VICTORIAN INDIGNENOUS CHILDREN'S RESPONSES TO MATHEMATICS NAPLAN ITEMS

PATRICIA MORLEY

Monash University

patricia.morley@monash.edu

It has often been reported that children of Australian Indigenous background do not perform as well as a group as the whole population. This paper addresses the question of whether Victorian Indigenous children have different patterns of responses from the general population. The analysis compares the responses on each item for Indigenous children with the responses for non-Indigenous children both directly and for those who achieved the same NAPLAN scores for the 2008 NAPLAN numeracy assessment for years 5, 7, and 9. The results indicate trends in the characteristics of items which successes or challenges for Indigenous children.

Introduction

It has often been reported that Indigenous Australian children do not perform as well as a group as their peers (Doig, 2001). Large-scale assessments such as the National Assessment Plan - Literacy And Numeracy (NAPLAN), while they may not provide the detailed understanding that smaller studies can, complement the findings of smaller studies by offering a wider perspective at a population level. This paper addresses the question of whether children of Indigenous background in Victoria, Australia, have different patterns of mathematical responses from the general population, which may have implications for teaching approaches.

The literature to date has been principally concerned with the socio-economic and environmental factors that contribute to the relatively low performance of Indigenous children. This analysis seeks to extend our understanding of the reasons underlying the lower performance by identifying differences in facility related to the topics and presentation of the mathematical NAPLAN items. This research shows that for items in the Space strand of the curriculum, Indigenous children are performing close to grade level. This paper seeks to provide a more detailed view of the NAPLAN results by examining the difference in facility of the items for the Indigenous children compared to the entire population. To ensure that the trends found are not due to differences in levels of mathematical understanding, the responses of Indigenous children and non-Indigenous children who achieved the same NAPLAN scores for the 2008 NAPLAN numeracy assessment are also compared. The data consistently show that Indigenous children show relatively strong performance on items in the Space strand, and have greater difficulty with items which are difficult for the general population.

Figure 1 below presents a visual representation of the relative performance of Indigenous and non-Indigenous children. The NAPLAN assessment score is calculated as the number of items correct that each child obtained. The box-plots, scaled in width to visually indicate relative population sizes, represent distribution the 2008 mathematical assessment scores for years 5, 7 and 9. The population size represented by each box is indicated at the bottom of each box-plot. Because the assessments are not directly comparable, either in the number of items or the relative difficulty, the score is represented as the number of standard deviations from a mean of zero.

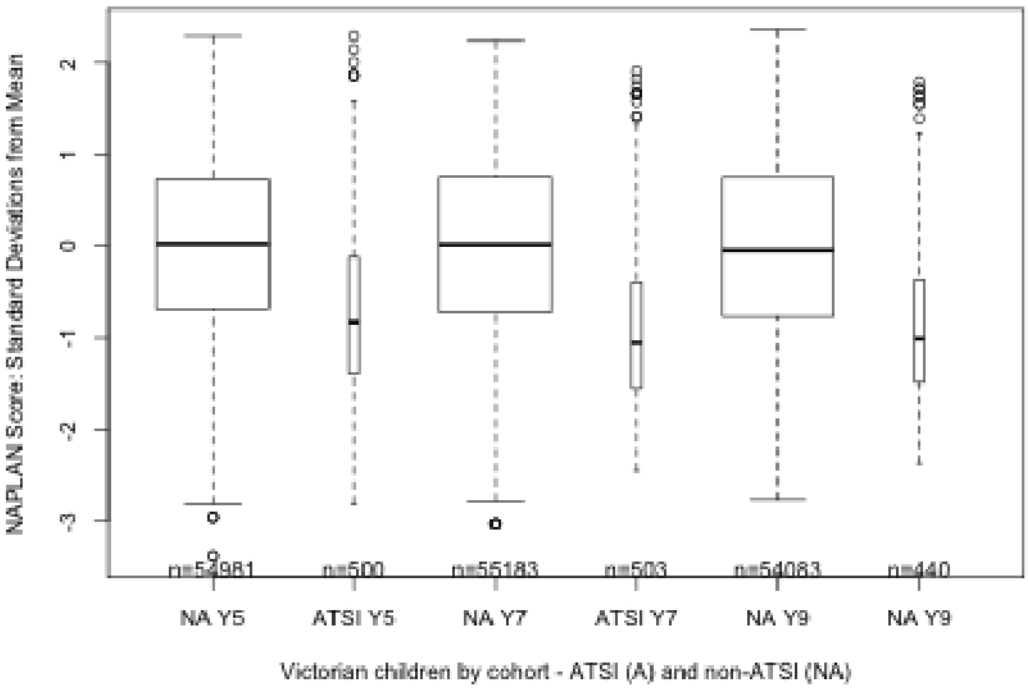

Figure 1: Distribution of 2008 NAPLAN scores.

Figure 1 shows the extent to which Indigenous children are performing less well as a group in comparison to the non-Indigenous population. For each year level represented, three quarters of Indigenous children scored in the same range as the lower half of the non-Indigenous population. The upper quartile included children who achieved very high scores, including one child who achieved a perfect score at the year 5 level. The second quartile of children are within the inter-quartile range of the main population, and the third and fourth quartiles are in the same range as the lower quartile of the main population, showing that although Indigenous children as a group perform less well than the general population, a significant proportion of the Indigenous children are performing within the main range of the general population, particularly in the earlier years.

Literature review

Hart (1980) demonstrates the use of results of assessments to obtain insights into children's understanding of mathematics. More recently, international large-scale assessments such as PISA and TIMMS have been used to measure and monitor

academic outcomes, allowing researchers to gain insights from analysis of large scale data sets. One example is Thompson, de Bortoli, Nicholas, Hillman, and Buckley (2010) who used the 2009 PISA results to make various inferences about successes and challenges within mathematics education in Australia.

Some of the issues that affect the learning of Indigenous students that have been addressed include issues such as remoteness, attendance, and language (Jorgensen & Sullivan, 2010), mismatches between expectations and pedagogies (Cooper, Baturo, & Warren, 2005), and aspects such as learning style (Reeve, 2010).

Methodology

The data used in this analysis comprise Victorian children's responses to all multiple-choice items on the Australian National Assessment Plan—Literacy And Numeracy (NAPLAN) 2008 Numeracy assessment for years 5, 7, and 9[1]. Because the capability of large-scale assessments to inform the mathematical education community is not fully known, an exploratory analysis allows us to detect patterns in the data, thus providing a base of knowledge upon which to form and confirm hypothesis in further research (Tukey, 1980). Specifically, the exploration focuses on identifying patterns in items for which the Indigenous population responds in a different way from the non-Indigenous population. The analysis presents an initial overview, using boxplots showing relative distributions of performance, scatter plots to show trends in item facility, and Lowess curves to show the trends in differences in facility between the two groups for different curriculum strands where there are sufficient data points.

As a way of getting better insights into the differences between groups, the facilities of items for the Indigenous and non-Indigenous children who attained the same score on the NAPLAN assessment were compared. The Welch t-test is a variation of the non-parametric Student t-test which is appropriate when the variances of the populations differ, especially when the population size is unequal, as in this case the Student t-test is less reliable (Ruxton, 2006). Both of these conditions apply in this case, with the Indigenous population being 500, 503, and 440 for the year 5, 7, and 9 cohorts respectively. The variance within each population of the total NAPLAN score was calculated for the Indigenous and non-Indigenous population, and were found to be 42 and 49 for the Year 5 cohort, 109 and 147 for year 7, and 155 and 114 for the year 9 cohort. The Welch t-test was applied for each score level for which there were more than 20 Indigenous students. A total of 1288 tests were carried out, and the 150 tests for which a p-value of less than 0.05 was obtained were noted. Since approximately 70 of these could be expected purely by chance, only items for which the Welch test was positive for three or more score levels have been considered.

Analysis

The focus of the analysis is to identify trends in the characteristics of items that Indigenous children find relatively challenging or have relative success with. Figure 2 shows, for each item, the proportion of the 55,481 Year 5 children who answered the item correctly, or facility, on the 2008 NAPLAN numeracy assessment grouped

[1] NAPLAN data are used and reproduced with permission of the Victorian Curriculum and Assessment Authority (VCAA). Analysis and findings using that data are not connected with or endorsed by the VCAA.

according to Indigenous background. The term facility is expressed here in terms of a proportion, rather than as a percentage, but is otherwise identical to the usage as introduced by Hart (1980).

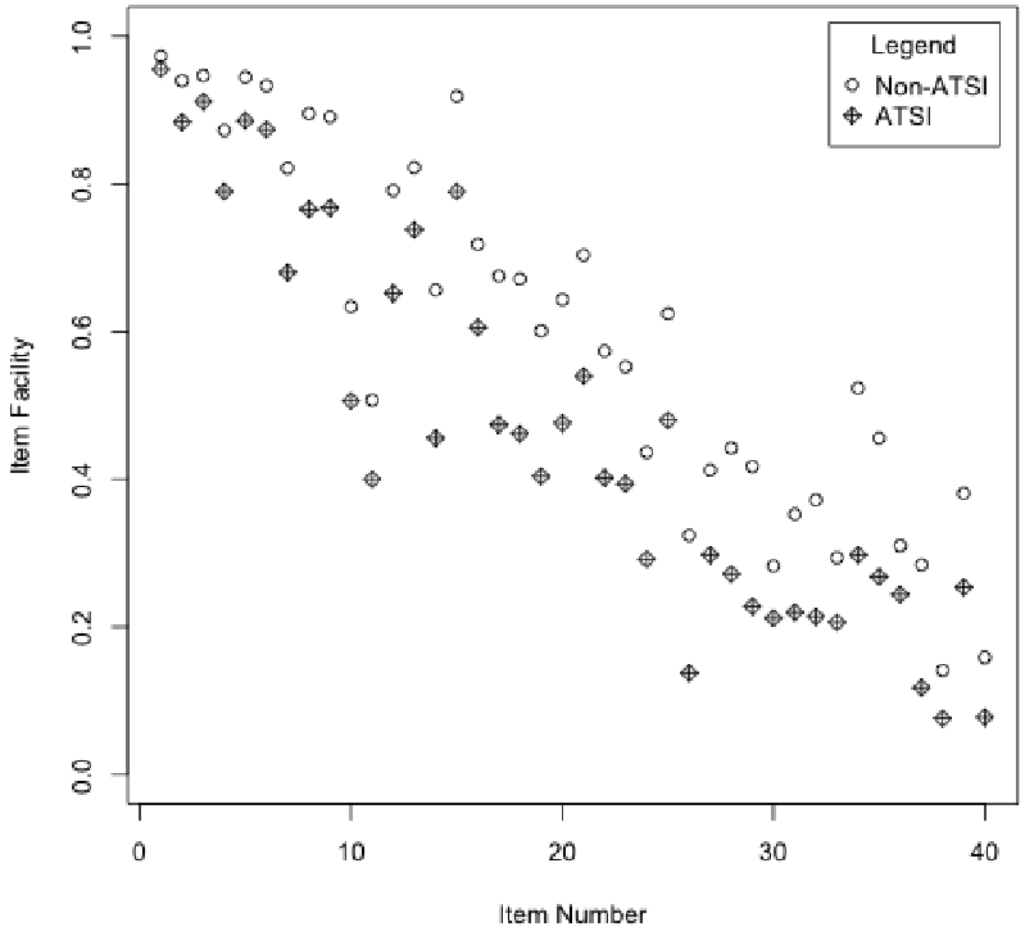

Figure 2. Proportion of Year 5 children who answered item correctly, grouped according to Indigenous background.

Figure 2 shows that the items are progressively more difficult to complete successfully throughout the assessment. The general trend appears to be the same for both groups, with children of both groups performing well on items at the beginning of the assessment, and finding the later items more difficult. The number of missed items is small—around 1%, for both Indigenous and non-Indigenous groups—even for items at the end of the assessment. The facility of each item is lower for Indigenous children than it is for non-Indigenous children, although some items have a smaller difference than others. These observations also hold for years 7 and 9 on the 2008 assessments.

Examining the magnitude of the difference in item facility provides more information. At year 5, the difference in facility between Indigenous and non-Indigenous children is greatest for the items of moderate facility. The items which are relatively easy, with a facility of greater than 0.9, are answered very well by Indigenous children, showing little difference between the two groups. Items of low facility only demonstrate that the items are very difficult for both groups, as the scope for differences

in facility becomes smaller. There is no obvious differentiation between curriculum strands. Some items pose greater difficulty for Indigenous students relative to non-Indigenous children, and the analysis of these may be of interest in future research.

Figure 3 shows the difference in item facility between the Indigenous and non-Indigenous populations on the year 7 NAPLAN assessment for each item. The three curriculum strands Measurement, Number, and Space are depicted respectively as circles, crosses and triangles. A locally-weighted scatterplot smoothing, or Lowess curve, has been drawn for each curriculum strand.

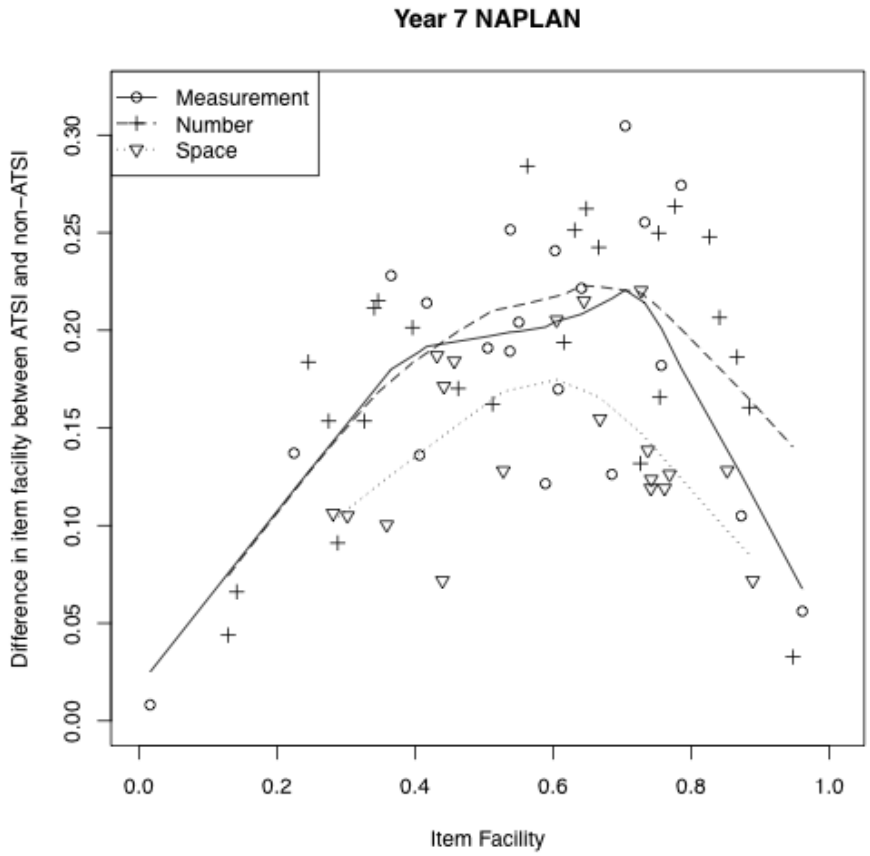

Figure 3. Difference in item facility between the Indigenous and non-Indigenous populations in the year 7 results.

Figure 3 shows that at Year 7, the differences of facilities between Indigenous and non-Indigenous children follow a similar pattern to Year 5. While Number and Measurement have similar Lowess curves, it is apparent that the Space strand of the curriculum has a smaller difference in facility between Indigenous and non-Indigenous groups. Lowess curves reflect all of the data, and the pointy peak of the Measurement curve indicates something unusual in the data. It draws our attention to the item which had the greatest difference in facility (0.3) between the two groups. Item 36 was answered correctly by 70% of non-Indigenous children, but only 40% of Indigenous children. This particular item assessed the ability to calculate the average of a number of items listed in a table. The most common error for both groups was to choose the sum of the numbers, rather than the average. In this case, the item could be solved by inspection, by eliminating all options other than the correct one using the knowledge that the average is in some way

representative of a set of numbers. Possible sources of difficulty include the potential unfamiliarity of the item context (walking a dog for a certain amount of time); unfamiliarity with the term "average"; or the number of steps involved in the task. Graphs such as this one highlight items that are particularly problematic to students, providing teachers with a focus for planning teaching strategies. This particular item would suggest that a problem-solving strategy that includes steps for inspecting an item before calculating, and afterwards checking if the result is reasonable, may be useful.

By year 9, some, but not all, of the items on the Space strand of the curriculum continue to be a strength of Indigenous children relative to other strands. The greatest difference is found in the Algebra strand, particularly in items of moderate difficulty. It is also apparent that some algebra items are more difficult for Indigenous children than others, independent of the facility of the item. It remains as a future research project to investigate those items to determine if there are identifiable factors that make these items relatively more difficult.

Analysis of items with achievement held constant

As a way of getting better insights into the differences between groups, the facilities of items for the Indigenous and non-Indigenous children who attained the same score on the NAPLAN assessment were compared. The items which were identified as having different facilities between the two groups on the Year 5 NAPLAN numeracy assessment are shown in Table 1. The skill is the description of the skill assessed by the item as given in the VCAA test answer booklet. The strand corresponds to one of the curriculum areas: Number (N); Space (S); Measurement (M); Chance and Data (D).

Table 1 shows that there are no items from the Measurement strand of the curriculum. The Space and Number strands of the curriculum contain some items, generally of high overall facility, that Indigenous children perform relatively well on, and items, generally of low overall facility that Indigenous children perform relatively less well on. These results indicate that the difficulty of the item has a greater impact on any difference between Indigenous and non-Indigenous children than the curriculum area of the item for children who are achieving the same NAPLAN score.

Table 1. Items on 2008 Year 5 NAPLAN Numeracy assessment for which Indigenous children and non-Indigenous children of the same NAPLAN score had different facilities.

Item	Strand	Multiple choice	Facility of item for Indigenous children	Overall facility	Skill
1	S	Yes	Higher	0.97	Identify symmetry in shapes
3	S	Yes	Higher	0.95	Compare the size of different angles
5	N	Yes	Higher	0.94	Complete number patterns based on simple criteria
6	S	Yes	Higher	0.93	Identify a 3D model given its individual components
13	N	No	Higher	0.82	Carry out simple money calculations
26	N	No	Lower	0.32	Recognise decimal numbers generated by dividing by 10
32	S	Yes	Lower	0.37	Identify and recognise properties of 2D shapes

38	N	No	Lower	0.14	Perform computations involving decimals
40	S	Yes	Lower	0.16	Recognise perspective in 2D representations of a 3D shape

The trend of Indigenous children doing well on the easier items and less well on the difficult items that was observed in the year 5 results is also evident in the year 7 data. Of the four items identified, only one was a multiple-choice item. Two items had an overall difference in facility that was less than chance and so the difference in facility can tell us little. Item 59 also had a low facility of 0.34. Item 3, with a facility of 0.83, was from the number strand, and asked children to select another way of writing 6^2 from the options 6×2; 6×6; $6 + 6$ and $2 \times 2 \times 2 \times 2 \times 2 \times 2$. Of the 503 Indigenous children in this year level, 125 chose the first option, 6×2. The correct answer, 6×6, was chosen by 291 children; only 12 chose $6 + 6$, and 67 chose the final option of $2 \times 2 \times 2 \times 2 \times 2 \times 2$. The remaining 8 children gave no legible response.

In the Year 9 data, the Algebra strand accounts for 2 of the 4 items identified as more difficult for Indigenous children. Overall, one of the most consistent findings is that Indigenous children score highly on high facility items, such as those where 90% or more of the population answer correctly. The implication of this finding is that children would benefit from increased exposure to more challenging mathematical material.

One of the assumptions made in measuring achievement by scores on multiple-choice tests is that a higher score reflects greater knowledge. As Sadler (1998) points out, albeit in the science domain rather than mathematics, this is not necessarily the case for difficult questions, where the performance dips from the expected performance level achieved by random guessing as the student gains an incomplete understanding of the topic being assessed, and is more likely to choose a distractor than the correct answer. This is an important issue, as for these very difficult items, the achievement of a higher score does not match with the goal of increased understanding of the topic unless the individual has achieved sufficient understanding to be able to answer correctly, making reliance on scores alone problematic, especially for low facility items. The items that fall outside of these patterns are also of interest, because these are the items that reveal opportunities to enhance teaching practices. For example, item 3 from the year 7 assessment, where children were asked to choose the option corresponding to 6^2, was answered correctly by 83% of children generally, but stood out as an item of difficulty for Indigenous children even when compared to children who scored at the same level. Since powers are an important component of algebra, this item is an early sign of the difficulties that Algebra poses for Indigenous children in year 9.

The implication drawn from this item, and the Year 7 item on calculating averages identified earlier, is that drawing the attention of children to the distinct use of language in mathematics, may be of benefit.

Conclusion

The exploratory analysis of children's responses to the 2008 NAPLAN numeracy assessment for years 5, 7, and 9 described in this paper confirms and extends previous findings in the research literature. The analysis demonstrates that there is wide variation in the individual achievement, and that Indigenous children perform well on items of high facility, and less well on items of low facility, suggesting that the children may be

more familiar with the simpler items. Indigenous children in year 7 have a relative advantage compared to their peers of similar achievement levels in the Space strand of the curriculum, but this advantage lessens for year 9, and that at the year 9 level, the Algebra strand is relatively difficult for Indigenous children, while the Number and Measurement strands are relatively difficult for all assessments. The implication for teaching is that a detailed analysis of results in large-scale assessments may provide insights that may be incorporated into teaching strategies.

References

Cooper, T., Baturo, A., & Warren, E. (2005). Indigenous and non-indigenous teaching relationships in three remote mathematics classrooms in remote Queensland. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 29th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 265–272). Melbourne: PME.

Doig, B. (2001). *Summing up: Australian numeracy performances, practices, programs and possibilities* (Literacy and numeracy research literature reviews series). Camberwell, Vic: ACER.

Hart, K. M. (1980). *Children's understanding of mathematics: 11–16.* London: John Murray.

Jorgensen, R., & Sullivan, P. (2010). Scholastic heritage and success in school mathematics: Implications for remote Aboriginal learners. In I. Snyder & J. Nieuwenhuysen (Eds.), *Closing the gap in education? Improving outcomes in southern world societies* (pp. 23–36). Clayton: Monash University Publishing.

Reeve, R. A. (2010). Using mental representations of space when words are unavailable: Studies of enumeration and arithmetic in Indigenous Australia. In *Teaching mathematics? Make it count: What research tells us about effective teaching and learning of mathematics* (pp. 62–66). Melbourne: ACER.

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology*, *17*(4), 688–690.

Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, *35*(3), 265–296.

Thompson, S., De Bortoli, L., Nicholas, M., Hillman, K., & Buckley, S. (2010). *Challenges for Australian education: Results from PISA 2009*. Melbourne: Australian Council of Educational Research.

Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, *34*(1), 23–25.