two types of assessment differ by year level, academic self-concept and the value placed on the assessment?
4. What are the relationships between student perceptions of teacher and parent value placed on doing well in the tests, and student responses to assessment?

7.1. Research question 1: How do students experience and respond to NAPLAN?

Responses to NAPLAN varied between students. The three most common responses were students feeling stressed, bored or confident. Some students selected more than one of these options.

This study found that around a quarter of students are concerned about the NAPLAN, and reported feeling stressed about the tasks. These findings are broadly congruent with perceptions of students’ responses from other stakeholders such as parents. Wyn et al. (2014) found that approximately 40 per cent of parents reported signs of stress or anxiety in their children prior to NAPLAN testing. This is somewhat more than the 27 per cent of students who directly reported that they felt stressed by NAPLAN tests in the current study. This
could be because only secondary students were surveyed in this study, whereas Wyn et al. (2014) surveyed parents of children from all year levels. Other research has identified anxiety and stress in students in response to national testing (e.g. Stiggins, 1999; Lewis, 2000) and research into teacher perceptions of NAPLAN by Dulfer et al. (2012) found that almost 90 per cent of teachers reported that they had observed anxiety or stress among some students. We could conclude that although testing does make some students stressed, at the secondary level at least, the stress may be felt by a minority of a class.

My findings were similar to Rogers et al. (2016) who investigated self-reported stress levels in response to NAPLAN in Year 3 and 5 students across 16 independent schools. They concluded that because of the small differences in stress responses, the size of the negative impact of NAPLAN was uncertain. The Rogers et al. (2016) study used a simple measure of emotional distress by asking students to rate their experience on a four-point scale (i.e., ‘not at all’, ‘a little bit’, ‘quite a bit’ and ‘a lot’), rather than the simple ‘Yes’ or ‘No’ scale used in the present study. Rogers et al. (2016) found that distress was higher during NAPLAN testing time. Despite stress being higher during NAPLAN, no mean value for any year level was higher than ‘a little bit’.

Previous research into NAPLAN has identified students experiencing negative physical responses to NAPLAN (Dulfer et al., 2012). The following section explores students who reported negative physical responses: sleeplessness the night before NAPLAN, freezing during the test, feeling sick or crying, and compares findings with previous research.

Overall, there was a small percentage of students who reported the negative physical responses of sleepless the night before, crying, freezing and feeling sick. The percentage of participants who reported being sleepless the night before NAPLAN (13 per cent) is similar to that found in studies surveying parents about sleeplessness before NAPLAN (Wyn et al., 2014). These types of strong and potentially concerning responses often feature in media
discussions regarding NAPLAN (Mockler, 2012). Dulfer et al. (2012) found that 40 per cent of teachers reported some students suffering from sleeplessness the night before. Three per cent of students participating in this study reported crying in response to NAPLAN. This response is consistent with a number of other studies. For example, Madaus et al. (2009) also reported students feeling ‘frustrated or crying’ in response to high-stakes tests, and identified this response in a small minority of students. Approximately six per cent of Australian parents in Wyn’s (2014) study reported their children crying due to NAPLAN. In contrast, Dulfer et al. (2012) indicated that approximately half of Australian educators in their sample reported that parents had complained about their child crying in response to NAPLAN. The difference between Dulfer’s findings and that of this study may be due to one of several factors. First, this study only surveyed secondary students, while the Dulfer et al. (2012) study surveyed primary and secondary teachers. Secondary students may be more familiar with testing and less likely to experience strong physical responses that their primary counterparts. Second, response rates for Dulfer et al.’s (2012) research were much lower than for the current study (less than 5% of teachers for Dulfer’s study compared to a response rate of 71% for students in this study). It is possible that those teachers with the most negative experiences of NAPLAN and strong feelings about it self-selected into Dulfer et al.’s (2012) study. Finally, the Dulfer et al. (2012) study asked participants whether they had “ever had any students” experience given reactions to NAPLAN (p. 33). Teachers participating in this study with some years of experience may have mentally grouped together all students they have ever taught (i.e. multiple cohorts) who had undertaken NAPLAN, leading to a higher percentage of participants affirming that they had observed these responses among their students. Thirteen percent of students in this study self-reported freezing in response to NAPLAN. This compares with twenty per cent of teachers reporting that at least one parent
had reported their child had a fear of ‘freezing up’ during the test (Dulfer et al., 2012). The reported rates from teachers and students, while not the same, suggest that a fairly small percentage of students have freezing as a response to NAPLAN.

Eleven percent of students in this study reported feeling sick in response to NAPLAN. This was reported by a similar percentage of teachers (12 per cent) in Dulfer et al.’s (2012) study. Overall, strong negative responses—including physical responses of crying, freezing, feeling sick or sleeplessness—were identified in a small percentage (between 3% and 13%) of our sample.

This can be compared to parent perceptions of students, with Dulfer et al. (2012) reporting that 81% of teachers reported that at least one parent had complained of their child feeling sick before the test. Again, this much higher figure may be due to the differences between this study and that of Dulfer, outlined previously.

Unlike many other studies, this study collected data about specific responses directly from students. The relatively small number of students who reported experiencing negative responses allows for a more accurate understanding of the impact of NAPLAN on students. The findings from this study suggest that negative responses to NAPLAN occur in a small group of students and that this group of students experiences similar responses to internal tests (see Section 7.2). Another key finding was that 50 per cent of students reported that they completed the test without any strong feeling either way or were bored when completing the test.

This research also identified a group of students rarely mentioned in the research and debates around testing programs. This group (25 per cent) reported feeling excited, enthusiastic or confident about undertaking NAPLAN. This contrasts with previous research on teacher perspectives on NAPLAN, with 50% of teachers reporting that no students in their class were looking forward to the test (Wyn et al., 2014). This suggests that the impact of
testing could be oversimplified if we assume all students will respond the same way, or if we fail to ask students themselves how they feel. Some students react positively and others respond negatively to the same test. This makes it difficult to describe the test as having either a positive or negative impact on students.

In keeping with the findings of this study, some research suggests that negative physical responses to national testing may exist only in a small sample of the student population (e.g. Wheelock et al., 2002). It may be that as a national testing program, NAPLAN produces fewer negative responses than some other international testing programs (see Section 7.2). This is likely to be because the stakes are low for the individual student.

Findings from this study suggest that feedback from NAPLAN is reaching students, but is not used to improve learning. For example, most students reported having seen their NAPLAN results, but generally they have not discussed them with their parents or teachers. The study also found that most students reported that they do not know how their results are used.

Most students reported some practice for NAPLAN. Three quarters of students reported they completed a small number of practice tests at school and 20 percent of students completed NAPLAN workbooks at home. Overall, there was little evidence of excessive NAPLAN practice.

7.1.2 Research question 2: What differences are there in students’ responses to internal tests and to NAPLAN?

Much concern has been expressed about the negative consequences of national testing programs on students (Brown et al., 2004; Dulfer et al., 2012; Klenowski, 2015; Polesel et al., 2012; Stiggins, 1999; Rice et al., 2015). However, it is not known if some of the negative physical and emotional responses reported in relation to such programs are due to the programs themselves, or are simply a more general student response to testing. To understand
which is likely to be the case, responses to testing programs (in this case, NAPLAN) must be compared to those to internal tests. The general process of testing causes a variety of responses from students and these responses are likely to vary with students’ expectancy, value and motivations. Presently, it would appear that there is little in the literature that compares student responses to internal tests to their responses to state and national testing.

There were differences in the three categories of student responses to NAPLAN and internal tests: positive, negative emotional and negative physical. Students were more likely to report positive responses to internal tests than to NAPLAN, as evidenced by the differences in the percentages of students reporting that they felt ‘confident’ or ‘excited’. This could be due to the closer relationship between the curriculum and internal tests. At secondary level in Australia, content is often presented in discrete units of work with a summative assessment after 3-6 weeks; this is likely to give students confidence that understand the required content and can predict what is likely to be on the test.

For negative responses, more students reported feeling nervous and stressed for internal tests than for NAPLAN; this suggests that it may be testing in general that causes some of the most discussed negative responses that can feature in media reports, rather than NAPLAN itself. As NAPLAN is likely to be lower stakes for students than their internal tests, they are less likely to be nervous about the test. A striking difference in responses was that students were twice as likely to report being bored in NAPLAN, compared to internal tests.

The higher levels of boredom for NAPLAN could demonstrate a lack of knowledge about the purpose of the test; students do not understand why they need to complete the test, do not discuss results, and cannot see if and how their results are used to improve their learning. The high levels of boredom in response to NAPLAN may be due to the perception that the NAPLAN results are meaningless for the individual student. As most students do not
know how the NAPLAN results are used, it may be possible that students do not see the purpose of NAPLAN, which lowers value and motivation but also generates fewer strong negative physical responses.

Other researchers have offered explanations for the decreased levels of negative physical responses and lower stress among older students. They suggest that for some students, increased familiarity with the test over time may reduce stress (Rice et al., 2015). Alternatively, lack of stress could also be due to testing fatigue or a perceived lack of test relevance (Rice et al., 2015).

There were no significant differences in negative physical responses to the two types of tests, with a small group of students reporting negative physical responses to both types of test. This suggests that the variable that determines negative physical responses is more likely to be the type of student, rather than the type of test.

Another difference is the level of expectancy for each test, with students reporting higher levels of expectancy in their internal tests than in NAPLAN. There could be many reasons for this.

I hypothesise that students’ expectancy is higher for internal tests because the assessed content relates to their recent class work and the higher frequency of the internal testing than NAPLAN will most likely lead to familiarity and understanding of the testing process. Students are also likely to receive explicit communication from their teachers about the content of internal tests whereas NAPLAN content is less predictable.

Regarding the feedback from the tests, students were significantly more likely to report discussing the results of their internal tests with their parents than their NAPLAN results. There may be many reasons for this, including the greater number of internal tests completed and the more detailed feedback likely to be given to parents from these tasks. This may also in part be due to the smaller turnaround time for internal testing or the sequential
flow of the curriculum in some subjects, where students need to master one content area at a time. In addition, the long delay between the sitting of NAPLAN tests (in May) and the provision of feedback reports to parents (in August) may also decrease the feedback students receive. These findings may reinforce students’ perceptions about the importance of internal assessment in school, as they reported that parents and teachers are far more likely to be involved in the learning, assessment and feedback cycle for internal tests than NAPLAN.

Findings from this study suggest that NAPLAN, at the secondary school level, is just another form of assessment and that largely, secondary school students complete the task with similar emotional responses to internal tests. At the secondary level at least, NAPLAN does not cause more stress than other tests, which may be due to the lower stakes associated with NAPLAN for the individual student. This is not consistent with other studies involving teachers, which have suggested NAPLAN causes additional stress in large numbers of students (Dulfer et al., 2012), but this may be due to differences in samples, with younger students experiencing greater stress in response to NAPLAN before they have fully developed the understanding that their NAPLAN performance does not carry any major consequences for them.

More broadly, the findings suggest that the debate about the responses to NAPLAN should not seek to remove all types of stress from students’ lives, but rather ensure there is a balance between the benefits of testing and the negative responses some students experience. External measures of performance—such as interviews, university exams and driver’s licence tests—are a part of life beyond secondary school and students may well benefit from controlled exposure to these types of situations in order to develop coping strategies.
7.1.3 Research question 3: To what extent do responses to these two types of assessment differ by year level, academic self-concept and the value placed on the assessment?

Students’ responses to NAPLAN and internal testing were found to vary greatly and depended on many factors, including academic self-concept, personality traits, year level, and the value placed on the assessments. The findings from this study suggest that students with higher levels of expectancy in terms of achievement, greater emotional stability or more positive academic self-concept are more likely to experience positive responses to assessment, whether internal or external.

Students with higher reported academic self-concept were more likely to report feeling confident and excited for NAPLAN and for internal testing than students with lower levels of academic self-concept. This could be due to the positive reinforcement loop that exists for these students when completing assessments. They have been successful in assessments in the past, then they perform well in a recent test, and this confirms their positive self-concept, giving them a high expectancy for the next test. This suggests that the testing process may accentuate differences between students. Students who are confident and enthusiastic may tend to become more confident and enthusiastic through this process, and those who are less confident may find this exacerbated. It may be that those students who are academically strong enjoy the testing process and those who struggle will disengage and are more likely to become bored during the test. The findings from this study suggest that NAPLAN might work to disadvantage lower achieving students; however, internal testing is likely to have the same effect.

Another key variable is the personality trait, emotional stability, as it had the strongest correlations with responses to NAPLAN and internal testing. Emotionally stable students were less likely than those with lower levels of emotional stability to report negative responses to testing. There were also positive correlations between emotional stability and
students reporting feeling excited and confident in response to NAPLAN and internal tests. It may be that students with higher levels of emotional stability are better prepared for the stress of assessment tasks, regardless of their academic ability. This suggests that when exploring the impact of national testing programs, we need to consider the personality traits of the students.

With regard to differences between year levels, both Year 7 and Year 9 students placed a higher value on internal tests than NAPLAN. Despite general similarities there exist differences between year levels in response to NAPLAN. Year 7 students were more likely to report feeling nervous, but also more likely to report feeling confident or excited compared to Year 9 students.

The Year 9 students reported significantly higher rates of boredom in NAPLAN than the Year 7s. This may be due to fewer students understanding the purpose of NAPLAN and the low stakes associated with their NAPLAN results. If students do not know the purpose, they will place a low value on the test and this could affect motivation. If a large portion, or a specific group (e.g. Year 9 boys), within the student population experience this, the results may not accurately measure student performance and this would affect the data. For the year 9 cohort, it could be expected that the high reported levels of boredom would be associated with lower motivation and performance. This may result in consistently low results in NAPLAN, particularly at the higher year level and could impact the validity of NAPLAN data. This could result is some policy decisions and interventions at the system, school or individual level not being effective.

This is consistent with comments by a school leader in existing research who stated that ‘students generally are not stressed about NAPLAN because neither they, nor their parents see much, if any value in it—Year 9 in particular’ (Rice et al., 2015, p. 76), indicating
that by Year 9, students are not as concerned by NAPLAN as they may be more focused on their internal tests.

It is possible that the very low stakes associated with NAPLAN at Year 9 have decreased motivation to a point that the test does not accurately measure student ability. If NAPLAN does not accurately measure student ability, then it does not meet the purpose of measuring student achievement as set out by ACARA.

These findings align with others in the international research. A study in the United States by Brown et al. (2004) found that by middle school, students ‘exhibit more scepticism about the value of national testing and discriminate between tests derived from their curricula and testing imposed from outside school’ (p. 146).

The pattern of varying responses to NAPLAN has been found in students from other year levels. Rogers et al. (2016) found that emotional distress in relation to NAPLAN was higher for Year 5 students than in Year 3 students. Rice et al. (2015) found that the percentage of teachers reporting negative student responses (e.g. feeling stressed or feeling sick) increased from Year 3 to Year 7, but decreased between Year 7 and Year 9. The current study found that from Year 7 to Year 9, confidence, nervousness and excitement all decrease, while boredom increases. Interestingly, the number of students who reported feeling stressed was similar in Year 7 and Year 9. It could be that there are a certain number of students who are stressed by the NAPLAN assessment across all year groups. However, it must be noted that there could be a cohort effect in this study as the Year 7 and Year 9s were different students; longitudinal research is required to determine whether students’ attitudes to NAPLAN change as they progress through school.

There are many reasons for the difference in value of NAPLAN between year levels. For example, students in secondary schools may have a greater awareness of the lower stakes for NAPLAN compared to primary school students. The younger students may be less
accustomed to the format of formal testing, in which they are on their own and teachers are no longer helpers, but merely supervisors and this could lead to increased stress and anxiety. Victorian secondary school students are also much more focused on the higher stakes testing at the end of school (e.g., the Victorian Certificate of Education, more commonly known as VCE). Conversely, because there are no other external assessments by which to judge school performance at the primary level, it is possible that NAPLAN may be higher stakes for primary teachers. It may also be that the stakes vary between teachers within a school, depending on their attitude to assessment and the subject and year levels they teach.

For the Year 9 cohort, the high reported levels of boredom are associated with lower motivation for the tests. This may lead to NAPLAN results that do not accurately reflect student ability, particularly at the higher year levels, and essentially make decisions based on the data questionable. If NAPLAN does not accurately measure student ability, then it does not meet the purpose of measuring student achievement as set out by ACARA (ACARA, 2011).

There was a strong positive relationship between student value for NAPLAN and positive responses to the test. The relationship with expectancy was similar, with students who expected to do well in NAPLAN (whether English or mathematics) more likely to feel excited, enthusiastic and confident, and less likely to feel bored. This pattern was consistent for internal tests. Students who placed a high value or high expectancy for doing well in internal tests were also more likely to report positive responses and less likely to report feeling bored.

The findings confirm part of what was predicted using expectancy-value theory. The expectancy and value students have for doing well in assessment tasks was associated with positive responses. Contrary to what was predicted, variations in expectancy and value did not show any significant relationships with reported negative physical responses. Overall,
students who placed a higher value and had a higher expectancy were more likely to experience positive responses to the test. They were however, more likely to feel nervous before the tests whereas low value students were more likely to report being bored.

7.1.4 Research question 4: What are the relationships between student perceptions of teacher and parent value placed on doing well in the tests, and student responses to assessment?

This study found students placed a higher value on doing well in internal tests than NAPLAN, and perceived that their parents and teachers also placed a higher value on doing well in internal tests than NAPLAN. Students perceived that all stakeholders value internal tests more than NAPLAN. Notably, students perceived that both parents and teachers valued NAPLAN more highly than they did, and this is consistent with previous research (Wyn et al., 2014).

The finding that students perceived their parents value NAPLAN more highly than themselves may be due to parents using the data as a chance to benchmark their children’s results with state averages. Wyn’s (2014) research showed that some parents do value NAPLAN, with 70 per cent of parents stating that they find the results useful. Interestingly, it appears parents have a more positive attitude to NAPLAN than teachers (Wyn et al., 2014). In contrast, Rogers et al. (2018), found that responses were mixed, with some parents very negative, and others positive, however overall the perception was that NAPLAN was only ‘somewhat’ useful (p. 509). According to this study, students reported that both parents and teachers place more value on internal tests than NAPLAN, while parents appear to place a higher value of NAPLAN than the teachers.

Previous research suggests that the relationship between parent and student attitudes to NAPLAN is complex. In surveys of parents, there are large differences between those in
favour of NAPLAN and those against it. For example, parents who were against NAPLAN were more likely to report stress and anxiety in their children than those parents who supported the program (Wyn et al., 2014). Specifically, the reporting of negative physical responses was five times higher for those parents against NAPLAN that those in favour (Wyn et al., 2014). The direction of this relationship is debatable. We do not know if parents are against the NAPLAN program because they have witnessed negative impacts on their children, or if parents’ negative attitudes to NAPLAN are somehow conveyed to their children, and this leads to negative student responses.

I also investigated the relationships between the perceptions of value placed on NAPLAN. There was a stronger correlation between student perceptions of parent value placed on tests and student value, than between student perceptions of teacher value and student value. It would appear that student attitudes to assessment may be more strongly influenced by their parents, than by teachers, but this would need to be confirmed through further research.

Students who perceived that their parents placed a high value on NAPLAN were more likely to experience positive responses to this assessment. Importantly this group was also more likely to report feeling nervous before the test, yet they did not report significantly higher levels of stress. This same group also reported much lower levels of boredom for NAPLAN. This could suggest that increased perceptions of parent value of NAPLAN are mostly associated with more positive responses to the test.

The study thus found that there were relationships between students’ perceptions of the value placed on various forms of testing and the values they themselves place on these tests. Expectancy-value theory (Eccles, 1983; Eccles & Wigfield, 2002) would predict that the value placed on doing well in assessments will vary between stakeholders and therefore, so may the motivations of those stakeholders. For example, the stakes for NAPLAN may be
highest for teachers and this study suggests that they place a higher value on NAPLAN than their students. This may be due to Australia’s national testing policies that encourage school leaders, department of education staff and parents to use NAPLAN as proxy for school quality (Howell, 2015; Mockler, 2015; Polesel et al., 2012).

Teachers know that performance in NAPLAN will have an impact on the reputation of the school and staff, which can affect enrolments. NAPLAN results are usually included in school annual reports and these is also some evidence of them being used for leadership performance bonuses (Howell, 2015). Comparatively, more recent research suggests that most teachers (in independent schools) did not report much impact on their well-being due to NAPLAN testing (Rogers et al., 2018). It may be that some schools place a high value on improving NAPLAN results, for example low-performing schools or those in areas with high competition for student enrolments, and these schools may have increased pressure on teachers to perform well in NAPLAN. These teachers may feel stressed about how the tests could be used to judge their effectiveness. These concerns may also contribute to students’ perception that their teachers place the highest value on doing well in NAPLAN.

Some teachers may be perceived to place less value on NAPLAN, due to their perception that NAPLAN lacks useful information about student performance. Previous research found that 54 per cent of teachers do not find the data from NAPLAN useful (Dulfer et al., 2012). Similarly, Paris and McEvoy (2000) suggested that high-stakes testing does not provide effective estimates of individuals’ learning and therefore, the usefulness of the data for teachers is limited. Athanasou (2010) reported similar findings when collecting data from school principals; only half of these school leaders reported finding NAPLAN results useful. Independent school teachers from Western Australia also did not report any large impact on their curriculum, nor did they perceive the testing to be particularly useful (Rogers et al.,
It is possible that teachers who do not find NAPLAN useful may communicate this to their students implicitly or explicitly.

Teachers’ perceptions of the lack of useful information generated by NAPLAN may influence students’ beliefs that there is low teacher value on NAPLAN compared to internal tests. Compared to NAPLAN, internal tests are usually designed by teachers, and are therefore more likely to be based on the subject matter teachers prioritise. There are also higher stakes for students with these internal tests, as grades are determined and reported directly to parents within a short time period. Another explanation for these findings may be the general negative attitude to NAPLAN in schools. A Newspoll study that found 70 per cent of teachers have negative feelings about NAPLAN (as cited in Dulfer et al., 2012). Students may detect these negative feelings and thus, place lower value on NAPLAN. For example, other research suggests that some teachers had individual conversations with students about their NAPLAN results, while others considered the NAPLAN tests irrelevant and did not commit any effort to using the testing data to inform teaching and curriculum planning (Ng et al., 2016). This variation in the uses of NAPLAN data may be symptomatic of a widespread lack of understanding about the purpose and reliability of the tests.

Another emerging theme is differences in attitudes to NAPLAN within stakeholder groups. Teachers who self-reported emotional distress during NAPLAN were more likely to perceive negative impacts on students (Wyn et al., 2014). This may mean that if an individual has a negative perspective of NAPLAN, they are more likely to suggest NAPLAN causes negative responses in other stakeholders—a classic example of social projection bias.

The variation in the value placed on NAPLAN may be due to many misconceptions about NAPLAN, these include both the purpose of the test or the correct use of the results. This study found that that most students reported that they do not know how their results are used. At the teacher level, it could be argued that teachers want useful feedback on their
students’ performance that will help them improve their teaching. The core issue here is that the test data are not particularly robust with regard to individual students. Groups of teachers (e.g., the English team) may be able to examine data across the years to identify areas of relative strength and weakness and then use this information to plan curriculum for the following year. However, small variations in small groups of students are not valid for decision-making (Wu, 2015).

### 7.2 Conclusion

The present research examined student expectancy, value, motivation and responses to NAPLAN and internal assessment in 206 students from Victorian schools. It also explored the student perceptions of parent and teacher value of NAPLAN and internal tests. In a review of the literature, there was a need for further research into NAPLAN that collected data directly from students. Prior studies included data collected from teachers (Dulfer et al., 2012; Thompson, 2013) and three studies with data collected from students (Rogers et al., 2016; Howell, 2012; Wyn et al, 2014). This study adds to the literature by examining in more detail student responses to NAPLAN and internal testing. The present study is the first time students have self-reported their responses to both NAPLAN and internal testing and gives policy makers an indication of the proportions of students responding to tests in certain ways. It also provided the opportunity to compare the impact of NAPLAN and the impact of internal tests. The findings contribute to our understanding of the impact of NAPLAN on students.

This study found that negative responses to NAPLAN are not evident amongst a large proportion of students in lower secondary school, and that these responses are similarly evident for internal tests. This is in contrast with some of the overseas research on national testing programs (Flores & Clarke, 2003; Lewis, 2000; Madaus et al. 2009 & Stiggins, 1999).
However, this may be due to differences in testing programs. One difference between NAPLAN and some other national testing programs is that NAPLAN performance is not a hurdle requirement for progression to the next year level, nor do NAPLAN test results appear to be routinely used to stream students. The stakes associated with NAPLAN thus tend to be high mainly for teachers and schools, whose performance may be judged on the results. The findings from this study suggest that if assessments have high stakes at the teacher level, teachers will most likely place a higher value on the tests.

At the individual level, two key findings were firstly, that NAPLAN has similar, and for some students, less of a negative impact than internal tests. This study found that NAPLAN did not generate higher reported rates of negative responses than internal tests, possibly due to the low stakes attached to NAPLAN for the individual student. Secondly, negative physical responses to NAPLAN occurred in a very small percentage of students. There were lower reported rates of negative physical responses in this study when compared to previous research with data collected from teachers (Dulfer et al., 2014) and parents (Wyn et al., 2014). However, recent research with student participants has also indicated relatively low levels of stress in students participating in NAPLAN (Rogers et al., 2016). It may be that the negative impact of NAPLAN on students is less severe than parents and teachers perceive.

It is likely that by having low stakes for the individual student, NAPLAN is less likely to produce negative responses in students. This means that some negative consequences of national testing programs overseas may not apply to NAPLAN in Australia. For example, concerns about anxiety, stress and dropping out (Stiggins, 1999), a sense of futility (Gregory & Clarke, 2003) and exhaustion, frustration and physical distress (Madaus et al., 2009) may well be less likely in the Australian context, due to the low-stakes nature of the test for students. While there may be other good arguments for the removal of NAPLAN, such as the
low reliability and validity of results for individual students, removal on the basis that it makes students stressed, cry or feel sick does not appear to have a strong foundation in the data from this study.

Overall, I argue that NAPLAN is low stakes for secondary school students, as the test does not have consequences for student grades, year level progression or subject selection. I suggest that the impact of NAPLAN on students is similar, and in most cases, less than the impact of internal tests. Therefore, the impact of a national testing program on students’ experiences of school is dependent on the complex relationships between the stakes associated with the test for each of the stakeholders, student academic self-concept, personality traits and year level.

More broadly, the national assessment program has been criticised for its negative impact on curriculum, pedagogy and student wellbeing. This study suggests the impact on most students’ wellbeing is not major. There is also a small group of individual student who find testing stressful and the fact that any students are badly affected by assessment, and potential impact on their school studies and school continuance is an area that needs to be considered. Policymakers still have to assess what the benefits of NAPLAN are, and question whether this is how best to spend educational dollars. Previous research into the negative impact on national testing programs has shown that testing has the power to increase inequality across students (Dee & Jacob, 2006; Teese, 2000). This study found that lower performing students were more likely to experience boredom during NAPLAN testing and higher achieving students more likely to report positive responses to NAPLAN. These findings are consistent with international studies which have found that national testing programs have a disproportionately negative effect on low performing students (Smyth & Banks, 2012). To continue with NAPLAN, policy makers must be aware of the negative
impact of NAPLAN on specific groups and ensure that testing does not increase current inequalities in Australian education.

7.2.1 Limitations and Implications for Further Research

The findings of this thesis need to be considered within the context of the following limitations. Further research should aim to replicate these findings with a larger and more diverse sample. Including students from both primary and secondary, as well as schools from various sectors, of differing sizes, and diverse geographical locales would allow for a greater understanding of the nuances, and support generalisability to the broader population. In addition, future research should seek to include selective schools or schools with a competitive ethos where there may be additional expectations for students to perform well. More research is also required to determine whether these findings can be replicated across a range of schools with high or low ICSEA values. Such replication is important given that Thompson and Harbaugh (2013) report that teachers perceive the negative impact of NAPLAN more often in lower socioeconomic status schools, while Dandy and Nettelbeck (2002) suggest that greater expectations placed on higher SES students may increase test-associated stress.

A methodological limitation of the current study is the dichotomous response format used to measure several variables. Although the instrument captured a variety of emotional responses, future research may seek to explore these emotions with more detailed measures using, for example, Likert scales. For instance, previous research (Rogers et al., 2016) has used the short form emotional distress scale (DASS-21). Utilising similar tools would provide a more robust measure of stress, boredom or confidence experienced by students. Combined with longitudinal data, this would also allow for a comparison of stress during NAPLAN with other baseline measures throughout the year.
Beyond the types of responses measured it is imperative that any future research continue to assess the similarities and differences between types of testing programs (including internal tests, NAPLAN and other national testing programs) and the impact these tests have on students. Australia is moving towards online NAPLAN testing and research is required to compare the responses between the ‘pen and paper’ and online NAPLAN tests. Furthermore, this research has found that responses to assessment may differ between year levels. Consequently, collecting data from students of different ages could show important differences in this area.

Current debates and writing on high-stakes testing often use the term very broadly. Overall, the current findings highlight a need to be specific about what we mean by high-stakes tests. This research highlights the need for further theoretical work that defines and identifies stakes and stakeholders, and explores the relationships between stakes, stakeholders and the impact of tests. In the case of NAPLAN, results may be high stakes for some teachers and schools, but appear to be low stakes for students. An investigation of testing programs in multiple settings would allow for a comparison of the stakes associated for different stakeholders and see their various impacts on students. This would allow for more meaningful comparisons between the many ‘high stakes’ tests, and add depth to our understanding of the impact of testing programs.

### 7.2.2 Implications for policy

The findings from this research have implications at many levels in education. At the national level, the reliance on NAPLAN may be symptomatic of a lack of other effective means to measure school quality. Test scores are relatively easy to measure; however, policymakers need to draw on a broader range of measures to judge school quality. There are ways of measuring other aspects of schools and these can be combined and weighted to
develop more holistic measures. These measures include indicators of student wellbeing, quality governance, relationships with parents, curriculum and teacher quality. If the purpose of schooling is to provide the best possible outcomes for our students, we need to find ways of measuring and improving all aspects of schools. While the My School website does contain a range of data beyond NAPLAN, and attempts to compare school performance against similar schools, NAPLAN is foregrounded as the achievement measure. In addition, while systems do collect a range of data on schools (for example, student engagement levels are measured in Victoria), anecdotal evidence suggests that NAPLAN data often has pre-eminence in discussions between schools and system-level staff around school performance.

Orchestrators of national testing programs must develop assessments that meet the needs of students, teachers, schools and national level data collection. This could require a move to online and adaptive testing (currently being implemented in stages for NAPLAN), which would give faster feedback to all stakeholders and allow them to see the purpose of the tests. It is also essential that school leaders and teachers have a clear, consistent message regarding the purpose of NAPLAN and its limitations. This includes a shift away from political statements and league tables that use NAPLAN for political gain or sensationalist headlines that sell media articles (e.g. Singhal, 2019).

At the school level, there are implications for teachers. Teachers may need to identify students who have negative responses to internal tests as they are more likely to experience similar responses to NAPLAN. The stress of students could be decreased by teachers explaining the purpose of the tests and the low stakes associated with it. In addition, there are techniques that can help students during stressful situations (e.g. positive self-talk, breathing techniques, mindfulness) and teaching students these skills may support them during assessment periods. It is also important that teachers, principals and policymakers understand
the purpose and limitations of NAPLAN and know that the data should not be considered the primary measure of a school’s effectiveness.

At the student level, the conversation about national testing is often independent from the conversation about internal testing in schools. This study argues that internal tests have a greater impact on students than NAPLAN. Yet, both NAPLAN and internal testing polarise student responses, with students who have strong academic self-concept experiencing positive responses, and other students either reporting negative responses such as feeling stressed or feeling bored. Because of this, it is possible that the process of testing benefits high achieving students over low achieving students. This can be observed in the broader context, as it has been previously reported that high-stakes testing may be a means through which academic success is transferred into social power among culturally advantaged families (Teese, 2000).

Overall, this study found that most secondary school students completed NAPLAN with minimal stress and are more concerned with their internal tests. Therefore, it argues that the policy debate about the value of NAPLAN should focus on the validity and reliability of the results, how results are used, and if the benefit to students and teachers is worth the cost of the program, rather than discussion about the stress it causes students.

Therefore, it can be argued that in its current form, contrary to findings from some other researchers, NAPLAN does not appear to have a direct negative repercussions in terms of student emotional responses for most students. There are issues regarding the limitations of the test, the creation of league tables and comparison mechanisms such as the My School website whether the allocation of resources to NAPLAN could be spent more efficiently. There may be many problems with NAPLAN, but for most secondary school students, negative emotional responses to the test do not appear to be a major one.
References


Australian College of Educators. (2010). *Submission to the Inquiry into the Administration and Reporting of NAPLAN Testing*. Canberra, ACT: Senate References Committee on Education, Employment & Workplace Relations.
Australian Curriculum, Assessment and Reporting Authority. (2010). NAPLAN Why NAP. 
NAP: National Assessment Program. Retrieved from

Australian Curriculum, Assessment and Reporting Authority. (2011). The tests. NAP: 

Australian Curriculum, Assessment and Reporting Authority. (2018). My School: A new and 

Australian Primary Principals Association. (2013). Primary principals: Perspectives on 


Education Policy Review, 113 (1), 17-25

Shepard, L. A. (2010). Problems with the use of test scores to evaluate teachers 


Freeman.


Education and Employment Reference Committee (2014). *Effectiveness of the National Assessment Program – Literacy and Numeracy (Final Report)*, Canberra, Australia.

Emery, K & Ohanian, S 2004, *Why is corporate America bashing our public schools?* Portsmouth, NH.


https://doi.org/10.1016/s0092-6566(03)00046-1


https://doi.org/10.1353/tip.2003.0005


https://doi.org/10.14507/epaa.v12n39.2004


https://doi.org/10.1016/j.stueduc.2011.03.003


https://doi.org/10.1080/00094056.1991.10521597


Preston, B. (2018). *The social make-up of schools, Family income, indigenous status, family type, religion, languages spoken, disability, home internet access, housing tenure, and geographic mobility of students in public, Catholic and independent schools*. A report for the Australian Education Union, ACT


https://doi.org/10.1162/edfp.2009.4.4.492


Sabol, R. F. (2010). No Child Left Behind: A study of its impact on art education. West Lafayette, IN: Purdue University, Indiana


Appendix A: Ethics Approval

14 July 2015

Dr S.M. Rice

Melbourne Graduate School of Education, The University of Melbourne

Dear Dr Rice,

I am pleased to advise that the Humanities and Applied Sciences Human Ethics Sub-Committee approved the following Project:

Project title: Student perceptions of the impact of high-stakes testing on their experience of school

Researchers: Dr S M Rice, M Dowley

Ethics ID: 1443196

The Project has been approved for the period: 14-Jul-2015 to 31-Dec-2015.

It is your responsibility to ensure that all people associated with the Project are made aware of what has actually been approved.

Research projects are normally approved to 31 December of the year of approval. Projects may be renewed yearly for up to a total of five years upon receipt of a satisfactory annual report. If a project is to continue beyond five years a new application will normally need to be submitted.

Please note that the following conditions apply to your approval. Failure to abide by these conditions may result in suspension or discontinuation of approval and/or disciplinary action.

(a) Limit of Approval: Approval is limited strictly to the research as submitted in your Project application.
(b) Variation to Project: Any subsequent variations or modifications you might wish to make to the Project must be notified formally to the Human Ethics Sub-Committee for further consideration and approval. If the Sub-Committee considers that the proposed changes are significant, you may be required to submit a new application for approval of the revised Project.

(c) Incidents or adverse effects: Researchers must report immediately to the Sub-Committee anything which might affect the ethical acceptance of the protocol including adverse effects on participants or unforeseen events that might affect continued ethical acceptability of the Project. Failure to do so may result in suspension or cancellation of approval.

(d) Monitoring: All projects are subject to monitoring at any time by the Human Research Ethics Committee.

(e) Annual Report: Please be aware that the Human Research Ethics Committee requires that researchers submit an annual report on each of their projects at the end of the year, or at the conclusion of a project if it continues for less than this time. Failure to submit an annual report will mean that ethics approval will lapse.

(f) Auditing: All projects may be subject to audit by members of the Sub-Committee.

If you have any queries on these matters, or require additional information, please contact me using the details below.

Please quote the ethics registration number and the title of the Project in any future correspondence.

On behalf of the Sub-Committee I wish you well in your research.
Acting Secretary Humanities and Applied Sciences HESC

Phone: 83442073, Email: k.murphy@unimelb.edu.au

RESEARCH, INNOVATION & COMMERCIALISATION

Office for Research Ethics and Integrity

The University of Melbourne, Victoria 3010, Australia

T: +61 3 8344 1539 (external) T: 40777 (internal) W: orei.unimelb.edu.au
10 August 2015

Mr Mark Dowley

3/797 Toorak Road, HAWTHORN EAST VIC 3123

Dear Mark

RE: Research in Catholic Schools in the Diocese of Sale

Thank you for your emailed application dated 29 July 2015 in which you have submitted documents to conduct research entitled Student Perceptions of the Impact of Assessment on their Experience of school involving one Catholic secondary school in the Diocese of Sale. Whilst I approve the survey instruments as documented, I do not give permission for you to distribute incentives to either students or staff in the participating schools. Consequently I request that any reference to this practice is removed from documents before it is distributed to Catholic schools in the Diocese of Sale.

I am happy for you to approach the schools in this diocese as indicated in your application. It is important that you understand that the final permission for you to undertake this work rests with the Principal.

As previously notified this in principle approval is subject to the attached Research in Catholic Schools -Standard Conditions. In particular, please note that, it is a requirement for researchers working on a one-to-one basis with children in schools, to present a current Working With Children Check (item 3) to the school principal.

Should you require further information please contact Marg Shiels at this Office, email mshiels@ceosale.catholic.edu.au or phone 5622 6648.

With best wishes

Yours sincerely,

Maria Kirkwood

DIRECTOR OF CATHOLIC EDUCATION

DIOCESE OF SALE

Faith…Learning…Growth

6 Witton Street Warragul VIC 3820 -PO Box 322 Warragul 3820 -Phone: (03) 5622 6600 -Fax: (03) 5623 4258 -Email: director@ceosale.catholic.edu.au
Appendix B: Plain Language Statement

Student Perceptions of the Impact of NAPLAN on Their Experience of School:

Plain Language Statement for Principals and Teachers

This research project is being conducted by the Melbourne Graduate School of Education at the University of Melbourne.

Background

National testing programs such as NAPLAN have a range of consequences for schools, teachers, parents and students. While there has been considerable research on the effects of such programs on curriculum and pedagogy, and on school focus, there has been little research that considers students’ perspectives on such programs. There is limited research into how tests may influence students’ experiences of school, and how students feel about undertaking the tests.

This research project seeks to add our current knowledge about national testing programs by seeking student views about their experiences. We are seeking to conduct student surveys and teacher interviews in each participating school. We would like to survey students from two Year 9 classes and two Year 7 classes in each school and interview one teacher from each of these year levels.

Benefits

The study will be the first large-scale examination of students’ experiences of NAPLAN from their own perspective, and will consider how expectations from parents,
schools and teachers influence these experiences. It will also fill a gap in international literature regarding how school, parental and student expectations shape student experiences of tests, and provide information that may allow schools and parents to frame their expectations so as to maximise student learning, motivation and wellbeing.

Findings from the study may support schools to identify those who could potentially be at risk of experiencing higher levels of stress from national and internal testing. This identification will be the beginning of a process of creating strategies to lessen these negative consequences and will identify factors associated with positive responses to test-based assessment.

Participants

As part of this research students from your school will be asked to participate in a survey that will take approximately 20 minutes. The surveys will be conducted online during a convenient school time using one of either a phone/computer/tablet and the students participating in the survey will be completely anonymous. Copies of the survey and the interview questions are attached for your information.

Any information the students provide is strictly confidential, within the limitations of the law, and all confidential data will be kept in a secure location at the University for five years and then destroyed. No individual persons or organisations will be identified by name in the database that will be developed or in the text of any report or article arising from the research. No comments will be attributed to specific informants. Participation in the research is voluntary and students can withdraw any time by speaking to their teacher or the researchers. When the research is complete an outline of the findings will be mailed to participating schools. Any data collected will only be used for the purpose of the research project.
This project has been approved by the Human Research Ethics Committee at the University of Melbourne and the Department of Education and Training. If you have any further questions, please do not hesitate to contact Dr Suzanne Rice on 8344 0950. If you have any other concerns about the conduct of this research project, please contact the Executive Officer, Human Research Ethics, the University of Melbourne, on phone: (03) 8344 2073 or fax: (03) 9347 6739.
Appendix C: Consent Forms—Principals and Parents

Melbourne Graduate School of Education

Consent Form for Schools Participating in a Research Project

Student Perceptions of the Impact of NAPLAN on Their Experience of School

Name of participant school:

Name of investigator(s): Mr. Mark Dowley, Dr. Suzanne Rice

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I understand that after I sign and return this consent form it will be retained by the researcher.

3. I understand that my participation will involve a survey at the school and I agree that the researcher may use the results as described in the plain language statement.

4. I acknowledge that:
   
   (a) The possible effects of participating in the survey have been explained to my satisfaction;

   (b) I have been informed that I am free to withdraw my school from the project at any time without explanation or prejudice and to withdraw any unprocessed data I have provided;

   (c) The project is for the purpose of research;

   (d) I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements;

   (f) Any names will be referred to by a pseudonym in any publications arising from the research.

Principal signature: ___________________________ Date: ______________

HREC: 1443196.1: 16/05/16; Version: 1.1

Melbourne Graduate School of Education
The University of Melbourne Victoria 3010 Australia

T: +61 3 8344 8285 F: +61 3 8344 8529 W: www.education.unimelb.edu.au
Melbourne Graduate School of Education

Consent Form for Parents of Children Participating in a Research Project

Student Perceptions of the Impact of NAPLAN on Their Experience of School

Name of student participant:

Name of investigator(s): Mr. Mark Dowley, Dr. Suzanne Rice

1. I consent for my child to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I understand that after I sign and return this consent form it will be retained by the researcher.

3. I understand that my child’s participation will involve a survey and I agree that the researcher may use the results as described in the plain language statement.

4. I acknowledge that:

   (a) the possible effects of participating in the survey have been explained to my satisfaction;

   (b) I have been informed that my child is free to withdraw from the project at any time without explanation or prejudice and to withdraw any unprocessed data my child has provided;

   (c) the project is for the purpose of research;

   (d) I have been informed that the confidentiality of the information my child provides will be safeguarded subject to any legal requirements.

Parent name:

Parent signature: Date:

HREC: 1443196.1: 16/05/16; Version: 1.1

Melbourne Graduate School of Education

The University of Melbourne Victoria 3010 Australia

T: +61 3 8344 8285 F: +61 3 8344 8529 W: www.education.unimelb.edu.au
Appendix D: Survey Instrument

Students’ Experiences of Assessment

We are interested in finding out more about how students experience tests like NAPLAN at school, and what they think of them. Thank you for taking the time to complete this survey.

1. First, we’d like to ask a few questions about yourself
   a. What is your year level? __________
   b. What is your gender? __________
   c. Do you speak another language at home? ________
      If so, what is it? __________________________________
   d. What is your home postcode? ________
   e. What type of school do you go to (government/grammar)?

2. Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other. Write the number next to the words.

<table>
<thead>
<tr>
<th>Disagree strongly</th>
<th>Disagree moderately</th>
<th>Disagree a little</th>
<th>Neither agree or disagree</th>
<th>Agree a little</th>
<th>Agree moderately</th>
<th>Agree strongly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

a. _____ outgoing, enthusiastic

b. _____ serious, stern

c. _____ trustworthy, self-disciplined

d. _____ anxious, easily upset

e. _____ open to new experiences, complex
f. _____ reserved, quiet

g. _____ sympathetic, warm

h. _____ disorganised, careless

i. _____ calm, emotionally stable

j. _____ predictable, uncreative.

**Now, we’d like to know more about your experiences of tests**

3. How important is it *to your teachers* for you to do well in NAPLAN?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Not important</th>
<th>A bit important</th>
<th>Important</th>
<th>Very important</th>
</tr>
</thead>
</table>

4. How important is it *to your teachers* for you to do well in internal tests?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Not important</th>
<th>A bit important</th>
<th>Important</th>
<th>Very important</th>
</tr>
</thead>
</table>

5. How important is it *to your parents* for you to do well in NAPLAN?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Not important</th>
<th>A bit important</th>
<th>Important</th>
<th>Very important</th>
</tr>
</thead>
</table>

6. How important is it *to your parents* for you to do well in internal tests?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Not important</th>
<th>A bit important</th>
<th>Important</th>
<th>Very important</th>
</tr>
</thead>
</table>

7. How important is it *to you* to do well in NAPLAN?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Not important</th>
<th>A bit important</th>
<th>Important</th>
<th>Very important</th>
</tr>
</thead>
</table>

8. How important is it *to you* to do well in internal tests?

<table>
<thead>
<tr>
<th>Not at all important</th>
<th>Not important</th>
<th>A bit important</th>
<th>Important</th>
<th>Very important</th>
</tr>
</thead>
</table>
Tell us how much you agree with the following statements:

9. I expect to do well in the NAPLAN English tests (spelling, language and persuasive writing) this year.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

10. I expect to do well in the NAPLAN mathematics test this year.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

11. I expect to do well in English in school this year.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

12. I expect to do well in mathematics at school this year.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
</table>

13. Thinking back to the last time you did NAPLAN, did you experience any of the following? You can say yes to more than one answer.

Circle Yes or No

a. Feeling nervous, Yes/No e. Crying, Yes/No
b. Sleepless the night before, f. Stressed, Yes/No
   Yes/No               g. Excited, Yes/No
c. Feeling sick before the test, h. Confident, Yes/No
   Yes/No               i. Enthusiastic, Yes/No
d. Freezing during the test, j. Bored, Yes/No
   Yes/No               k. Other? …………
14. Did you experience any of the following before your last internal test? (for example, a mathematics or English test)?

You can say yes to more than one answer.

a. Feeling nervous, Yes/No
b. Sleepless the night before, Yes/No
c. Feeling sick before the test, Yes/No
d. Freezing during the test, Yes/No
e. Crying, Yes/No
f. Stressed, Yes/No
g. Excited, Yes/No
h. Confident, Yes/No
i. Enthusiastic, Yes/No
j. Bored, Yes/No
k. Other? ............

15. Have your parents ever spoken to you about your NAPLAN results in the past?

Yes         No

If yes, was this to:

(You can say yes to more than one answer if you need to)

a. congratulate you, Yes/No
b. point of areas for improvement, Yes/No
c. suggest you work harder, Yes/No
d. point out why the tests are important, Yes/No
e. Other (please explain) ......................

16. Have your parents ever spoken to you about your internal test results in the past?

Yes         No
If so, was this to:

(You can say yes to more than one answer if you need to)

a. congratulate you, Yes/No
b. point of areas for improvement, Yes/No
c. suggest you work harder, Yes/No
d. point out why the tests are important, Yes/No
e. Other (please explain) …………………

17. Have your teachers ever spoken to you about your NAPLAN results in the past?

Yes No

If so, was this to: (You can say yes to more than one answer if you need to)

a. congratulate you, Yes/No
b. point of areas for improvement, Yes/No
c. suggest you work harder, Yes/No
d. point out why the tests are important, Yes/No
e. Other (please explain), …………………

18. Have your teachers ever spoken to you about your internal test results in the past?

Yes No

If so, was this to: (You can say yes to more than one answer if you need to)

a. congratulate you, Yes/No
b. point of areas for improvement, Yes/No
c. suggest you work harder, Yes/No
d. point out why the tests are important, Yes/No
e. Other (please explain) …………………

19. Have you ever seen your NAPLAN results in the past? – (Tick as many as apply)

☐ My teacher/s discussed them with me
20. Tick one. Which statement is most correct for you:

- Internal tests are more important to me than NAPLAN
- NAPLAN tests are more important to me than internal tests
- Both are equally important to me

Why?

21. How much emphasis does this school place on doing well in NAPLAN?

<table>
<thead>
<tr>
<th>Very little emphasis</th>
<th>Some emphasis</th>
<th>A lot of emphasis</th>
</tr>
</thead>
</table>

22. Have you completed any NAPLAN practice tests for at school this year?

- No, never
- 1–2 tests
- 3–5 tests
- 6 or more

23. Have you ever completed NAPLAN workbooks at home?

- Yes
- No

24. Have you ever done anything else at home to prepare for NAPLAN?

- Yes
- No

If so, please write what you’ve done

25. Has your school ever cancelled other classes to give your class time to practise for NAPLAN?

- Never
- Once or twice
- Three or more times

26. Which sentence best describes how you are going in mathematics? (Pick one)

a. I am doing really well in this subject
b. I’m not one of the very top students but I’m above average
c. I’m about average in this subject

d. I find this subject difficult

27. Which sentence best describes how you are going in English? (Pick one)

a. I am doing really well in this subject
b. I’m not one of the very top students but I’m above average
c. I’m about average in this subject
d. I find this subject difficult

28. How hard did you try in the NAPLAN mathematics test?

<table>
<thead>
<tr>
<th></th>
<th>I didn’t try at all</th>
<th>I put in a bit of effort</th>
<th>I put in a reasonable effort</th>
<th>I put in my best effort</th>
</tr>
</thead>
</table>

29. How hard did you try in NAPLAN spelling test?

<table>
<thead>
<tr>
<th></th>
<th>I didn’t try at all</th>
<th>I put in a bit of effort</th>
<th>I put in a reasonable effort</th>
<th>I put in my best effort</th>
</tr>
</thead>
</table>

30. How hard did you try in the NAPLAN language conventions (punctuation and grammar) test?

<table>
<thead>
<tr>
<th></th>
<th>I didn’t try at all</th>
<th>I put in a bit of effort</th>
<th>I put in a reasonable effort</th>
<th>I put in my best effort</th>
</tr>
</thead>
</table>

31. How hard did you try in the NAPLAN persuasive writing test?

<table>
<thead>
<tr>
<th></th>
<th>I didn’t try at all</th>
<th>I put in a bit of effort</th>
<th>I put in a reasonable effort</th>
<th>I put in my best effort</th>
</tr>
</thead>
</table>

32. How hard did you try in English tests at school?

<table>
<thead>
<tr>
<th></th>
<th>I didn’t try at all</th>
<th>I put in a bit of effort</th>
<th>I put in a reasonable effort</th>
<th>I put in my best effort</th>
</tr>
</thead>
</table>

33. How hard did you try in mathematics tests at school?

<table>
<thead>
<tr>
<th></th>
<th>I didn’t try at all</th>
<th>I put in a bit of effort</th>
<th>I put in a reasonable effort</th>
<th>I put in my best effort</th>
</tr>
</thead>
</table>
34. How are NAPLAN results used at your school? (Tick as many answers as apply)

☐ To place students in classes

☐ To decide who can enrol in particular subjects

☐ They aren’t used at all

☐ I don’t know.

35. How are the results from your internal tests used in your school? (Tick as many answers as apply)

☐ To place students in classes

☐ To decide who can enrol in particular subjects

☐ They aren’t used at all

☐ I don’t know.

36. Is there anything else you’d like to say about NAPLAN or internal tests?

___________________________________________________________________________

___________________________________________________________________________