The Effect of Language, Gender and Age in NAPLAN Numeracy Data

Tim Wilson
Monash University
<tim@connectbt.com.au>

Tasos Barkatsas
RMIT University
<tasos.barkatsas@rmit.edu.au>

This study investigates the relationship between students ability to answer reduced language dependency mathematical questions with their overall numeracy level. It investigates whether a student's success at reduced language mathematical questions translates into better overall numeracy scores. It was found, students have up to two years advancement if able to correctly answer reduced language dependency questions. This phenomenon was clearly apparent in the overall findings, but was most pronounced at the Year 3 level test, and for female students.

Some researchers believe a large proportion of schools are still focussing on mathematical skills, algorithms and processes to get the correct answer at the expense of comprehension and problem solving skills (Perso, 2009). Other interest groups believe that teaching of mathematics should move back to basics (Perso, 2007). Furthermore, many in both the academic community and the general community have an underlying belief that boys are better than girls at mathematics (Hargreaves, Homer, & Swinnerton, 2008). However some researchers struggle to show significant differences between boys and girls in regard to numeracy (Hargreaves et al., 2008, Lachance & Mazzocco, 2006).

This study seeks to examine part of this debate. The data used to compare students' is based on the National Assessment Program Literacy and Numeracy (NAPLAN) results. Students who found the reduced language dependency (RLDM) questions difficult, were compared to the students who were able to correctly answer the same questions.

A broad aim of this research is to attempt to verify a correlation between student's ability to answer RLDM questions and their nationally measured numeracy level? Are people justified when they complain questions on the numeracy papers are not fair (Perso, 2009)? Are there statistically significant differences between girls and boys in their grasp of numeracy? Another goal of this research is to determine whether further simplification of literary requirements in NAPLAN numeracy testing may be warranted. For example, students with language backgrounds other than English (LBOTE) and student with English language reading, understanding or learning difficulties may have better mathematical knowledge than their NAPLAN scores indicate (Abedi & Lord, 2001). The significance of this information may lead to investigations regarding the teaching of basic tools during the earlier years of schooling. Other outcomes of this research may suggest further investigation of better pedagogical processes for teaching mathematical concepts to different students groups.

Literature Review

Although numeracy has various different meanings and usages (Sullivan, 2011), it is not just another name for school mathematics (AAMT, 1997). Numeracy is sometimes described as quantitative literacy (Clements, Bishop, Keitel, Kilpatrick, & Leung, 2013), and as such is a priority for mathematics education (Aubrey, Dahl, & Godfrey, 2006). The Australian Ministerial Council for Education, Employment, Training and Youth Affairs

2014. In J. Anderson, M. Cavanagh & A. Prescott (Eds.). Curriculum in focus: Research guided practice *(Proceedings of the 37th annual conference of the Mathematics Education Research Group of Australasia)* pp. 653–660. Sydney: MERGA.

(MCEETYA) have defined numeracy to be the effective use of mathematics for participation in life (Ministerial Council on Education, Employment, Training and Youth & Affairs (MCEETYA), 1997). In Australia, numeracy benchmarks are designed to improve accountability and inform parents and the broader community regarding individual numeracy achievement.

Some teachers do not always believe that teaching students to read and interpret questions is part of teaching numeracy (Perso, 2009). Answering questions correctly involves reading comprehension, identifying mathematics that will help, choices about 'how', doing it and analysing whether the solution is contextually sensible (Perso, 2009). Some students do not effectively understand numeracies mathematical situations, and are unable to move to abstract thinking from concrete thinking (Aubrey et al., 2006). Lesh (2000) claims the difficulty with traditional worded problems is that students have to understand described situations of which it is difficult to make meaningful symbolic descriptions (Lesh, 2000).

Over the last twenty years, mathematics curriculum has de-emphasised procedural and computational skills in favour of an emphasis on deep student understanding of ideas and contextual usage (Monash University, 2007). Unfortunately this trade-off has not translated into benefits for Australian students (Monash University, 2007). Furthermore, questions have been raised about whether hands-on activities lend a realism to the mathematics being taught, or even cause cognitive confusion for students when they struggle to relate between the analogue and the quantity (Boulton-Lewis, 1998).

There are clear implications regarding literacy issues associated with learning mathematics (Monash University, 2007). One-step written problems must be read, comprehended, transformed to a mathematical model, processed and then encoded. The correct answer is usually not obtained if there is a failure at any step (Clements, 1980). A student's lack of familiarity with the context within which a problem is set, has the potential to be a disadvantage (Monash University, 2007) or conversely a student's very familiarity with a context may be a hindrance to obtaining the correct answer (Boaler, 1994). Furthermore, familiarity with language can be associated with socio-economic factors and of course students' language background.

Another factor that challenges numeracy learners is that mathematical linguistic structure is often different from everyday usage of language. Some words, such as sum and fraction are mathematical words, but others, such as borrow and product have special meaning in mathematics (Schleppegrell, 2007). Knowing words such as *more* and *less* is insufficient in numeracy situations, the language pattern associated with these words must be both explicitly taught, and learnt (Schleppegrell, 2007, Perso, 2009). Greenlees (2010) noted the difficulty students experienced with the word 'fewer' in national testing, whereas most students in her study were correctly able to understand the same question when 'fewer' was replaced with 'less.'

The literacy demands required for understanding of some item content within NAPLAN numeracy tests has been raised as of particular concern (Perso, 2009). However, some degree of literacy skill combined with mathematical skill is required to be numerate. Over the last 20 years, this combination of skills has been part of the national curriculum, so poor NAPLAN results suggest to some researchers that lessons are still focusing on processes and algorithms to the neglect of comprehension and problem solving skills (Perso, 2009).

Gender, SES and LBOTE Differences

Liu and Wilson (2009) examined reasons for gender performance differences across mathematical domains using the U.S. portion of a large international mathematics study conducted in 2000 and 2003. They refuted previous findings that suggested male superiority in standard multiple-choice items and show that the largest gender difference is with complex multiple-choice item types, and in the Space and Shape domain. Other academics argue that evidence of continuing and re-emerging gender differences warrant ongoing attention and research into gender and mathematics education (Vale & Bartholomew, 2008).

Since at least the mid-1980s, researchers have been pointing to ways that language is implicated in the learning of mathematics (Schleppegrell, 2007). Even proficient speakers of English face challenges from the language of mathematics (Schleppegrell, 2007). The language factor in mathematics tests is reportedly not a significant issue by gender according to Adebi and Lord (2001). English language learners scored lower than native speakers, as did low socio-economic status (SES) learners in Abedi and Lord's (2001) testing of 1,174 eighth grade students. The impact of SES seems to have a greater impact than gender. McConney and Perry (2010) found that increases in school SES are consistently associated with substantial increases in science and mathematics performance.

Design and Methodology

In an attempt to make sense of an aspect of the world, the epistemology used within this research is a natural science lens as a post-positivist (Boden et al., 2005). The outcome of the research is based on selecting the answers to questions, which are specifically chosen to test the hypotheses. The ontological position taken is realism. This position holds that the existence of an object is independent of the observer (Cohen, Manion, & Morrison, 2007). It is expected that the quantitative approach used in this study will give reliable evidence that may refine and clarify current understanding of our social world (Phillips, 2009).

One of the limitations of quantitative research is with regard to sample selection. This study uses 'convenience sampling' with the full understanding that it only approximates the wider population. Within the context of this research, the participation of students was not directly voluntary, however, identifying features of individuals were removed from the data, before the researcher used it, thus no direct ethical considerations between participant and research was seen to exist. The first hypothesis for this study was that students' ability to successfully answer questions containing reduced language dependency questions is directly proportional to their National Benchmark numeracy score obtained in NAPLAN testing. The second hypothesis of this study is that there are no statistically significant differences between male students and female students in their ability to successfully answer reduced language dependency questions as demonstrated by their National Benchmark numeracy score obtained in NAPLAN testing.

Implementation

A simple model of data can be obtained by using the mean as a summary (Field, 2005). In this research, we compared means of different groups. The manipulation between groups is their performance on a type of question within the NAPLAN numeracy tests, and the variable being measured is the overall numeracy score obtained. The type of design for this scenario is a repeated-measures design (Field, 2005). If this experimental manipulation is considered successful, then the sample groups should appear to have come from different populations (Field, 2005). In other words, the 95% confidence interval should not overlap between the two groups (Field, 2005).

The data selected for this study were based on a convenient sample, rather than a random sample. The data used is from six small independent schools. The schools include a rural school, and schools in the east, west and south west of Melbourne. The student cohort included a wide ethnic mix and therefore LBOTE are well represented in the sample, however there are no indigenous students represented by the sample data. This data is not considered truly representative of the population of Australian students, however it does explore whether further research is warranted. The sample was made up of 1106 students, 556 female, and 550 male. The selection of test papers from which the questions were drawn was made randomly. The following papers were chosen: 2008 Year 3, Year 5 and Year 9, 2009 Year 9 calculator allowed, 2010 Year 7 and Year 9 both calculator allowed and not allowed, 2011 Year 5 and Year 7 both calculator allowed and not allowed. An example question is: "Add 17 to find the next number in this pattern. 41, 58, 75,__ (Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA), 2008a) While it was found to be quite difficult, and perhaps slightly subjective, to select NAPLAN test questions that contain reduced language dependency, the questions selected were chosen based on 'a minimal role' for the worded part of the question in the solution process. The goal in selecting questions was to choose those that could be answered by a student with limited understanding of English.

Presentation and Analysis

The total sample NAPLAN participants used were 1106 students in total from five schools in Victoria. The sample contained 550 male students and 556 female students. At Year 3 level, there were 95 students made up of 48 male and 47 female. At Year 5 level, there were 289 students made up of 143 male and 146 female. At Year 7 level, there were 306 students made up of 162 male and 144 female. At Year 9 level, there were 416 students made up of 197 male and 219 female.

Participants were placed in a group who were successful at reduced language dependency questions (SRLDQ) if at least one of their answers to the RLDM questions were correct. Participants were placed in a group who were unsuccessful at RLDM questions (URLDQ) if at least on of their answer to the RLDM questions were incorrect. Many students appear in both outcome groups. This will result in there being a higher total combined count of students than the total participating in the study, but reduces the effect of *ability*.

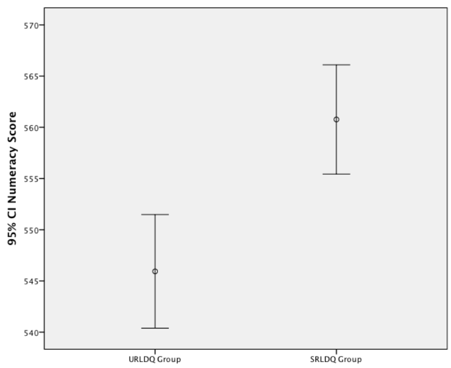

Figure 1. Combined Levels

The graph shown (Figure 1) for all students at all levels clearly indicates the presence of differences between the means of the two groups. This suggests that the manipulation is successful. The corresponding graphs for all students at Year 3, 5 and 7 (not shown) also indicated clear difference between the two groups. At the Year 9 level, however, the graphs are overlapping which does not suggest as clear a difference between the two groups at this level. Thus in the case of Year 9, even though there appears to be differences between the means of each group, it is possible the population mean for each group could be the same value because of the overlap in the statistical error (Field, 2005).

The graph of all female students at Years 3 and the graph of all female students at Year 5 show clear difference between the two groups. At Year 7 and 9 the graphs showed that the difference between the two groups is less clear, and although the means of each group are different, the population mean value could be the same. The graph of only male students at Year 3 (Figure 9) shows clear difference between the two groups, indicating that it is more than 95% likely the population means would also show two distinct means and therefore two distinct groups. The Year 5 and combined results are exceedingly close to showing distinct groups within the population, however the Year 5, 7 and 9 graphs clearly overlap thus allowing the possibility of the same mean value within the population for both groups.

These graphs for the different cases studied suggest that at the lower year levels, further analysis is definitely warranted, but that the differences between the two groups, although seemingly visible, become less distinct, and not quite as convincing as students mature.

Comparison of Means – Independent t-tests

This section reports results on the 1106 students (550 boys and 556 girls) who answered the reduced language dependency mathematical questions. To remove the effect of *ability*, 956 of these students were in the URLDQ group, and 1043 were in the SRLDQ group.

The breakdown of the numeracy scores of these groups revealed that showed that successful RLDQ students had statistically significantly higher scores (Table 1.)

Table 1
Population Statistical Results

	Successful RLDQ Group	Unsuccessful RLDQ Group
Whole Group [Mean Diff. = −14.832, t(1997) = −3.783, p < 0.01]	$M_{(SRLDQ)}$ = 560.77 $SD_{(SRLDQ)}$ = 87.702 $SE_{(SRLDQ)}$ = 2.716	$M_{(URLDQ)}$ = 545.93 $SD_{(URLDQ)}$ = 87.392 $SE_{(URLDQ)}$ = 2.826
Female Group [Mean Diff. = −14.004, t(1001) = −2.614, p < 0.01]	$M_{(SRLDQ)}$ = 553.89 $SD_{(SRLDQ)}$ = 84.327 $SE_{(SRLDQ)}$ = 3.702	$M_{(URLDQ)}$ = 539.88 $SD_{(URLDQ)}$ = 85.300 $SE_{(URLDQ)}$ = 3.877
Male Group [Mean Diff. = −15.440, t(994) = −2.708, p < 0.01]	$M_{(SRLDQ)}$ = 567.58 $SD_{(SRLDQ)}$ = 90.488 $SE_{(SRLDQ)}$ = 3.953	$M_{(URLDQ)}$ = 552.14 $SD_{(URLDQ)}$ = 89.151 $SE_{(URLDQ)}$ = 4.103

Comparison of Means at Year 3-9 Level showed that successful RLDQ students had statistically significantly higher scores at all levels until Year 9.

Discussion of outcomes/findings

The main finding from this research is that students who are able to successfully complete reduced language dependency mathematics questions will on average do better at numeracy testing than students who cannot. It informs teaching; such that time spent covering number and symbol based mathematics may translate to better overall numeracy skills in most students. The study supports that girls are hardest hit if they lack understanding or skills in reduced language dependency mathematics, and that all students are affected most during their early years of schooling, but that the effect becomes less as students reach mid high school.

The outcomes also suggest that the national numeracy testing, which includes separate technology assisted and unassisted papers should perhaps include reduced language dependency mathematics papers or sections, and that the results be separated out to include reduced language dependency mathematics results as well as numeracy results in the same way separate reading, writing, grammar and spelling scores are made for the literacy components of NAPLAN testing.

Achievement Differences

The state average NAPLAN numeracy score from 2008 to 2011 at the Year 3 level was 414, and at Year 5 level was 498, a difference of 84 over the two years of schooling. Given the finding that the difference between female students who could successfully answer reduced language dependency mathematics questions and those who were unsuccessful was 75.318, it suggests a magnitude of almost two years schooling difference in achievement level. The male difference of 53.286 is also concerning, appearing to be more than one year difference in achievement level.

These results confirm the findings of Boaler (1994) who also reported contextualized mathematics as particularly damaging to female students. The outcomes show significant differences between male and female students at the Year 3 level, however these differences between genders appear to be minimal from Year 5 onwards although NAPLAN numeracy results consistently show male students scoring slightly higher than female students.

Another outcome of these results is that two questions on a Year 3 NAPLAN paper have selected students who have statistically significantly better numeracy outcomes; above the national mean by more than a Year level.

Another study of this nature should be undertaken in the future, but with some of the inherent bias and effects ameliorated. A larger sample size of possibly 1000 students per Year level would result in more confidence regarding the population. Even testing at the population level may be possible if access can be obtained to the data. Another change to a future study would involve more precise test item selection, and better methods of reducing or eliminating the effects of other variables, such as *ability*.

Recommendations

The major recommendation to the education community suggested by this study would be that research regarding causality of the results be undertaken as soon as possible. Some questions that need to be considered are:

Can students better understand the numeracy requirements of assessments if they have an increased understanding of the mathematics behind the situational problem? Are assessment questions, especially at the lower grade levels, based on too complex literacy requirements? Should schools introduce another subject called numeracy? How much numeracy is really taught in other subjects than mathematics? Do teachers of other subjects have the skill and confidence to teach numeracy adequately within their classes? Is it equitable to assess a student who may have high mastery of mathematical manipulation skills but low mastery of the English language as having low numeracy outcomes because of how we assess mathematics in Australia? Should NAPLAN testing provide a 'mathematics' score as well as a numeracy score? Given that cognitive development of students occurs at different rates, are the numeracy test questions appropriately framed? Do girls need different methods of instruction to boys?

Given that early year success at mathematics translates into better outcomes in later years (Geary, 2011), mathematics educators should be very eager to research ways of moving students from the less able to the more able group identified by this research.

Conclusion

The results of this study seem most compelling for younger students. The research also suggests that those who recommend the education profession concentrate on teaching the basic mathematical skills may have a point. The research shows disparity between girls and boys in this study. The female students who understood how to process RLDM questions were significantly better at test questions than their peers who did not.

The long held academic belief that contextualising learning aids educational outcomes (Dalgarno & Lee, 2010) might not be completely correct in learning Mathematics, especially during early years. There is no doubt about a link between engagement and learning (Marks, 2000), but the mathematical literacy involved in some NAPLAN questions may be too complex for the early years of schooling. It pre-supposes that the student is able to understand the words and imagine the context.

Cooper and Dunne (2000) found that service-class children performed better than working and intermediate-class children on 'realistic' items compared to 'esoteric' items. They also found that boys performed better than girls on the 'realistic' items, and thus the biggest difference between service-class boys and working-class girls (Cooper & Dunne, 2000). However, NAPLAN numeracy-based year level testing assumes that assessment should be made in a form that parallels the real world, that all Australian students should be able to understand the context as expressed, that the question adequately expresses the context and that students will not superimpose their own understanding of the context in ways unexpected by the question author.

The efficacy of adjusting the school subject currently called 'Mathematics' into two subjects, mathematics and numeracy, should be considered by educators. Also the addition of a NAPLAN domain called mathematics is strongly suggested by the findings of this study. Within the noted limitations and biases of this study, the overall finding that students who are good at reduced language dependency mathematics will on average get a higher NAPLAN numeracy score by at least a year level is a certainty. The statistical variable of being able to manipulate symbolic mathematics makes a significant difference to student outcomes.

References

AAMT. (1997). *AAMT Policy on Numeracy Education in Schools*. Australian Association of Mathematics Teachers Inc. Retrieved from http:// www.aamt.edu.au/content/download/724/19518/file/numpol.pdf

Abedi, J., & Lord, C. (2001). The Language Factor in Mathematics Tests. *Applied measurement in education*, *14*(3), 219–34.

Aubrey, C., Dahl, S., & Godfrey, R. (2006). Early Mathematics Development and Later Achievement: Further Evidence. *Mathematics Education Research Journal*, *18*(1), 27–46.

Boaler, J. (1994). When Do Girls Prefer Football to Fashion? An Analysis of Female Underachievement in Relation to "Realistic" Mathematics Context. *British educational research journal*, *20*(5), 551–64.

Boden, R., Kenway, J., & Epstein, D. (2005). The research process. In *Getting started on research* (pp. 36–49). London; Thousand Oaks, CA: New Delhi: Sage Publications.

Clements, M. (1980). Analyzing Children's Errors on Written Mathematical Tasks. *Educational Studies in Mathematics*, *11*(1), 1–21. doi:10.2307/3482042

Clements, M., Bishop, A., Keitel, C., Kilpatrick, J., & Leung, F. (2013). *Third International Handbook of Mathematics Education*. New York; London: Springer.

Cohen, L., Manion, L., & Morrison, K. (2007). Validity and reliability. In *Research methods in education* (6th ed., pp. 133–164). London : Routledge.

Cooper, B., & Dunne, M. (2000). *Assessing Children's Mathematical Knowledge - Social class, sex and problem-solving*. Buckingham: Open University Press.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

Field, A. (2005). *Discovering Statistics Using SPSS* (Second Edition.). London: SAGE Publications Ltd.

Geary, D. (2011). Cognitive Predictors of Achievement Growth in Mathematics: A 5-Year Longitudinal Study. *Developmental Psychology*, *47*(6), 1539–1552.

Greenlees, J. (2010). The Terminology of Mathematics Assessment. In *Shaping the future of mathematics education*. Freemantle: MERGA. Retrieved from http://www.merga.net.au/publications/counter.php?pub=pub_conf&id=849

Hargreaves, M., Homer, M., & Swinnerton, B. (2008). A comparison of performance and attitudes in mathematics amongst the "gifted". Are boys better at mathematics or do they just think they are? *Assessment in Education: Principles, Policy & Practice*, *15*(1), 19–38.

Lachance, J. A., & Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences*, *16*(3), 195–216. doi:doi: DOI: 10.1016/j.lindif.2005.12.001

Lesh, R. (2000). Beyond Constructivism: Identifying Mathematical Abilities that are Most Needed for Success Beyond School in an Age of Information. *Mathematics Education Research Journal*, *12*(3), 177–195.

Liu, O. L., & Wilson, M. (2009). Gender Differences in Large-Scale Math Assessments: PISA Trend 2000 and 2003. *Applied Measurement in Education*, *22*(2), 164–184.

Marks, H. (2000). Student engagement in instructional activity: patterns in teh elementary, middle, and high school years. *American Education Research Journal*, *37*(1), 153–184. doi:10.3102/00028312037001153

McConney, A., & Perry, L. (2010). Science and Mathematics Achievement in Australia: The Role of School Socioeconomic Composition in Educational Equity and Effectiveness. *International journal of science and mathematical education*, *8*(3), 429–452.

Ministerial Council on Education, Employment, Training and Youth, & Affairs (MCEETYA). (1997). *National Report on Schooling in Australia 1997*. Melbourne.

Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA). (2008a). National assessment program: Literacy and numeracy. Year 3 numeracy test. Melbourne: Curriculum Corporation.

Monash University. (2007). *National numeracy review background paper* (Background Paper).

Perso, T. (2007). *"Back to Basics" or "Forward to Basics"?* Australian Mathematics Teacher , 63(3), 6-11.

Perso, T. (2009). Cracking the NAPLAN Code: Numeracy and Literacy Demands. *Australian primary mathematics classroom*, *14*(3), 14–18.

Phillips, T. (2009). Analysing quantitative data. In *Social research methods* (2nd ed., pp. 217–242). South Melbourne, Vic.: Oxford University Press.

Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading & writing quarterly*, *23*(2), 139–159.

Sullivan, P. (2011). *Teaching Mathematics: Using research-informed strategies* (p. 80). Australian Council for Educational Research.

Vale, C., & Bartholomew, H. (2008). Gender and mathematics: theoretical frameworks and findings. In *Research in Mathematics Education in Australasia 2004-2007* (Vol. 12). Rotterdam / Taipei: Sense Publishers.