Comparing student motivations for and emotional responses to national standardised tests and internal school tests: The devil in the detail

Abstract:

National testing of students has become an increasingly prevalent policy tool, often implemented to drive improvement through increased accountability and heightened competition between schools. Such testing has been found to generate negative emotional responses among students, including increased stress and anxiety (e.g. Thompson, 2013). However, there is little examining whether such responses are associated specifically with national testing regimes or are more general responses to testing situations. This study surveyed 206 students in Australian secondary schools to compare responses to NAPLAN and internal school tests. Students reported higher expectations for their performance in internal school tests than for NAPLAN, higher levels of boredom for NAPLAN, and greater levels of confidence for their internal school tests. While most students reported low levels of negative emotional responses to NAPLAN, a small group of students reported strong negative emotional responses to both NAPLAN and internal school tests, suggesting that negative responses to national testing programs may be more dependent on the individual student.

Word count: 8835 including references

Keywords: Testing programs, student experience, test anxiety, student attitudes, educational policy, student perceptions.

Introduction:

The drive for increased accountability for schools and teachers has led many education systems around the world to implement state and national testing regimes. Countries that have established national or state testing programs for school students include England, the United States, Scotland, Indonesia and Australia (Supovitz, 2009; Gorur, 2015). While some countries and states use sample testing, many opt for population testing that arguably provides a complete picture of student achievement and allows comparisons to be made between different groups of students (boys compared to girls, wealthier students compared to disadvantaged ones), between schools and between sectors or states.

Such large-scale programs require the investment of significant resources, and impact on millions of students and teachers. They have understandably been the subject of considerable research to determine how such testing programs influence pedagogies used, the curriculum, the focus and energies of principals and teachers, and students' experiences of schools. This body of research considers what the

positive and negatives of testing programs are, and whether the benefits of such testing programs balance or outweigh their potential and experienced downsides.

A range of potential positive effects from national and state testing programs have been identified in the literature. Supovitz (2009) notes that two key theoretical reasons for introducing national and state testing programs are that they will improve alignment within and across schools, and will increase teacher motivation to teach well and student motivation to learn. Improving alignment increases the likelihood that students will achieve essential competencies regardless of where they learn, and whether they transfer between schools, and some researchers have indeed reported increased alignment between the documented curriculum and the taught curriculum (Phelps, 2006; Clarke et al., 2003, Jones, 2007). Some researchers also report that testing does in fact lead to greater student motivation to study (Sloane & Kelly, 2003). Supovitz (2009) also notes that an additional reason for implementing testing programs is that they provide broad information about student achievement at national, state, school, teacher and student levels, and the argument that testing improves teaching and learning by providing this type of information is often used by testing authorities (e.g. see ACARA, 2020). However, while national and state testing regimes appear to increase the capacity of systems and schools to track changes in student achievement over time and adjust resourcing accordingly to address student need (Santiago et al., 2011), there is evidence that the information they provide is much less reliable or useful at the level of the teacher or individual student (Wu, 2015; Supovitz, 2009).

In addition to these positive impacts, the literature highlights a range of negative effects generated by national or state testing programs. These negative flow-on effects can be categorised into impacts on first, schools, second, teachers, third, teaching and fourth, students. At the school level, while national and state testing programs may sharpen the focus on what policymakers see as essential competencies, they have been found to be associated with reduced student access to and emphasis on curriculum areas that are not included within the testing program, encouraging teachers to teach to the test and narrowing the learning focus (Thompson, 2013; Thompson & Harbaugh, 2013; Polesel, Rice & Dulfer, 2014; Rentner et al., 2006). Some research has found evidence of reduced time spent on non-tested areas (such as the arts) and there is evidence that these types of effects may be most keenly felt in disadvantaged schools, which may be under most pressure to bring their students “up to scratch” (Thompson & Harbaugh, 2013).

Also at the school level, other research indicates that, particularly where high stakes are attached to schools or students doing well in the tests, national and state testing programs can lead to various forms of cheating and gaming the system (Vasquez Heilig & Darling-Hammond, 2008; Jacob & Levitt, 2003). Where poor-performing schools are named and shamed through publication of results, schools may use strategies such as suspending poorly-performing students for the duration of the testing period to inflate

artificially the school's scores (Nichol, Glass & Berliner, 2007). In the longer term, schools may use tactics such as encouraging low-achieving students to drop out (Vasquez Heilig, Young, & Williams, 2012), placing them into unassessed special education classes (Allington & McGill-Franzen, 1992; Figlio & Getzer, 2002), or diverting resources away from students who will not be assessed (Yiu, 2020), all forms of cream-skimming.

There is also evidence of an impact on teachers themselves. Some studies have found that national and state testing programs are associated with a reduction in teachers' morale and sense of professional autonomy (Crocco & Costigan, 2007; Wright & Choi, 2005). Thompson's (2013) respondents indicated that NAPLAN tests in Australia had generated increased pressure on teaching staff. Ashadi and Rice (2016) documented impacts on teachers' professional opportunities and career pathways in Indonesia, with those teachers perceived to be more competent or better qualified being assigned to classes sitting for the national tests and being offered more extensive professional learning than other teachers in the same schools.

With regard to teaching, the research suggests that, within curriculum areas targeted for testing, national and state testing programs can lead to a focus on low-level skill acquisition, rather than deeper and more meaningful learning (Chang et al., 2013; Thompson & Harbaugh, 2013; Macmillan, 2005). Lobascher (2011) found that national and state testing programs were associated with use of "defensive pedagogies" that are teacher-centred and reduce student activity and agency. This effect has since been found in other research with teachers on their teaching practice following the introduction of national and state testing regimes (e.g. Polesel, Rice and Dulfer, 2014).

Finally, there are the impacts on students. Concerns have been raised about negative emotional impacts on students due to the implementation of national and state testing, and there is a considerable body of research into this topic. Paris and McEvoy (2000) reported that students experienced anxiety and freezing in response to national tests, while exhaustion, crying and frustration responses were found among some children by Madaus, Russell and Higgins (2009). Athanasou (2011) reported that wellbeing issues, such as the inability to sleep, were associated with NAPLAN tests among Australian primary students. Other Australian research with teachers finds that they identify NAPLAN as having a negative impact on student wellbeing (Cummings et al., 2018), with reports of increased student anxiety and stress, or even crying (Thompson, 2013; Rice, Dulfer, Polesel & O'Hanlon, 2015). Howell's (2017) research with primary school students in Australia found that NAPLAN tests were construed as high stakes by students and associated with negative emotional responses such as anxiety, even though the policy intention is that such tests are low stakes for students. Cummings et al. (2018) found that around one third of students reported feeling worried or sad before and during NAPLAN testing. However, it should be noted that some studies find no such effects: Rogers, Barblett & Robinson (2016)

found only slightly increased stress levels among the great majority of students sitting for national tests, with high stress levels reported by only around 3% of participating students.

The finding that national and state testing programs may impact negatively on student wellbeing by causing distress has been a major point of contestation in discussions around the programs, with the media particularly highlighting research with these types of findings (e.g. Bell, 2018; Bowden, 2014; BBC, 2018). However, while there has been a considerable body of research that has sought to explore student emotional reactions to national and state testing, this research has had several limitations. The first is that much (although not all) of the research has sought the views of administrators, teachers, and parents about student responses to tests and is thus second-hand and does not capture the students' perspectives. As Rogers et al. (2016) note, it is possible that teachers and parents with negative attitudes towards state and national testing programs perceive greater levels of distress among children and adolescents than actually exist, or that parents and teachers who have observed student distress disproportionately volunteer for research on the impact of testing programs. The second limitation is that the research on student reactions has for the most part assumed that students' emotional responses to national and state testing programs are a function of the imposition of such tests rather than potentially being a more general reaction that some students have in response to any testing situation. Both points are important. It is possible that teachers and parents are drawing conclusions about the degree of distress students feel in relation to testing, that may not accurately reflect students' own perceptions of their responses. It is also possible that students who have negative emotional responses to large-scale high-stakes testing may be at least as distressed by regular internal school testing, and thus state or national testing programs are placing little further emotional burden on such students that they do not already experience in school.

Segool et al. (2013) did seek to compare elementary school student anxiety responses to *No Child Left Behind* tests with their responses to regular in-school testing. This study found significantly higher anxiety among students in response to NCLB tests compared to in-school tests. But there appears to be little else examining this issue. Further, Howell's (2017) study suggests that younger students might respond more anxiously to national testing than older students because they have little understanding of how the tests are used or what the consequences of testing are. This aligns with the research of Rice et al. (2015), which found that negative emotional responses to NAPLAN testing were reported much more frequently by teachers of students in primary/elementary schools than those in secondary schools.

Finally, there is relatively little that seeks to explain theoretically why students might respond to tests (whether national or regular internal assessment experiences in schools) in the ways that they do, and how their perceived need and motivation to do well in tests might inform such responses.

The current study sought to address these gaps in the literature. The study drew on Wigfield and Eccles' (2000) expectancy-value theory to examine student responses to testing. Expectancy-value theory posits that our motivation for a goal will be a function of the value we place on the goal, and the expectancy we have of being able to attain it with effort. If, for example, a person strongly desires a goal, but has very low expectations of being able to attain it, motivation will be low. If the person perceives that they can easily attain a goal, but places little value on it, motivation is also low. In contrast, where the person places a high value on a goal, and has a strong expectation that they will be able to attain it with effort, motivation will be high.

We expanded upon Wigfield and Eccles' theory by hypothesising that students' emotional responses to testing (both NAPLAN and internal school tests) would be related to their expectancies and values. We hypothesised that where students place a high value on doing well in a test, they are more likely to report strong emotional responses (both positive and negative) to testing. We further hypothesised that students with high expectancies to do well would be more likely to report positive emotional responses to testing. We also hypothesised that students reporting high levels of motivation as expressed as effort in a test would be more likely to report positive emotional responses to testing, while students reporting low levels of motivation (as expressed by effort) would be more likely to report boredom in response to testing.

We sought to compare students' emotional responses to the Australian national literacy and numeracy tests (NAPLAN) and their internal school assessments in mathematics and English. We predicted that student expectancies, values and motivation would influence student emotional responses to testing for both national tests and formal school based assessments.

The research questions framing the study therefore were:

- Are students' self-reported levels of value, expectancy and motivation (as reported effort) the same for national testing (NAPLAN) and internal school tests?
- Are students' emotional responses to national testing (NAPLAN) and to internal school tests the same? If not, what differences are evident?
- How are student self-reported levels of value, expectancy and motivation (as self-reported effort) related to the types of responses students report in relation to testing?

Context for the study:

The study was conducted in the state of Victoria. The Australian National Assessment Program (Literacy and Numeracy), or NAPLAN, was established in Australia in 2008 and assesses all students in Years 3, 5, 7 and 9 (that is, students who are usually 8, 10, 12 and 14 years of age respectively). NAPLAN is an annual test that measures students' proficiency in reading, writing, language conventions (spelling, grammar and punctuation), and numeracy through five separate tests, administered across three days. Test results are sent to parents and schools approximately five months later, although online testing for some schools began in 2019 and is expected to produce a faster turnaround time as it is rolled out across the country¹. Criterion-referenced reports are delivered to parents giving an indication of how an individual student's achievement compares to a typical student in their year, and schools receive both individual and aggregate achievement data for all tests. School-level results are also made public on the national *My School* website, and the level of achievement of students in each school is mapped on the website against 59 similar schools in a format that was designed to be easily read by the public (Gorur, 2015), so that parents can use this information to make a judgement about how well their child's school is performing. While the *My School* website was designed to make it difficult to construct league tables of school performance, media outlets quickly developed school ranking tables ('league tables') following the release of NAPLAN results (Gorur, 2015).

An individual student's NAPLAN results are for the most part simply used to provide information to parents and inform teaching, although there are reports of some schools using NAPLAN results to inform selection for entry into the school or into specific academic streams or subjects (for example, advanced maths courses). Overall, NAPLAN might be considered low-stakes, then, for most of the students undertaking the tests. However, the publication of school results online, and the potential for this to impact on school prestige and student enrolments, together with the common use of NAPLAN results by departments of education to monitor school performance, means that the tests are arguably high-stakes for schools, principals and teachers.

Materials and methods:

Approach:

A total of 15 schools was targeted to participate in the study. Secondary schools were targeted for two reasons. First, students at secondary schools have considerably more experience with NAPLAN tests, having already completed them at primary school in Years 3 and 5. Second, secondary schools conduct more formal and frequent internal tests with students that are overtly labelled as tests, allowing a better

¹ It should be noted that NAPLAN was suspended for 12 months in 2020 due to the impact of COVID-19 on schools.

comparison to be made between students' perceptions of their reactions to NAPLAN and their school's internal tests than would be possible with primary school students.

Schools were contacted initially by an email that included the Plain Language Statements for the project and sample consent forms. The email was followed up with a telephone call a week later to determine whether the school was interested in participating. The study targeted students in Years 7 and 9 as these are the years that undertake NAPLAN tests, and survey was conducted in the month after NAPLAN to ensure that students' memories of the experience were recent. Where a school indicated its willingness to participate in the study, a member of the research team visited the school to speak with students, provide the link to the online survey and answer any questions from students and teachers. Survey completion took around 15 minutes, and was undertaken during class time in the presence of the teacher.

Students were offered an incentive of a small confectionary item on returning the parental permission form to increase participation; they received this incentive regardless of whether or not their parent agreed to their participation (i.e. the incentive was used simply to ensure the signed form was returned). No other incentives were involved.

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. Participation in the student survey was anonymous, and individual schools have not been identified.

Sample

To facilitate the research, schools where the researchers had an existing professional relationship were targeted. Of the fifteen schools approached, four agreed to participate in the study. The sample of schools included single sex and co-educational schools from rural and metropolitan areas. Two participating schools were government schools and two were Catholic; while several independent schools were approached, none was able to participate in the study.

Participants consisted of 131 Year 7 students and 75 Year 9 students and the sample had a greater number of male students (131) than female students (71), given that one participating school was a boys-only school. A breakdown of schools and participants can be found in Table 1.

Table 1
Sample characteristics

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

Note. ¹ The ICSEA (Index of Community Socio-Educational Advantage) is a scale measuring the level of educational advantage in a school. ICSEA values are calculated on a scale that has a mean of 1000 and a standard deviation of 100. The higher the ICSEA value, the higher the level of educational advantage for students at the school (ACARA, 2020).

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 school students in this study compared to the state's 17 per cent. The schools in the study were all around average socio-economic status, as indicated by the ICSEA.

Instrument

The survey instrument used in this study built on 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). The developed survey was trialled with a small sample of students in a non-participating school and adjusted prior to use with the schools participating in the study.

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

- demographic information (age, gender)
- student perceptions of teacher attitudes to assessment (NAPLAN, school tests)
- student perceptions of parent attitudes to assessment (NAPLAN, school tests)
- student expectations of own performance in assessment (NAPLAN, school tests)
- student levels of motivation for the assessment (NAPLAN, school tests). This was assessed through self-reported effort ("How hard did you try...") for each test.
- student self-reported value for doing well in the assessment (NAPLAN, school tests)
- student experience of feedback from the tasks (NAPLAN, school tests)
- student experience of preparation for the tasks (NAPLAN, school tests)
- student self-efficacy in the assessed subjects
- emotional responses to tests (negative, positive and neutral) for NAPLAN and for the most recent school test in the equivalent subject
- personality (as measured by the Ten-Point Personality Inventory TIPI scale (Gosling et al., 2003)).

The survey also included some brief open-ended sections for students to make any additional comments about their experiences of NAPLAN and school testing. Data from the surveys were transferred to and analysed in SPSS, using paired t-tests, McNemar's tests, and Gamma and chi-square analyses.

Results

Comparing student value, expectancy and motivation for NAPLAN and internal school tests

We first sought to compare students' responses to the national standardised tests, and their internal school tests. We compared the value students placed on doing well, their expectancy that they would do well, and their level of motivation as measured by self-reported effort in NAPLAN and internal school tests.

To determine the value participants placed on doing well in the two types of test, participants were asked two questions, 'How important is it to you, for you to perform well on NAPLAN tests?' and "How important is it to you, for you to perform well on internal tests?", responding on a 5 point Likert scale ranging from 1 (Not at all important) to 5 (Very important) Table 2 compares the responses to these survey items.

Table 2

Comparison of the Value Students Placed on Doing Well in NAPLAN and Internal Tests

	1: Not at all important	2: Not important	3: A bit important	4: Important	5: Very important	Mean	SD
NAPLAN	12%	10%	23%	31%	24%	3.32	1.37
Internal tests	2%	1%	8%	28%	60%	4.46	0.84

A paired t-test found that there was a significant difference between the value students placed on doing well in NAPLAN ($M = 3.32$, $SD = 1.37$) and the value students placed on 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' to them. Twelve percent of students selected the 'Not at all important' option regarding NAPLAN, compared to only two percent of students choosing the 'Not at all important' option for internal tests.

We also compared how well students expected they would do in each of the two forms of test (NAPLAN and internal). There were separate questions for subject domains, given that students may be strong maths students and weak English students, or vice versa. Students were asked about the extent to which they agreed with statements about their expectancies to do well in the four NAPLAN tests (mathematics, language conventions, writing, reading) and in internal English and mathematics tests (e.g. "I expect to do well in NAPLAN mathematics this year", "I expect to do well in mathematics tests in school this year"). Students responded on a 5-point Likert scale from (1) 'Strongly disagree' to (5) 'Strongly agree'. To improve comparability, the three NAPLAN English domains were averaged to create a student expectancy in NAPLAN English composite variable. This was done as internal school English tests often assess multiple domains including reading, writing and language conventions. Table 3 shows the student expectancies in mathematics for NAPLAN and internal tests and Table 4 shows the average expectancies of students across the three domains in English.

Table 3

Student Expectation to do Well in Mathematics for NAPLAN and Internal Tests

	1:	2:	3:	4:	5:	Mean	SD
	Strongly disagree	Disagree	Neither Agree or Disagree	Agree	Strongly agree		
Maths (NAPLAN)	5%	7%	23%	34%	31%	3.76	1.11
Maths (Internal)	4%	4%	19%	31%	43%	4.02	1.04

Table 4

Student Expectations to do Well in English (composite) for NAPLAN and Internal Tests

	1:	2:	3:	4:	5:	Mean	SD
	Strongly disagree	Disagree	Neither Agree or Disagree	Agree	Strongly agree		
English (NAPLAN)	4%	6%	27%	42%	22%	3.68	1.02
English (Internal)	2%	4%	17%	40%	29%	3.93	0.94

A paired sample t-test found students reported significantly higher expectations to do well in internal mathematics tests ($M = 4.02$, $SD = 1.04$) than in NAPLAN mathematics ($M = 3.76$, $SD = 1.11$, $t(205) = -2.52$, $p = 0.006$, $d = 0.24$). A second paired sample t test found students reported significantly higher expectations to do well 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 types of test (NAPLAN and internal school). However, 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 school tests (8 per cent).

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. Given that both value and expectancy were higher for internal tests than for NAPLAN, it was anticipated that student motivation as measured by self-reported effort would be higher for internal tests than for NAPLAN.

Participants were asked to respond on a 5-point Likert scale to the questions on effort in each of the NAPLAN tests (e.g. “How hard did you try in the NAPLAN spelling test?”) and on usual effort in school-based tests (e.g. “How hard do you usually try in English tests at school?”). As with expectancy, the three responses to the three NAPLAN English test items were averaged to form a composite variable for motivation in NAPLAN English. Student responses are outlined in Tables 5 and 6 below.

Table 5

Self-reported Motivation for NAPLAN Mathematics Tests compared to Motivation for Internal Mathematics Tests

	1: I didn't try at all	2: I put in a small effort	3: I put in an average effort	4: I put in a large effort	5: I put in my best effort	Mean	SD
NAPLAN mathematics	5%	5%	12%	32%	47%	4.05	1.13
Internal mathematics tests	2%	1%	8%	24%	65%	4.52	0.82

Table 6

Self-reported Motivation for NAPLAN English Tests (composite) compared to Motivation for Internal English Tests

	1: I didn't try at all	2: I put in a small effort	3: I put in an average effort	4: I put in a large effort	5: I put in my best effort	Mean	SD
NAPLAN English	4%	4%	17%	27%	49%	4.09	1.12
Internal English tests	1%	2%	8%	32%	56%	4.39	0.86

A paired sample t-test found students reported significantly higher motivation for internal mathematics tests ($M = 4.52$, $SD = 0.82$) than for NAPLAN mathematics ($M = 4.05$, $SD = 1.13$, $t(205) = -4.83$, $p < 0.001$, $d = 0.48$). A paired sample t-test found students reported significantly higher motivation for internal English tests ($M = 4.39$, $SD = 0.86$) than NAPLAN English ($M = 4.09$, $SD = 1.12$, $t(205) = -3.45$, $p < 0.001$, $d = 0.30$). 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 mathematics. The pattern is similar although not as strong

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 tests.

These findings were consistent with what would be predicted by our application of expectancy–value theory. Students placed more value on doing well in internal tests than doing well in NAPLAN, had higher expectations of doing well in internal tests than in NAPLAN, and were therefore more motivated for and reported putting in more effort for internal tests of both English and mathematics.

Students in this study demonstrated higher expectancies to do well, value placed on doing well and motivation for internal school tests than for NAPLAN tests. We now compare the emotional responses of students to each form of assessment.

A Comparison of Student Emotional Responses to NAPLAN and Internal Tests

To understand better whether the negative student responses to national and state testing programs are specific to those testing situations, or occur as readily in in-school tests, we compared student responses to internal tests and NAPLAN. We also sought to determine whether there were other emotional responses to testing among students, not just negative emotions such as nervousness.

Students were asked: 'Thinking back to the last time you completed a [NAPLAN/Internal] test, which of the following did you experience?'. The response items comprised three positive emotional responses to testing (feeling confident, excited, enthusiastic), one neutral item (feeling bored) and six negative emotional responses (feeling nervous, feeling stressed, experiencing sleeplessness the night before, feeling sick, freezing during the test, and crying). Students were asked to tick any items they had experienced, and were able to select more than one option.

Table 7 shows the response to these survey items.

Table 7

Students' Reported Emotional Responses to NAPLAN and Internal School Tests

	NAPLAN	Internal tests	p value
Nervous	48%	57%	.025*
Confident	29%	43%	<.001***
Stressed	27%	34%	.044*
Bored	52%	25%	<.001***
Excited	17%	25%	.005**
Enthusiastic	18%	21%	.243
Sleepless the night before	13%	10%	.424
Freezing during the test	11%	8%	.701
Feeling sick before the test	10%	7%	.332
Crying	3%	2%	.250

Note: As students were able to select more than one answer, 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.

For positive emotional responses to testing, 43 per cent of students in the sample reported feeling confident before their internal tests, compared to 29 per cent reporting confidence prior to sitting NAPLAN ($p < .001$). More students reported feeling excited before internal tests than before NAPLAN ($p = .005$). There were no statistically significant differences between students feeling enthusiastic prior to NAPLAN and internal tests.

For the negative emotional responses, there was a small but statistically significant difference between students feeling nervous prior to NAPLAN and internal tests, with fewer students reporting feeling nervous for NAPLAN than for internal tests ($p = .025$). Students were also more likely to report feeling stressed for internal school tests than for NAPLAN ($p = .044$).

The greatest difference in student responses to the two forms of testing was that students were much more likely to report being bored for NAPLAN than for internal tests, with almost double the number of students reporting feelings of boredom associated with doing NAPLAN compared to internal tests ($p < .001$). Small percentages of students reported feeling sick, freezing, and sleeplessness the night before for both NAPLAN and for internal tests. There were no significant differences in the number of students reporting these responses for NAPLAN and internal tests.

Relationships between value, expectancy and motivation, and student emotional responses to tests:

We also sought to examine the relationships between value, expectancy, motivation and student emotional responses to tests. We consider each of these below in relation to the NAPLAN tests.²

Expectancy: The association between student expectancy and emotional responses to NAPLAN was assessed using Goodman and Kruskal's gamma. Analysis indicated that student expectancy to do well 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. Expectancy to do well in NAPLAN English had a negative relationship with reporting boredom ($G=-0.523, p<.001$). There was no relationship between expectancy to do well and students reporting negative physical responses of sleepless, feeling sick, freezing in the test or crying.

In relation to NAPLAN mathematics, there were similar strong positive relationships between expectancy to do well and positive emotional 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 to do well in NAPLAN mathematics and boredom ($G=-0.529, p<.001$). No relationships were found between expectancy and students reporting negative physical responses.

In summary, 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 than students who had less expectation to do well.

Value:

Analysis of the value students place on doing well in NAPLAN and their emotional responses indicated that the value students placed on doing well was positively correlated with positive emotional responses, including feeling excited ($G=0.687, p<.001$), confident ($G=0.564, p<.001$) and enthusiastic ($G=.523, p<.001$). There was also a strong positive correlation between feeling nervous ($G=.442, p<.001$), and the value students placed on doing well in NAPLAN, yet there was no statistically significant relationship between the value students placed on doing well in NAPLAN and feeling stressed. The only significant negative correlation was between the value students placed on doing well 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 tests.

There were no statistically significant relationships between the value students placed on doing well and the other negative responses (sleeplessness, feeling sick, freezing during the test or crying). In sum, students who

² For reasons of space, we report these results only in relation to NAPLAN tests, but similar relationships were found between value, expectancy, motivation and emotional responses for internal school tests.

placed a higher value on doing well in NAPLAN were more likely to experience feeling excited, confident and enthusiastic, but also more likely to report feeling 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.

Motivation

Motivation for performing well in NAPLAN as measured by reported effort was correlated with positive responses, but not negative responses. Students responded to the questions, ‘How hard did you try in NAPLAN mathematics tests this year?’ and ‘How hard did you try in NAPLAN English tests this year?’. Relationships between reported motivation and students’ emotional responses were investigated using Goodman and Kruskal’s Gamma. Motivation as assessed by self-reported effort in NAPLAN English was positively correlated with students reporting feeling excited ($G=.370$, $p=.012$), confident ($G=.350$, $p=.004$) and enthusiastic ($G=.354$, $p=.029$). There were similar positive correlations between reported motivation in NAPLAN Mathematics and feeling excited ($G=.347$, $p=.003$), confident ($G=.435$, $p<.001$) and enthusiastic ($G=.339$, $p=.039$). The only significant negative relationship was between motivation to do well in NAPLAN mathematics and feeling bored ($G=.293$, $p=.032$). There were no statistically significant relationships between the motivation students have for NAPLAN and the other negative responses (stressed, feeling nervous, sleepless the night before, feeling sick, freezing or crying). As would be predicted by our application of expectancy value theory, students with greater levels of motivation for NAPLAN are more likely to report positive emotional responses to the tests.

Discussion:

National and state testing programs are now an important feature of the education policy landscape in many countries, whose implementation consumes significant resources, and shapes students’ experiences of school and learning. For these reasons research into the impact and efficacy of such programs is vital and ongoing. Arguably, as the key stakeholders in the endeavour of education, the impact on students should be a key focus of this research.

As national testing programs have been implemented around the globe, numerous studies have reported that students suffer emotionally from them as evidenced by their negative emotional responses to the tests – high levels of student stress, crying, sleeplessness and so on (Cummings et al., 2018; Athanasou, 2010; Simpson, 2018). In some countries, these reports have been strong enough to lead to calls to abandon or modify significantly national testing programs (see for example, Bell, 2018; BBC, 2018).

However, testing in various forms is a part of students’ experiences in school. It is possible that some students respond negatively to any testing situation, and removing national testing programs will not alter this response, simply provide one less occasion on which the response will occur. There is little in the research that helps us understand whether students’ observed and reported negative emotional responses to national and state testing programs are created by these programs, or if they are simply the response that some students have to any form

of testing, whether this testing is national or carried out by the school as part of its routine instructional practices. The extent of such responses among students is also somewhat under-researched, and has often been determined through seeking parents' and teachers' views on the proportion of students affected, and to what degree. Students' voices have been far less frequently captured.

There is also a need for a much more nuanced understanding of why students respond as they do to testing. Missing in many of these debates on national testing is an understanding of how students think about testing, the stakes that are associated with various tests, the value that students place on doing well, and their own understanding of how likely they are to achieve a result that matters to them. Further, these values, expectancies and levels of motivation are likely to link strongly to the types of emotional responses students have to testing. No matter what the testing program or situation, if a student does not value doing well in a test, they are unlikely to be very highly stressed by sitting it.

This study sought to examine and compare student responses to national tests and internal school tests, and to examine how the value they placed on doing well, their expectancies to do well, and their levels of motivation for the tests were related to their emotional responses, within an Australian context. The first finding was that the negative emotional responses that have featured in international debates around national and state testing programs were in fact more likely to be experienced for internal school tests than for national tests. Students in our study were more likely to report both nervousness and feeling stressed prior to internal school tests than prior to NAPLAN. The most striking difference between student responses to the two forms of testing was that there was a much greater proportion of students reporting feelings of boredom in response to national testing than to internal school tests.

The study also considered positive emotional responses to testing, both for national and internal school tests, something that is not routinely covered in the extant literature. Significant numbers of students reported feeling enthusiastic or excited in response to testing, possibly because they enjoy the challenge, or because they are strong students who anticipate doing well and having a strong academic self-concept validated by their test results. These positive emotional responses were more frequent for internal school tests than for national tests, but suggest that for some students at least, testing may be a positive rather than negative experience.

As predicted by our application of expectancy-value theory, students who expected to do well in a test were more likely to report feeling excited, enthusiastic and confident, and students who placed a high value on doing well in a test were similarly more likely to report feeling excited, enthusiastic and confident. More concerning was the finding that students who were not expecting to do well, and those who placed a low value on doing well, were more likely to report feelings of boredom in relation to doing the test. It is possible that weaker students might be responding to their perceptions of their own academic capacity by devaluing doing well and emotionally switching off from an experience that might confirm any feelings of inferiority. This type of disengagement has the capacity to becoming a vicious circle of low performance leading to low motivation and lower future performance. Further research is needed to confirm if this is the case.

The study further found that students' self-reported levels of effort were lower for national tests than they were for their school's internal tests, which, together with large numbers of students reporting feeling bored in national tests, raises important questions about the positioning of such tests both in policy and in public debates. National and state standardised tests are frequently presented by politicians and the media as objective, reliable, and accurate, with the implication that school-level teacher-developed tests and teacher judgements about student progress are not, or at least not to the same degree. For example, the rationale for Scotland's Scottish National Standardised Assessments (SNSA) program argues that they have been introduced to "provide teachers with *objective* [authors' emphasis], comparable information about [student] progress" (SNSA website, n.d.). Writers such as Au (2013) have noted how this positioning has the effect of reframing structural inequalities within education as being due to individual differences in merit. However, in addition to such concerns, the higher levels of boredom, and lower levels of effort for national tests than for their internal school tests that students reported in this study suggest that their achievement in national tests might potentially be a poor indicator of their actual level of ability. Depending on students' beliefs around national and state tests, and the value they place on doing well, it may be that the results from supposedly objective externally-developed standardised national and state tests are in fact a *less* accurate indicator of students' achievement levels than results from their internal, teacher-developed school tests.

Our research also found that some of the more extreme documented negative student emotional responses to national testing (for example, sleeplessness) may be limited to a relatively small group of students for whom both national and internal school tests are stressful experiences. This contrasts with some previous research indicating that large numbers of students are unduly distressed by national and state testing programs (see for example, Rice et al., 2015), although the findings are in keeping with the work of Rogers, Barblett & Robinson (2016), who found that only small numbers of students reported high stress levels for national tests. There are two potential reasons that might explain this discrepancy. First, as noted earlier, much of the research on student emotional responses draws on the perspectives of teachers, school leaders, psychologists and parents, but does not incorporate student views (see for example Heiser et al., 2015, cited in Simpson, 2018). It may be that adults overestimate the level of distress that these programs cause for the students (adults may particularly notice upset students and foreground these students in their thinking) or there may be sampling bias in that parents or teachers strongly against national and state testing may disproportionately volunteer to participate in research on the topic. Second, different national and state testing programs have different stakes attached to them. In some instances, stakes may be very low or not valued by students, which would suggest that students are unlikely to be distressed by the tests, while in other contexts in which the stakes for students are much higher and more highly valued by students (for example, a student who does not achieve highly enough might have to repeat a year of schooling) may potentially lead to much higher levels of student anxiety, nervousness, sleeplessness and so on. NAPLAN has relatively low stakes for students, so our findings may simply be a reflection of this.

This also suggests the need for further research on national and state testing programs that incorporates and critically examines the stakes attached to doing well in any testing program, and the value that is placed on

those stakes by the full range of stakeholders. Often such research assumes that those involved in testing will value tests and the stakes associated with them equally, and will respond in similar ways. However, different stakeholder groups – students, parents, teachers, school leaders and policymakers – will have different stakes in the outcomes of these tests, and within each of these groups, the value placed on the stakes is likely to vary among individuals. How these stakes differ between stakeholders, the role that the value placed on stakes plays, and the effects of these aspects on responses to testing, often remain unexamined. There is a need to expand nuanced approaches that involve systematic examination of the stakes and the value placed on them to understand fully the impact of such programs on all involved.

The study also has some limitations that point the way to future research. The research was conducted with students in secondary schools, and there is some evidence that suggests that younger students, with a much less clear understanding of the stakes associated with tests and with less exposure to formal testing in their day-to-day school experience than older students, might be more strongly affected emotionally by national testing programs (e.g. Howell, 2017; Rice et al., 2015). Findings from this research need to be verified with a broader range of students, including younger students and those at senior secondary and tertiary levels of education, whose responses might differ from those outlined here. The low stakes associated with NAPLAN tests for students (but not for teachers and principals) in Australian schools also suggest that student responses, especially the extent of negative emotional responses to national tests, would differ, perhaps greatly, in states and countries where there are much higher stakes attached to doing well for the students. Again, these differences need to be explored in greater detail by subsequent research.

Finally, the study has implications for policymakers and schools. For schools, the findings indicate that there is one small group of students that is likely to find formal testing – whether in national testing programs or during normal school testing of student learning – stressful and upsetting. The findings also suggest that for some lower achieving students, testing is associated with feelings of boredom and reported low effort, potentially leading to a cycle of lower student motivation, engagement and achievement. School psychologists and teachers need to be aware of such students and offer them appropriate support to minimise these effects, and the literature suggests a range of potential interventions (e.g. von der Embse, 2013). For policymakers, the findings suggest that some of the concerns about negative student emotional responses to testing might not hold, at least for older students in situations where the tests are low-stakes for them.

This is not to say that there are not many valid reasons to question the use of high-stakes testing programs, as the literature documents many other negative impacts of testing programs. These include distortion of curricula and pedagogy (e.g. Avalos, Perez & Thorrington, 2020), the failure of such programs to reduce inequalities in achievement (e.g. Vasquez Heilig, Jameson Brewer & Ojeda Pedraza, 2018) or their ability to mask such inequality by promoting notions of objectivity and merit (Au, 2013). Rather, the negative student emotional impacts generated by such testing might not, under some circumstances, be a major reason for their discontinuation.

Acknowledgements: This article reports on findings that arose from work undertaken as part of the first author's Doctor of Education thesis. The authors would like to acknowledge the feedback of Professor David Gurr and Associate Professor Lawrie Drysdale in shaping this study and article. There was no funding grant associated with the study.

References:

- Allington, R. L., & McGill-Franzen, A. (1992). Unintended effects of educational reform in New York. *Educational Policy*, 6(4), 397–414.
- Ashadi, A. & Rice, S. (2016). High stakes testing and teacher access to professional opportunities: Lessons from Indonesia. *Journal of Education Policy*, 31(6), 727-741.
- Athanasou, J. (2010). *NAPLAN and MySchool survey*. Barton, Australian Capital Territory: Independent Education Union of Australia.
- Au, W. (2013). Hiding behind high-stakes testing: Meritocracy, objectivity and inequality in US education. *The International Education Journal: Comparative Perspectives*. 12(2) 7-19. Retrieved from [EJ1017702.pdf \(ed.gov\)](#)
- Author. (2019). *The Impact of Testing on Students: Australian Students' Perspectives on NAPLAN and Internal Assessments*. Unpublished Doctor of Education thesis.
- Australian Curriculum Assessment and Reporting Authority (ACARA) (2020). *Guide to understanding the index of community socio-educational advantage*. Retrieved from <https://www.myschool.edu.au/media/1820/guide-to-understanding-icsea-values.pdf>
- Avalos, M. A.; Perez, X. & Thorrington, V. (2020). Comparing secondary English teachers' ideal and actual writing practices for diverse learners: Constrained professionalism in figured worlds of high-stakes testing. *Reading & Writing Quarterly*, v36 n3 p225-242 2020
- Bell, S. (2nd August, 2018) Scottish Government urged to abandon 'cruel' standardised testing following accounts of pupils' distress Retrieved from <https://sourcenews.scot/scottish-government-urged-to-abandon-cruel-standardised-testing-following-accounts-of-pupils-distress/>
- Bowden, T. (2014) NAPLAN study finds testing causes students anxiety, program not achieving original goals. Australian Broadcasting Corporation website. Retrieved from <https://www.abc.net.au/news/2014-05-19/naplan-study-finds-school-testing-program-not-achieving-goals/5463004?nw=0#:~:text=NAPLAN%20study%20finds%20testing%20causes%20students%20anxiety%2C%20program%20not%20achieving%20original%20goals,-7.30&text=A%20new%20study%20on%20the,is%20having%20unintended%20negative%20consequences.>
- British Broadcasting Corporation (BBC) (28th June, 2018). National tests bordering on child abuse, claims head. Retrieved from <https://www.bbc.com/news/uk-wales-44633354>
- Chang, M. C., Al-Samarrai, S., Ragatz, A.B., Shaeffer, S., De Ree, J., & Stevenson R. (2013). *Teacher reform in Indonesia: The role of politics and evidence in policy making*. Washington, D.C.: World Bank Publications.

- Clarke, M., Shore, A., Rhoades, K., Abrams, L., Miao, J. & Li, J. (2003). Perceived effects of state-mandated testing programs on teaching and learning: findings from interviews with educators in low-, medium-, and high-stakes state', National Board on Educational Testing and Public Policy, Boston College, Retrieved from <http://www.bc.edu/research/nbetpp/statements/nbr1.pdf>
- Crocco, M. & Costigan, A. (2007). The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, 42(6), 512-535.
- Ferguson-Patrick, K. (2018). The importance of teacher role in cooperative learning: The effects of high-stakes testing on pedagogical approaches of early career teachers in primary schools. *Education 3-13*, 46(1), 89-101.
- Figlio, D. N., & Getzer, L. S. (2002). *Accountability, ability, and disability: Gaming the system?* Cambridge, MA: National Bureau of Economic Research.
- Gorur, R. (2015). The performative politics of NAPLAN and MySchool. In B. Lingard, G. Thompson, & S. Sellar (Eds). *National Testing in Schools: An Australian Assessment*. Melbourne: Routledge. Pp. 30-43.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality*, 37, 504-528.
- Howell, A. (2017) Because then you could never ever get a job!': children's constructions of NAPLAN as high-stakes, *Journal of Education Policy*, 32(5), 564-587, DOI: [10.1080/02680939.2017.1305451](https://doi.org/10.1080/02680939.2017.1305451)
- Jacob, B. & Levitt, S. (2003). *Catching cheating teachers: The results of an unusual experiment in implementing theory*. National Bureau of Economic Research. NBER Working Paper 9414. Retrieved from <http://www.nber.org/papers/w9414>
- Jones, B.D. (2007). The unintended outcomes of high-stakes testing, *Journal of Applied School Psychology*, 23(2), 65-86.
- Lobascher, S. (2011). What are the potential impacts of high-stakes testing on literacy education in Australia? *Australian Journal of Language & Literacy*, 34(2), 9-19.
- Madhaus, G., Russell, M. and Higgins, J. (2009). *The paradoxes of high-stakes testing: How they affect students, their teachers, principals, schools and society*. Charlotte NC.: Information Age Publishing.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, 14.
- Paris, S. (2000). Trojan horse in the schoolyard: The hidden threats in high stakes testing, *Issues in Education*, 6, 1/2, 1-8.
- Paris, S. & McEvoy, A. (2000). Harmful and enduring effects of high stakes testing, *Issues in Education*, 6, No. 1/2, 145-160.

- Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*, Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.
- Phelps, R.P. (2006) Characteristics of an effective student testing system, *Educational Horizons*, 85(1), 19-29.
- Polesel, J., Rice, S. & Dulfer, N. (2014). The impact of high stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of Education Policy* 29(5), 640-657.
doi:10.1080/02680939.2013.865082.
- Rentner, D. S., Scott, C., Kober, N., Chudowsky, N., Chudowsky, V., Joftus, S., et al. (2006). *From the capital to the classroom: Year 4 of the no child left behind act*. Washington, DC: Center on Education Policy.
- Rice, S., Dulfer, N., Polesel, J. & O'Hanlon, C. (2015). NAPLAN and student wellbeing: Teacher perceptions of the impact of NAPLAN on students. In B. Lingard, G. Thompson, & S. Sellar (Eds). *National Testing in Schools: An Australian Assessment*. Melbourne: Routledge.
- Rogers, S. L., Barblett, L., & Robinson, K. (2016). Investigating the impact of NAPLAN on student, parent and teacher emotional distress in independent schools, *The Australian Education Researcher*. DOI 10.1007/s13384-016-0203-x
- Rogers, S., Barblett, L., Robinson, K., (2018), Parent and teacher perceptions of NAPLAN in a sample of Independent schools in Western Australia. *The Australian Educational Researcher*, 45(4), 493-513, DOI: 10.1007/s13384-018-0270-2
- Santiago, P., Donaldson, G., Herman, J. and Shrewbridge, C. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia*, Paris: OECD.
- Segool, N., Carlson, J., Goforth, A., von der Emse, N., Barterian, J. (2013). Heightened test anxiety among young children: Elementary school students' anxious responses to high-stakes testing. *Psychology in the Schools*, 50(5), 489-499.
- Sloane, F. and Kelly, A. (2003) Issues in high stakes testing, *Theory into Practice*, 42 (1), 12-17.
- Simpson, C. (2018). *Effects of standardized testing on student well-being*. Retrieved from [c. simpson effects of testing on well being 5 16.pdf \(harvard.edu\)](#)
- Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform, *The Journal of Educational Change*, 10(2-3)211-227.
- Thompson, G. (2013). NAPLAN, MySchool and accountability: Teacher perceptions of the effects of testing. *The International Education Journal: Comparative Perspectives*, 12(2), 62–84.

- Thompson, G., & Harbaugh, A. G. (2013). A preliminary analysis of teacher perceptions of the effects of NAPLAN on pedagogy and curriculum. *Australian Educational Researcher*, 40, 299–314.
- Vasquez Heilig, J., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75–110
- Vasquez Heilig, J., Young, M., & Williams, A. (2012). At-risk student averse: Risk management and accountability. *Journal of Educational Administration*, 50(5), 562–585.
- Vasquez Heilig, J., Jameson Brewer, T., Ojeda Pedraza, J. (2018). Examining the myth of accountability, high-stakes testing, and the achievement gap. *Journal of Family Strengths* 18(1). Retrieved from ["Examining the Myth of Accountability" by Julian Vasquez Heilig, T. Jameson Brewer et al. \(tmc.edu\)](#)
- Von der Embse (2013). Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000-2010. *Psychology in the Schools*, 50(1), 57-71.
- Walker, T. (2014). NEA Survey: Nearly half of teachers consider leaving profession due to standardized testing. Retrieved from <http://neatoday.org/2014/11/02/nea-survey-nearly-half-of-teachers-consider-leaving-profession-due-to-standardized-testing-2/>
- Wigfield and Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81. <http://dx.doi.org/10.1006/ceps.1999.1015>
- Wright, W. & Choi, D. (2005). Voices from the classroom: A statewide survey of experienced third-grade English language learner teachers on the impact of language and high-stakes testing policies in Arizona. Retrieved from [ERIC - ED508521 - Voices from the Classroom: A Statewide Survey of Experienced Third-Grade English Language Learner Teachers on the Impact of Language and High-Stakes Testing Policies in Arizona, Language Policy Research Unit, 2005-Dec](#)
- Wu, M. (2015). What national testing data can tell us. In B. Lingard, G. Thompson, & S. Sellar (Eds). *National Testing in Schools: An Australian Assessment*. Melbourne: Routledge.
- Yiu, L. (2020). Educational injustice in a high-stakes testing context: A mixed methods study on rural migrant children's academic experiences in Shanghai public school. *Comparative Education Review*, 64(3), 498-524

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Dowley, M;Rice, S

Title:

Comparing student motivations for and emotional responses to national standardised tests and internal school tests: The devil in the detail

Date:

2022-01-03

Citation:

Dowley, M. & Rice, S. (2022). Comparing student motivations for and emotional responses to national standardised tests and internal school tests: The devil in the detail. AUSTRALIAN JOURNAL OF EDUCATION, 66 (1), pp.92-110. <https://doi.org/10.1177/00049441211061889>.

Persistent Link:

<http://hdl.handle.net/11343/296010>