Do Students Solve Graphic Tasks with Spatial Demands Differently in Digital Form?

Tom Lowrie
University of Canberra
<tlowrie@gmail.com>

Ajay Ramful
University of Canberra
<a.ramful@mieonline.org>

Tracy Logan
University of Canberra
<tracyloga@gmail.com>

Siew Yin Ho
University of Canberra
<siewyin_ho@yahoo.com.sg>

This study compares Singaporean Grade 6 students' performance and strategy preference on two graphic-rich mathematics tasks, presented via pencil-and-paper and iPad modes. There were statistically significant differences between students' performances on the two tasks, one in favour of the paper mode and the other in favour of the iPad. Students who possessed higher spatial ability were more likely to solve the tasks correctly. The implications of the study are timely given the fact that high-stakes tests are likely to be presented in a digital form in coming years.

The recent announcement that the 2016 National Assessment Program – Literacy and Numeracy (NAPLAN) will be implemented in a digital online environment, raises questions about accessibility for all students as well as test-design features. It calls for renewed interest in the debate about whether students perform differently on pencil-and-paper and computer-based tests. Given the fact that international assessment jurisdictions are moving toward different modes of representing mathematical information (Organisation for Economic Co-operation and Development [OECD], 2013), it is important to pose questions about how students process information on pencil-and-paper and digital modes.

Visuospatial reasoning, that is, thinking with visual and spatial elements concurrently, is critical for success when students encounter graphic-rich mathematics tasks as highlighted by recent research (Lowrie & Diezmann, 2007; Lowrie & Logan, 2007). This form of reasoning involves the use of "internal [encoding] or external [decoding] visual or spatial representations from visual imagery to diagrammatic reasoning" (Shah & Miyake, 2005, p. xi). In order to decode a graphic, an individual must contend with multiple sources of information which may include text, keys or legends, axes, and labels (Kosslyn, 2006), as well as perceptual elements of retinal variables (e.g., depth of shading and pattern) (Bertin, 1967/1983). In order to encode a graphic, an individual typically draws pictures, diagrams or represents and manipulates images mentally. This study investigates how students process graphic tasks with spatial demands on pencil-and-paper and iPad.

Literature Review

Research investigating the difference between pencil-and-paper and computer-based tests (i.e., test mode effect) has been somewhat inconsistent in their findings. Although some studies reported differences in students' performance (Bennett et al., 2008; McDonald, 2002), a comprehensive meta-analysis involving 44 independent experiments revealed no test mode effect (Wang, Jiao, Young, Brooks, & Olson, 2007). The meta-

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. 429–436. Sydney: MERGA.

analysis criteria was based on psychometric measures that examined change in terms of total mean scores, time taken to complete tests and students' familiarity in using computers. In a similar vein, a study by Threlfall, Pool, Homer and Swinnerton (2007) investigated the different strategies students utilised when test items were translated from pencil-and-paper to a computer-based mode. They found that there were minimal differences on overall test performance between the two modes. However, for some items, students' performance differed significantly between the two modes. Following Threlfall et al.'s (2007, p. 346) suggestion to analyse test items "on a question by question basis", this paper considers two assessment items which focussed on students' processing of spatial concepts as described in the Method section. Besides the test mode effect, the display of information may also influence the ways in which students decode information. For instance, Hegarty, Canham and Fabrikant (2010) demonstrated that salience (i.e., the noticeable element in the display of information) has a large effect on performance in the interpretation of graphic tasks. In fact, they argued that the display design interacts with knowledge to influence the comprehension of visual displays.

Another important element in this study is spatial visualization which is defined as "the ability to mentally manipulate, rotate, or twist, or invert a pictorially presented stimulus object" (McGee, 1979, p.893). In spatial visualization, we use imagery to reason about an object when it is transformed. In simple terms, a mental image is "a mental representation of a mathematical concept or property containing information based on pictorial, graphical or diagrammatic elements" (Gutiérrez, 1996, p.5).

Research Questions

The current research report focuses on two items (Symmetry and Street map tasks, see Appendix) from the Mathematics Processing Instrument (MPI) designed and developed by Lowrie and colleagues (Lowrie, 2013). The relatively low students' performance on these two graphic items motivated the present study. Specifically, these two items involve considerable spatial demands, i.e., requiring the manipulation of objects, their positions and orientations in the mind's eye. This study addresses the following research questions:

1. Is there a relationship between spatial visualization ability and performance on graphic tasks with spatial demands?
2. How does the performance of students compare when solving graphic tasks with spatial demands on iPad versus pencil-and-paper?
3. Is there a difference in strategy use on the iPad and pencil-and-paper format of the two items?

Method

Participants

807 Grade 6 students (aged 11-12) from 8 Singapore schools (six government and two government-aided) took part in the study. There were 392 boys and 415 girls in the sample.

Instrument 1: Mathematics Processing Instrument (MPI)

As highlighted earlier, this study focused on two graphic items from the MPI (see Appendix), namely the Symmetry task (Part A, Item 9) and the Street Map task (Part B, Item 23). In its complete form, the MPI comprised 24 mathematics tasks (graphic and nongraphic), with corresponding processing instrument questions that allowed students to identify the problem-solving strategies they employed to solve the respective tasks. The MPI consists of two parts: Part A (Items 1-12) and Part B (Items 13-24).

The MPI was administered in Singapore schools in March-April 2013.Three staff from the research team attended schools during the normal curriculum class time. Students in each school were randomly divided into two groups. They were given 2 hours to complete the MPI. During the first hour, half the class answered one part of the MPI on iPad while the second cohort completed the tasks in a pencil-and-paper mode. Students working on the iPad were given working-out-paper for any calculations or representations. After a short break, the students answered the second part of the MPI for the next hour - via the other representational mode. Consequently, each student completed 12 items on the iPad and 12 items in the pencil-and-paper mode. The two items investigated for further study (Items 9 and 23) were among the items the students had most difficulty in solving.

Instrument 2: Paper Folding Test - Measurement of spatial visualization ability

We measured the spatial visualization ability of the students from the widely-used Paper Folding Test (Ekstrom, French, & Harman, 1976). In this instrument, students are required to visualise the folding action of a square sheet of paper. A hole is then punched in one part of the fold and students are to identify how the punched sheet would appear when fully reopened. An example of an item in this test is shown in Figure 1. The Paper Folding Test consists of 20 items. A correct item is given a score of 1 mark. The total score is calculated as follows: Number of items marked correctly minus one-fifth the number marked incorrectly. The minimum score is -4 and the maximum score is 20. The Paper Folding Test was administered on a different day from the MPI. Figure 1 gives a sample item from the test.

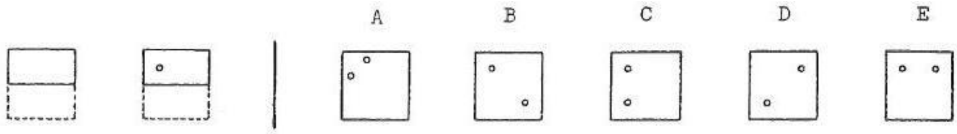


Figure 1. Paper Folding Test[1]

Results

Descriptive Analysis

A correct item was scored as "1" and an incorrect one as "0". The mean and standard deviation for the Symmetry task and the Street map task were M = .57 (SD =.496) and M = .47 (SD =.499), respectively. A test of proportions was used to determine if there were significant differences in performance. Girls (M =.60, SD =.49) did significantly better than boys (M =.53 , SD =.50) on the Symmetry task, χ^2 (1) = 4.18, p < .05. On the other hand, on the Street map task, boys (M =.52 , SD =.50) did significantly better than girls (M =.42 , SD =.49), χ^2 (1) = 8.44, p < .01.

[1] *The Paper Folding* test is reproduced with license and permission of Educational Testing Service, New Jersey, USA.

Research Question 1: Is there a relationship between spatial visualization ability and performance on graphic tasks with spatial demands?

The mean and standard deviation on the Paper Folding Test were 9.98 and 4.17 respectively. The Kolmogorov-Smirnov test (D(804) = .06, p = 0.00) suggested that the distribution of marks on the Paper Folding Test was significantly non-normal, as could be observed from the box-and-whisker plot.

We categorized the performance of the students on the Paper Folding Test as low, high and average (Kozhevnikov, Hegarty & Mayer, 1999) as follows: (i) Low: bottom 25% of the distribution, (ii) High: top 25% of the distribution and, (iii) Medium: the middle 50% of the distribution. The intent of the categorization was to determine if there was a difference in spatial visualization ability (as measured by the Paper Folding Test) and performance on the two spatial visualization items. Figure 2 illustrates that students with different spatial visualization ability (i.e., Low, Medium and High) performed differently on the Symmetry and Street map tasks.

We also computed the correlation between the Paper Folding Test scores and the scores on the two items. Since the performance on the Symmetry and Street map tasks were graded dichotomously as either 0 (incorrect) and 1(correct) and performance on the Paper Folding Test involved continuous scores, biseral correlation coefficients (Field, 2009) were computed. The performance on the Paper Folding Test was significantly correlated to that of the Symmetry task ($r_b = .36, p < .01$) and Street map task $r_b = .31, p < .01$). Further, the performance on the Symmetry task was significantly related to that of Street Map Task, ($r_s = .166$ (Spearman's correlation coefficient), p = .000).

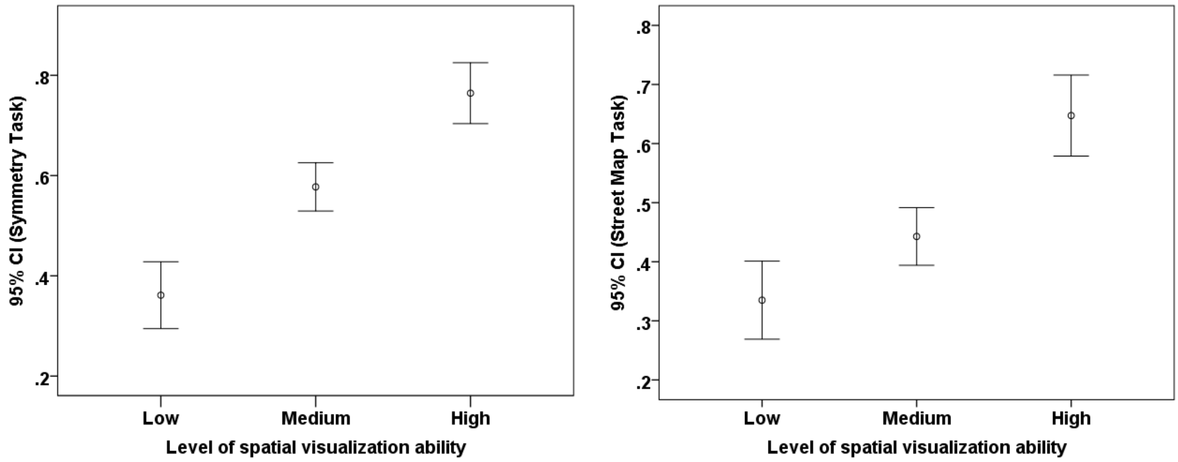


Figure 2. Plot of confidence interval for the two tasks for Low, Medium and High ability groups

Research Question 2: How does the performance of students compare when solving graphic tasks with spatial demands on iPad versus pencil-and-paper?

There were significant differences in proportion of successes between iPad and pencil-and-paper test modes for both tasks. For the Symmetry task, the performance on iPad (M = .61 and SE = .49) was higher than that on pencil-and-paper (M = .53 and SE = .50), χ^2 (1) = 5.22, p < .05. For the Street Map task, the performance on pencil-and-paper (M = .53 and SD = .50) was higher than that on iPad (M = .40 and SD = .49), χ^2 (1) = 15.23, p < .01.

We equally explored whether students with different levels of spatial visualization ability, as measured by the Paper Folding Test, performed differently on iPad and pencil-and-paper on the two items. There were no significant differences in performance between the iPad and pencil-and-paper media for the Symmetry task within each of the three levels of spatial visualization ability (Table 1). However, for the Street map task, differences were significant for students with low and medium spatial visualization ability (Table 2).

Table 1
Performance on iPad and pencil-and-paper on the Symmetry Task

Spatial visualization ability	Number of students	% correct on iPad	% correct on pencil-and-paper	χ^2 (1)	p-value
Low	202	41.2	30.7	2.27	0.13
Medium	407	60.1	54.5	0.97	0.33
High	194	82.0	70.3	3.07	0.08

Table 2
Performance on iPad and pencil-and-paper on the Street Map Task

Spatial visualization ability	Number of students	% correct on iPad	% correct on pencil-and-paper	χ^2 (1)	p-value
Low	200	25.7	40.2	5.51	0.02
Medium	402	36.6	50.0	8.04	0.05
High	193	59.3	69.0	1.91	0.17

Research Question 3: Is there a difference in strategy use on the iPad and pencil-and-paper format of the two items?

By asking students to describe their solution strategies, the MPI aimed at capturing the ways in which the iPad and pencil-and-paper test mode prompted students to use particular strategies on the two spatial items. In the Symmetry task, students were given three options, denoted S1, S2 and S3 (see Appendix). In the Street map task, students were given five options S1-S5. Table 3 shows the distribution of the strategies used and the percentage success for each strategy.

The most common strategy for the Symmetry task on both iPad and pencil-and-paper was S2: "I visualized/imagined folding the paper along the dotted line." However, there was more success on iPad than on pencil-and-paper in using strategy S2. We wonder if students were prompted to visualize by the iPad. For the Street map task, S3 (I visualized/imagined where the compass indicating the North direction will be on the graphic) was the most common strategy on iPad whereas on pencil-and-paper it was S1 (I solved the task by drawing a compass indicating the North direction on the graphic). Almost two thirds of students from the pencil-and-paper mode, who chose S1, correctly solved the task. By contrast only 43% of students who employed this strategy on the iPad produced a correct response. This shows that the medium in which a mathematical assessment (involving a spatial task) takes place may have an impact on the strategy used by students, and indeed their likelihood of success.

Table 3
Strategies used and percentage success

	S1		S2		S3		S4		S5	
	% chosen	% correct	% chosen	% correct	% chosen	% correct	% chosen	% correct	% chosen	% correct
Sym iPad	27.2	70.5	69.4	55.9	3.4	78.7	--	--	--	--
Sym Paper	16.7	69.7	78.5	48.7	4.8	36.8	--	--	--	--
Map iPad	28.9	43.0	12.4	51.0	33.2	29.0	18.2	50.0	7.3	10.3
Map Paper	42.7	63.6	11.2	41.3	28.6	35.6	9.5	66.7	8.0	21.2

Discussion and Conclusion

It is widely acknowledged that spatial ability is a predictor of performance in mathematics (Wai, Lubinski, & Benbow, 2009), especially in situations where students are required to decode graphic information (Lowrie & Logan, 2007). For the two tasks in the study, students with high spatial visualization ability performed at a much higher level than those who possessed medium or low spatial visualization ability. We investigated how the test mode (i.e., iPad versus pencil-and-paper) influenced students' performance and strategies in solving the two graphic tasks with spatial demand. The results of the study reveal significant differences in performance across the two test modes. There were equally differences in strategy use across the two graphic items on iPad and pencil-and-paper.

For the Symmetry task, there were performance differences in favour of those students who used iPads. We hypothesise that the iPad encouraged students to mentally reflect the given object (RZ) in their mind's eye. Those students who completed this task in a pencil-and-paper mode were also more likely to use this approach, but surprisingly with less success. Given that the students could not physically fold the object across the line of symmetry on iPad (as they could have done on paper), we speculate that the digital mode prompted them to mentally reflect the given object (RZ).

By contrast, those students who completed the Street Map task in a pencil-and-paper form scored higher than those students who solved it on iPad. This task required students to superimpose and rotate a visual compass from its prototypical North position on the given graphic. The iPad students used a variety of strategies to solve the task, with the highest proportion using imagery to evoke a mental representation of a compass indicating the North direction. The pencil-and-paper mode tended to encourage students to draw a compass on the diagram. Such an encoding strategy produced a high proportion of correct responses. Despite having working-out-paper, iPad users were less likely to draw, and success rates were lower when they chose this strategy (i.e., S1).

The results of the study indicate three educational implications. Firstly, the test mode (i.e., pencil-and-paper versus iPad) appears to be influential in students' mathematics performance on graphic tasks with spatial demands. We also observed that strategy selection varied across the two test modes. Given that NAPLAN will be online, and in a digital form by 2016, it may be problematic to compare digital-based performance to earlier paper-based NAPLAN results. It may be the case that a change in test mode means we are measuring different skills. Secondly, visuospatial reasoning skills will need to be taught more explicitly - and from an Australian perspective - in a new national curriculum that does not draw explicit attention to any such skills or abilities. This suggestion is motivated by the differences in strategy use that were observed in the students' responses to the two spatial tasks. Thirdly, students were less likely to draw diagrams to monitor their thinking on iPads.

As digital forms of representation become commonplace, it is important to ascertain that students continue to use and value critical encoding skills (e.g., drawing diagrams) that support monitoring and sequential processing in problem solving. It is acknowledged that the current study used only two items to investigate students' strategies in relation to graphic tasks with spatial demands on pencil-and-paper and iPad. Further, only one measure of spatial visualization ability was used in the study. The findings, however, are illuminating, especially at a time where not much is known about students' performance on the two test modes. This is particularly the case with investigation of graphic-rich tasks which are becoming more prevalent in numeracy tests (Lowrie & Diezmann, 2009). Future research could look at the ways in which students solve mathematical tasks on iPad in terms of variables such as spatial ability, task types and gender.

References

Bennett, R.E., Braswell, J., Oranje, A., Sandene, B., Kaplan, B., & Yan, F. (2008). Does it matter if I take my mathematics test on computer? A second empirical study of mode effects in NAEP. *Journal of Technology, Learning, and Assessment, 6*(9). Retrieved 1 march 2014 from http://www.jtla.org

Bertin, J. (1983). *Semiology of graphics* (W.J. Berg, Trans.). Madison, WI: The University of Wisconsin Press. (Original work published 1967).

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton, NJ, Educational Testing Service.

Field, A. (2009). *Discovering statistics using SPSS*, 3rd Ed. London: Sage Publications.

Gutiérrez, A. (1996). Visualization in 3-Dimensional geometry: In search of a framework. In L. Puig & A. Gutiérrez (Eds.), *Proc. 20th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 3-19). Valencia, Spain: PME.

Hegarty, M., Canham, M., & Fabrikant, S. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphics interface task. *Journal of Experimental Psychology: Learning, Memory and Cognition, 36*(1), 37-53.

Kosslyn, S. M. (2006). *Graph design for the eye and mind.* New York: Oxford University Press.

Kozhevnikov, M., Hegarty, M., & Mayer, R. (1999). *Students' use of imagery in solving qualitative problems in Kinematics*. Washington DC: US Department of Education (ERIC Document Reproduction, 433 239).

Lowrie, T. (2013). Cross-country comparisons of student sense making: The development of a mathematics processing framework. In V. Steinle, L. Ball & C. Bardini (Eds.), *Mathematics education: Yesterday, today and tomorrow* (Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia, eBook, pp. 751-754). Melbourne, Australia: MERGA Inc.

Lowrie, T., & Diezmann, C. M. (2009). National Numeracy Tests: A graphic tells a thousand words. *Australian Journal of Education. 53*(2), 141-158.

Lowrie, T., & Diezmann, C. M. (2007). Solving graphics problems: Student performance in the junior grades. *The Journal of Educational Research, 100(6), 369-377.*

Lowrie, T., & Logan, T. (2007). Using spatial skills to interpret maps: Problem solving in realistic contexts. *Australian Primary Mathematics Classroom, 12*(4), 14-19.

McDonald, A. (2002). The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments. *Computers and Education*, *39*, 299-312.

McGee, M.G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, *86* (5), 889-918.

Organisation for Economic Co-operation and Development (OECD). (2013). PISA 2015: Draft mathematics framework. OECD publishing, retrieved 12 March 2014 from http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Mathematics%20Framework%20.pdf

Shah, P. & Miyake, A. (Eds.). (2005). *The Cambridge handbook of visuospatial thinking*. New York: Cambridge University Press.

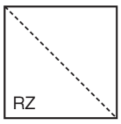
Threlfall, J., Pool, P., Homer, M., & Swinnerton, B. (2007). Implicit aspects of paper and pencil mathematics assessment that come to light through the use of the computer. *Educational Studies in Mathematics*, *66*(3), 335-348.

Wai, J., Lubinski, D., & Benbow, C. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative knowledge solidifies its importance. *Journal of Educational Psychology, 101,* 817-835.

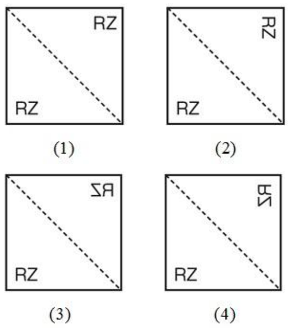
Wang, S., Jiao, H., Young, M.J., Brooks, T., & Olson, J. (2007). A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, *67*, 219-238. DOI: 10.1177/0013164406288166.

Appendix

<table>
<tr><th>Symmetry Task</th><th>Street Map Task</th></tr>
<tr><td>
Ron paints these letters on a piece of paper.



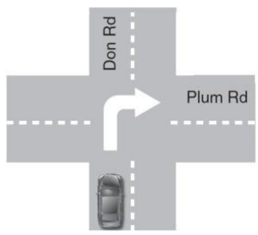
While the paint is still wet, he folds the paper along the dotted line.
When Ron unfolds the paper, what will it look like?



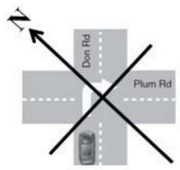
S1: I folded the paper along the dotted line.

S2: I visualized/imagined folding the paper along the dotted line.

S3: I did not use any of the above methods – I attempted the task in this way.
</td><td>
A car is travelling north-east along Don Road. The car is about to turn right into Plum Road.
In which direction will the car be travelling after it turns right?



S1: I solved the task by drawing a compass indicating the North direction on the graphic.



S2: I solved the task by turning the page so that the North direction is vertical to the edge of my table.

S3: I visualized/imagined where the compass indicating the North direction will be on the graphic.)

S4: I knew that a right hand turn is 90°, therefore 90° from north-east is South-East.

S5: I did not use any of the above methods – I attempted the task in this way.
</td></tr>
</table>