The Distinction Between Mathematics and Spatial Reasoning in Assessment: Do STEM Educators and Cognitive Psychologists Agree?

Danielle Harris
University of Canberra
<Danielle.harris@canberra.edu.au>

Tom Lowrie
University of Canberra
<Thomas.lowrie@canberra.edu.au>

Mathematics in Australia specifies spatial reasoning as a general capability within the curriculum. However, psychological research to date limits spatial assessment to psychometric tests leaving little room for a well-defined spatial curriculum. Although there are clear relationships between mathematics and spatial thinking, the independence in the measurement of the two constructs in research literature is rarely explored. In the present study, professionals in the fields of STEM Education and Cognitive Psychology evaluated mathematics and spatial assessment items. The results show evidence for a distinction between the two constructs in the content of the items, however with a caveat that thoughtful selection of assessment items is crucial to ensure independence in the measures.

Throughout the world there is growing advocacy for developing spatial thinking within school curriculums (Mulligan, 2015; National Research Council, 2016; Newcombe, 2017). Although some countries remain grounded in traditional computational mathematics, others, such as Australia are moving towards a wider view of mathematical competency. The definition of Numeracy in the Australian curriculum addresses the need to develop students' mathematical skills that can be applied to the real world (ACARA, n.d.). This includes specific reference to spatial reasoning. The focus on numeracy in the Australian agenda has resulted in a shift in assessment content (Logan & Lowrie, 2017). Although the nature of assessment is designed as a barometer for educational outcomes, there is little doubt that these outcomes influence classroom practice, thus driving the assessment-curriculum cycle (Doig, 2006).

Mathematics and Spatial Assessment

Within standardised mathematics assessment in Australia the balance between traditional mathematics and spatial content is changing (Lowrie & Diezmann, 2009). In mathematics problem solving, spatial processes are advantageous for assessment success (Lowrie, Logan & Ramful, 2016a). Likewise, for items that appear fundamentally spatial, there is evidence to suggest some numerical processing may be required (Maybury & Do, 2003). Mix and Cheng (2012) proposed that research needs to identify "psychological distinction" (p. 205) between mathematical and spatial processing in order to distinguish between the constructs. There is scant literature on the classification of mathematics tasks by processing requirements, instead mathematical categories tend to be defined by content strands within curriculums (Mix & Cheng, 2012). While different countries support the teaching of different techniques and heuristics for assessment with varying degrees of success (Lowrie, Logan & Ramful, 2016b), spatial processes have been found to support the development of mathematical proficiency, particularly when encountering novel tasks (Lowrie & Kay, 2001). Although neither spatial reasoning nor mathematics can be thought of as unitary constructs (Mix & Cheng, 2012) the enduring relationship cited within the

2018. In Hunter, J., Perger, P., & Darragh, L. (Eds.). Making waves, opening spaces *(Proceedings of the 41st annual conference of the Mathematics Education Research Group of Australasia)* pp. 376-383. Auckland: MERGA.

literature begs the question, are they distinct proficiencies when it comes to academic assessment or do they exist on a continuum?

The nature of the components that comprise spatial reasoning have been explored in multiple ways but are often grounded within factor analytic studies based on multiple measures (Carroll, 1993). One theoretical distinction within spatial reasoning is the idea that spatial reasoning can be defined in terms of mental transformations of an object compared with transformations of one’s own viewpoint (Sorby, 1999). The differences between these two types of spatial processing may be the result of different task demands or in individual strategy preferences (Hegarty & Waller, 2005). Despite the categorisation of processing differences, the definitions are still bound within psychological measures. Spatial reasoning has a reputation for its lack of theoretical underpinnings or frameworks, often being restricted to narrow definitions aligned to the psychological tests used to measure the constructs (Hegarty & Waller, 2005). Despite the increasing acceptance that spatial reasoning is embedded within many aspects of numeracy, there is still limited research establishing how to incorporate traditional measures of spatial thinking into mathematics curriculums (Mulligan, 2015).

Therefore, the question remains, if space and number are inherently linked and there is a push towards spatial content within school curriculums and assessment, how do we separate traditional psychological measures of spatial thinking from practical curriculum-based assessment?

Aim of the Present Study

The relationship between mathematics and spatial reasoning has been explored extensively through correlational and longitudinal studies (see Mix & Cheng, 2012 for a review; Casey et al., 2015). Researchers are confident that the development of strong spatial skills equips students in STEM fields and mathematics in particular (Mulligan, 2015). Whether this relationship is due to shared underlying processes or problem-solving strategies is still under debate (Mix & Cheng, 2012). Nonetheless, research has rarely questioned the independence of the two constructs as output measures, even for geometry items that are inherently spatial (Clements & Battista, 1992). Validated measures in research carry an assumption of content validity as they have undergone rigorous testing and peer-review (Peter & Churchill, 1986). However, in the case of spatial reasoning and mathematics where there is theoretical overlap and the lines distinguishing the two skills are blurred, this assumption of content validity can no longer be guaranteed.

Peter and Churchill (1986) define two characteristics of valid measures, 1) the measures do not understate the intended constructs and 2) they do not assess extraneous characteristics. It would seem from their conclusions that there is a midpoint in which measures should sit and that the only evaluation of content validity at our disposal is subjective judgement. In research exploring the mathematics-spatial relationship this position may be hard to find (particularly with reference to point 2), yet researchers rely on the content validity of the separate measures in declaring construct independence. Therefore, the aim of the present study was to explore the content validity of mathematics and spatial measures through the ratings of discipline professionals in order to validate the separation of numeracy and spatial assessments. Secondary to this, we aim to examine the psychological distinction between the two constructs within assessment.

Method

Participants

A participation request was sent out to the authors' network within the fields of mathematics education, educational psychology, cognitive psychology, pre-service teacher education and school teachers. Eighty-four responses were collected. Thirteen participants were removed from the analysis as they completed less than 50% of the survey. This left a total of 71 participants from four countries, see Table 1 below for geographical breakdown. Based on the supplied information, participants were categorised as STEM Education professionals (N = 58) or Cognitive Psychology professionals (N = 13).

Table 1
Participant Geographical Demographics

Country	Australia	U.S.A.	U.K.	New Zealand	El Salvador	Saudi Arabia	Unknown	Total
N	51	14	2	1	1	1	1	71

Measure and Procedure

Participants completed an online survey via an anonymous Qualtrics link. Demographic information was collected and then participants were asked to rate 38 items on a continuum of purely mathematics (score = 0) to purely spatial (score = 100). Participants moved a horizontal slider to rate each item, there was no numerical input required however data was recorded as the numerical equivalent of slider placement. The items rated were from the Spatial Reasoning Instrument (SRI; Ramful, Lowrie & Logan, 2017) and a set of numeracy items developed for a larger project examining mathematics and spatial reasoning in primary school students. These items were designed to reflect NAPLAN items (ACARA, n.d.) and covered geometry and measurement and number and algebra content strands. The items were randomly presented, and no information was given as to whether they were from the numeracy or spatial assessment.

Results and Discussion

For all items except two, there were no differences (using t-tests) between STEM Education professionals and Cognitive Psychology professionals in the ratings of individual items (p >.05). Two spatial items requiring reflection across a diagonal line (see Figure 1 for an example) were rated significantly higher (i.e., closer to purely spatial) by Cognitive Psychologists (*Means* = 93.33 and 91.47) than STEM Educators (*Means* = 69.95 and 76.00), t(46.89[1]) = 23.39, p < .001, d = 1.28 and t(48.71[1]) = 3.38, p = .001, d = .97. Cognitive Psychologists on average rated these items as more spatial, while STEM Educators placed them closer to a mixture of the two constructs due to the mathematical conventions (i.e., reflection, diagonal) in the question.

Most respondents were Australian (71.8%; see Table 1) and as a result of unequal sample sizes in other countries of origin the sample was classified as Australian and non-Australian for comparison purposes. For all items except one there were no significant differences (based on t-tests) in ratings based on country of origin (p < .05). One numeracy

[1] Adjusted values used due to violation of Levene's test for Equality

item (see Figure 2) was rated lower on the scale (i.e., closer to purely mathematics) by Australian professionals (*M* = 18.38) than non-Australian (*M* = 38.31) respondents, $t(51) = 3.51$, $p = .001$, $d = 0.99$. This may be a result of the role of the number-line in spatial literature (Edmonds-Wathen, 2012). This item may be considered a vertical number-line.

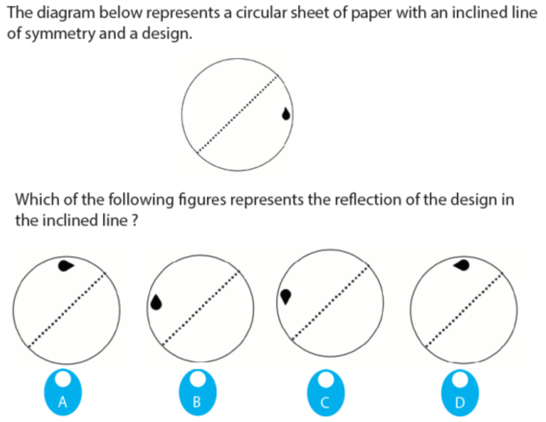

Figure 1. Spatial item (SRI; Ramful et al., 2017) rated differently by STEM Educators and Cognitive Psychologists

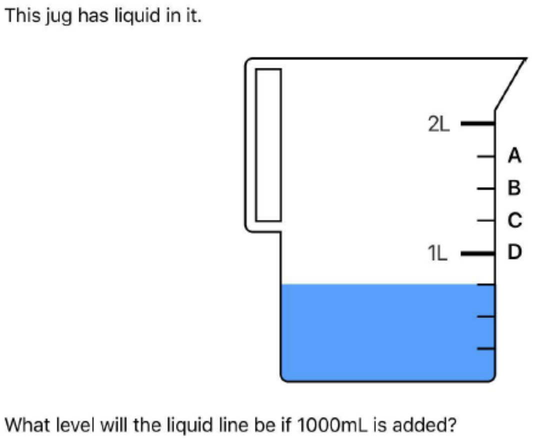

Figure 2. Numeracy item rated differently by Australian and non-Australian professionals

Factor Analysis

An exploratory factor analysis was performed because it was unclear how many factors may underlie the assessed items. A Principal Components extraction method was used and the scree plot of Eigenvalues was examined to establish the number of factors for consideration. The point at which the scree slope flattened suggested that there were three factors present in the dataset. A forced orthogonal rotation did not produce component correlations greater than .32 ($r = .017$), therefore no rotation was performed (Tabachnick & Fidell, 2001). The resulting factor loadings are presented in Table 2. Factor loadings of 0.32 or greater are commonly regarded as adequate for establishing the existence of a factor (Tabachnick & Fidell, 2001). Items that load greater than 0.32 on more than one factor are cross-loaded. Cross-loaded items are identified in Table 2, these items were attributed to the dominant factor when calculating factor scores (see Table 3). The three factors accounted for 60.32% of the variance in the model.

Table 2
Factor Loadings >0.32 Based on a Principal Components Analysis with a Promax Rotation of 38 Items from the Item Assessment Survey (N = 75)

Item Source	Description	1	2	3
SRI	Rotation of a figure	.90		
SRI	Rotation of a figure	.89		
Numeracy	Rotation of an object in Euclidean plane	.84		
SRI	Rotation of a figure	.82		
SRI	Directions on a map	.80		

SRI	Alternative perspective of a figure	.79		
SRI	Identifying the irregular shape of a net	.77		
SRI	Map rotation	.77		
Numeracy	Direction on a map	.75		.51
SRI	Identifying the irregular shape of a net	.75		
SRI	Rotation of a bike image	.75		
SRI	Rotation of a pentagon	.74		
SRI	Alternative view of a single block	.74		
SRI	Paper fold	.73		
SRI	Determining position from alternative perspective	.72		
SRI	Rotation of a dog image	.71		
SRI	Directions on a map	.71		
SRI	Alternative view of a set of blocks	.71		
SRI	Alternative view of a set of blocks	.69		
SRI	Identifying a slice of a 3-D shape.	.68		
SRI	Identifying the irregular shape of a net	.66		
SRI	Map navigation	.65		
Numeracy	Net of Cube	.63		
SRI	Lines of symmetry	.54		
SRI	Determining order from alternate perspective	.52		.49
SRI	Reflection across a diameter** - reflected object only	.43	.33	
SRI	Reflection across a diagonal** – reference object	.42		
SRI	Perspective taking	.36		.75
SRI	Perspective taking	.32		.71
Numeracy	Bar Chart		.83	
Numeracy	Fractions		.81	
Numeracy	Calculation on a vertical number line*		.76	
Numeracy	Number patterns		.76	
Numeracy	Number patterns		.70	
Numeracy	Number line		.67	.45
Numeracy	Bar Chart		.66	
Numeracy	Rotating fuel gauge		.57	
Numeracy	Rotation of an object in degrees		.40	

Note. * denotes difference by Nationality, ** denotes difference by profession

The tests were chosen to assess two separate constructs, numeracy and spatial reasoning. The exploratory factor analysis, however, revealed three underlying factors. The pattern that emerged was in line with the literature on spatial thinking that the constructs while related do not load on a single factor (Carroll, 1993; Sorby, 1999). Factor one is made up of 27 items that link to spatial transformations such as mental rotation and spatial visualisation (Sorby, 1999), for example rotation of a dog image or identifying a slice of a 3-D shape. These items included numeracy items that could be solved using spatial transformations alone (such as identifying one side of the net of a cube). The nine items in factor two are numeracy items that involve a degree of computation and understanding of numerical conventions. It is noteworthy that the cross loaded item is a reflection of a diagram with no reference object in the answer options. This is one of the items previously discussed that differed in rating based on assessor profession. Although intended as a spatial reasoning item, STEM Education professionals rated this item as more mathematical than Cognitive Psychologists due to the inclusion of numeracy concepts such as reflections. However, this cross-loaded item lacked a reference object in the answer options (compared with a similar reflection item with a reference object, see Table 2). Why these two items did not both cross-load on the two factors is unclear but may be influenced by the presence (or lack of) a reference object in the answer options. Factor three contained five items that cross-loaded on the other two factors (only two of these items were dominant) but could be characterised as requiring a viewpoint change, as opposed to a mental transformation. Of note is that not all perspective change items fell into this factor, this could be attributed to the ease of doing a mental transformation for some spatial items rather than a perspective change (Hegarty & Waller, 2005). The cross-loading of the number-line item and with this secondary spatial factor is interesting. Research has shown strong links between representations of number and space in Australian populations (Edmonds-Wathen, 2012), particularly when displayed on a horizontal left to right path. The vertical number-line item assessed did not produce the same relationship, however as previously noted was rated differently according to country of origin. It is noteworthy that the spatial factor aligned to the number-line item was the orientation factor, perhaps owing to the egocentric spatial features of location and direction embedded within the number-line (Edmonds-Wathen, 2012). Average factor scores are presented in Table 3.

Table 3
Average Rating Scores for Each Factor

	Spatial Transformations (27 items)	Numeracy (9 items)	Perspective Taking (2 items)
Rating Average (S.D.)	81.53 (12.07)	28.91 (19.11)	87.58 (15.06)

The loading of items onto three factors rather than two supports the notion that on a content validity level, neither spatial reasoning nor mathematics are unitary constructs or completely independent as literature suggests (Mix & Cheng, 2012). The cross-loading of particular items also support Mix and Cheng’s (2012) conclusion that numeracy and spatial reasoning assessments on the surface are not entirely distinct measures. The scale on which the items were measured had 0 as purely mathematical and 100 as purely spatial. The ratings presented in Table 3 lie towards the ends of the continuum as would be expected from separate measures of mathematics and spatial reasoning. However, the scores were not at the extreme ends of the continuum, suggesting that despite the surface delineation of the ratings of the two constructs, there are common components in mathematics and spatial reasoning assessment. Although separate factors, spatial transformations and perspective taking factors were both within the spatial end of the continuum, and significantly different in their ratings, $t(57) = .247$, $p = .02$, $d = .33$. Given there are only two items in the perspective taking factor, there is a limit to the conclusions which may be drawn. Regardless, it might be the case that items in this factor are further removed from mathematics due to the ego-centric transformations as opposed to the object-centric transformations more closely aligned to mathematical problem-solving (Kozhevnikov & Hegarty, 2001).

Conclusions and Recommendations

If Australian mathematics education research is to reflect the Australian curriculum it is important to explore the overlap between the psychological construct of spatial reasoning and the embedded curriculum content, as reflected by curriculum assessment.

The results of this study suggest that despite the apparent similarity of some numeracy (based on traditional Australian NAPLAN items) and spatial reasoning items (Ramful et al., 2017), there was a distinction between the two as rated by STEM Education and Cognitive Psychology professionals. As with Maybury and Do (2013), it appeared that some of the spatial items required a degree of mathematical content knowledge and some of the numeracy items were dependent on spatial processing. However, overall the two constructs were independent. The present work relied on ratings from professionals across STEM Education and Cognitive Psychology and therefore the conclusions drawn reflect on the content validity of the items and not the processing demands of the items per se. Given the reliance on content validity in defining measures for comparison (Peters & Churchill, 1986) and the close relationship of the two constructs (Mix & Cheng, 2012) it is important to separate the content of assessment items to draw conclusions about the nature of the mathematics-spatial relationship. Future research may address the connection between the processing of the two assessment types in addition to the content validity.

Although the sample was heavily Australian where the curriculum explicitly identifies the role of spatial reasoning in numeracy, there were few differences across countries where the role of spatial reasoning in the curriculum is less explicit (e.g., the United States; National Research Council, 2006). The present items assessed where drawn from the Australian curriculum despite a large volume of longitudinal work originating in the United States (e.g., Casey et al., 2015). There are interesting cross-cultural opportunities to explore the mathematics-spatial distinction across curriculums in future.

The results of the study shed light on some of the defining characteristics that distinguish mathematical from spatial assessment. Regardless of country of origin or industry specialisation, mathematical conventions appear to mark the delineation between mathematics and spatial reasoning. Within spatial reasoning there is a further divide between mental transformations and perspective changes. While curriculum achievement measures and psychometric tests of spatial reasoning continue to be used as the foundation of relational research, it is important to distinguish the line between the two. Validating their independence enables researchers to move towards a better understanding of the underlying relationships between mathematics and spatial thinking and how they can be developed to support one another.

References

ACARA (n.d.), General Capabilities, retrieved February 2, 2018 from Australian Curriculum website: https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/

Carroll, J.B. (1993). *Human cognitive abilities- A survey of factor-analytic studies.* New York: Cambridge University Press.

Casey, B. M., Pezaris, E., Fineman, B., Pollock, A., Demers, L. & Dearing, E. (2015). A longitudinal analysis of early spatial skills compared to arithmetic and verbal skills as predictors of fifth-grade girls' math reasoning. *Learning and Individual Differences, 40,* 90-100.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420-464). New York: Macmillan.

Doig, B. (2006). Large-scale mathematics assessment: Looking globally to act locally. Assessment in Education: Principles, Policy and Practice, 13, 265–288.

Edmonds-Wathen, C. (2012). Spatial Metaphors of the Number Line. In J. Dindyal, L. P. Cheng & S. F. Ng (Eds.), *Mathematics education: Expanding horizons* (Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia). Singapore: MERGA.

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah, & A. Miyake (Eds.), *Handbook of visuospatial thinking* (pp. 121-169). New York: Cambridge University Press.

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & cognition, 29*(5), 745-756.

Logan. T. & Lowrie, T. (2017). Gender perspectives on spatial tasks in a national assessment: a secondary data analysis. *Research in Mathematics Education*, *(19)*2, 199-216.

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., & Kay, R. (2001). Relationship between visual and nonvisual solution methods and difficulty in elementary mathematics. *The Journal of Educational Research, 94*(4), 248-255.

Lowrie, T., Logan, T., & Ramful, A. (2016a). Spatial Reasoning Influences Students' Performance on Mathematics Tasks. In White, B., Chinnappan, M. & Trenholm, S. (Eds.). *Opening up mathematics education research* (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia), pp. 407–414. Adelaide: MERGA.

Lowrie, T., Logan, T. & Ramful, A. (2016b). Cross cultural comparison of grade 6 students' performance and strategy use on graphic and non-graphic tasks. *Learning and Individual Differences, 52,* 97-108.

Maybury, M.T., & Do, N. (2003). Relationships between facets of working memory and performance on a curriculum-based mathematics test in children. *Educational and Child Psychology, 20*(3), 77-92.

Mix, K. S., & Cheng, Y.-L. (2012). The relation between space and math: developmental and educational implications. *Advances in child development and behaviour, 42,* 197-243.

Mulligan, J. (2015). Looking within and beyond the geometry curriculum: Connecting spatial reasoning to mathematics learning. *ZDM Mathematics Education, 47,* 511-517.

National Research Council. (2006). *Learning to think spatially: GIS as a support system in the K—12 curriculum.* Washington, DC: National Academy Press.

Newcombe, N. S. (2017). Harnessing Spatial Thinking to Support Stem Learning. OECD Education Working Paper, No. 161, OECD Publishing, Paris. doi: 10.1787/7d5dcae6-en

Peter, J.P. & Churchill, G.A. (1986). Relationships among research design choices and psychometric properties of rafing scales: A meta-analysis. *Journal of Marketing Research, 23,* 1-10.

Ramful, A., Lowrie T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. *Journal of Psychoeducational Assessment, 35*(7), 709-727.

Sorby, S. A. (1999). Developing 3D spatial visualization skills. *Engineering Design Graphics Journal, 63*(2), 21–32.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York, NY: Allyn and Bacon.