

Spatial Cognition and Science: The Role of Intrinsic and Extrinsic Spatial Skills from Seven to Eleven Years

Alex Hodgkiss (alex.hodgkiss.14@ucl.ac.uk)

Department of Psychology and Human Development, UCL Institute of Education, 25 Woburn Square
London WC1H 0AA, UK

Katie A. Gilligan (katie.gilligan.15@ucl.ac.uk)

Department of Psychology and Human Development, UCL Institute of Education, 25 Woburn Square
London WC1H 0AA, UK

Michael S. C. Thomas (m.thomas@bbk.ac.uk)

Department of Psychological Sciences, Birkbeck College, Malet Street, London, WC1E 7HX

Andrew K. Tolmie (a.tolmie@ucl.ac.uk)

Department of Psychology and Human Development, UCL Institute of Education, 25 Woburn Square
London WC1H 0AA, UK

Emily K. Farran (e.farran@ucl.ac.uk)

Department of Psychology and Human Development, UCL Institute of Education, 25 Woburn Square
London WC1H 0AA, UK

Abstract

The current study investigated the relationship between children's spatial ability and their scientific knowledge, skills and understanding. Children aged 7-11 years ($N=123$) completed a battery of five spatial tasks, based on a model of spatial ability in which skills fall along two dimensions: intrinsic-extrinsic; static-dynamic. Participants also answered science questions from standardised assessments, grouped into conceptual topic areas. Spatial scaling (extrinsic static spatial ability) and mental folding (intrinsic dynamic spatial ability) each emerged as predictors of total science scores, with mental folding accounting for more variance than spatial scaling. Mental folding predicted both physics and biology scores, whereas spatial scaling accounted for additional variance only in biology scores. The embedded figures task (intrinsic static spatial ability) predicted chemistry scores. The pattern was consistent across the age range. These findings provide novel evidence for the differential role of distinct aspects of spatial ability in relation to children's science performance.

Keywords: spatial ability; science; STEM; children.

Introduction

Large-scale longitudinal studies spanning the past 50 years provide convincing evidence that spatial ability in adolescence predicts later science, technology, engineering and mathematics (STEM) achievement; both in academic and career outcomes (Wai, Lubinski & Benbow, 2009). As well as often cited examples of scientific discoveries resulting from creative spatial thought, a growing body of research with adults and adolescents highlights a link between spatial ability and scientific reasoning (e.g., Kozhenikov & Thornton, 2006). However, with a few exceptions (e.g. Tracy, 1990), the relationship between

spatial ability and science learning in younger children has been largely neglected. This is important to address, given that early science learning involves specific areas of conceptual understanding, and because knowledge of how spatial ability and science relates in younger children has implications for early intervention. The focus of this study was therefore to investigate the relationship between a range of spatial skills and primary-school aged children's scientific knowledge, skills and understanding.

Spatial ability

Spatial ability, which relates to "the location of objects, their shapes, their relation to each other, and the paths they take as they move" (Newcombe, 2010, p30), has long been recognised as an ability at least partly independent of general intelligence, reasoning and verbal ability (Hegarty, 2014). As well as being apparently distinct from other cognitive abilities, spatial thought itself is generally conceptualised in a multidimensional fashion, consisting of several separate but correlated skills. Two categories of multidimensional models have emerged: ones based in the psychometric tradition (e.g. Carroll, 1993) and other more recent, theoretically driven models (e.g. Uttal et al., 2013).

Psychometric analyses of spatial ability have often resulted in inconsistent findings, with the number of identified factors ranging between two and twelve (Hoffler, 2010). Uttal et al. (2013) argue that some of the inconsistencies result from factor analysis models not being theory-driven. In contrast, Uttal et al.'s (2013) model is based on top-down, theory driven understanding of spatial skills, and draws upon developments in cognitive neuroscience. They propose that spatial skills can be

categorised along two dimensions: static-dynamic and intrinsic-extrinsic. Intrinsic spatial abilities and extrinsic spatial abilities broadly map onto a within-object and between-object classification, respectively. Intrinsic/extrinsic skills can be further categorised as either static or dynamic abilities; dynamic abilities include transformation or movement, whilst static skills do not.

Intrinsic-static skills involve the processing of objects or shapes without further transformation or movement of parts of the object or shape. Tasks that measure this skill often require this processing to occur amidst distracting background information. For example, in disembedding tasks, participants search for a specified 2D shape in a larger distracting image. *Intrinsic-dynamic* skills involve the processing, and manipulation or transformation of objects or shapes. 2D and 3D mental rotation fit into this category.

Extrinsic-static skills require the processing and encoding of the spatial relations between objects or configurations of objects, without further manipulation or transformation of these relations. The extrinsic-static category includes the ability to find corresponding locations between shapes of equal proportion but differing sizes (e.g. scaling and map use). *Extrinsic-dynamic* skills involve the apprehension, processing and manipulation of more than one object, or the relationship between objects and frames of reference. Spatial perspective taking, in which a participant visualises what an object would look like from a different viewpoint, is an extrinsic-dynamic skill because it involves the manipulation of the relationship between an object and another frame of reference/viewpoint.

Spatial ability and science

Spatial skills support understanding and learning of conceptual areas that are very spatial in nature (e.g. astronomy) yet even apparently non-spatial topics are often represented in a spatial format.

Most prior research with adults point to visualisation skills as being related to science learning: the ability to mentally transform spatial information about single objects, assessed through intrinsic-dynamic skills such as mental rotation and mental folding. For example, studies report a link between intrinsic-dynamic spatial skills and conceptual understanding in aspects of biology (e.g. Garg, Norman, Spero & Mashewari, 1999), chemistry (e.g. Stull, Hegarty, Dixon & Stieff, 2012) and physics (e.g., Kozhenikov & Thornton, 2006). In Stull et al. (2012), for instance, 3D visualisation positively correlated with undergraduate students' ability to translate between different diagrammatic representations of chemical structures.

Other spatial skills within Uttal et al.'s (2013) model may play a role in science learning. However, this relationship has been largely neglected to date; the role of extrinsic-static skills such as scaling, for example, has yet to be addressed.

Spatial ability and science in development

Research relating spatial ability and science learning in younger children is sparse (e.g. Jarvis & Gathercole, 2003;

Tracy, 1990). Tracy (1990) assessed science performance in a sample of 10- and 11-year-old students who were split into high and low spatial ability. The study revealed that the high spatial ability group outperformed the low spatial ability group. However, this study did not include any measure of IQ or other cognitive factors, and thus did not discount general ability as an alternative explanation. It also used a composite spatial measure.

One unpublished study that compared the role of different intrinsic-dynamic spatial ability measures on scientific understanding in children, and controlled for verbal ability, found that mental folding, but not mental rotation, predicted five-year-olds' performance in a task involving understanding of force and motion. However, this was still limited to intrinsic-dynamic skills (Harris, 2014).

There is also some mixed evidence to suggest that spatial skills may be more important during the early stages of science learning, rather than later stages (Hambrick et al., 2012). During initial learning a learner may use spatial processing to establish mental maps and models or to problem solve (Mix et al., 2016). With experience, domain-specific knowledge may become more important. Such a hypothesis is supported in the science literature by the finding that visuospatial working memory is less predictive of 14-year-old versus 11-year-old students' science performance (Jarvis & Gathercole, 2003), and that mental folding ability predicted children's but not adults' understanding of forces in the previously described study by Harris (2014). The relationship between spatial thinking and science performance, and how this varies between the ages of 7 and 11, has hitherto remained unclear.

Current Study

Although prior research indicates that spatial ability predicts aspects of science learning in older populations, little research has been conducted with younger children. Research that has done so, has either focused on visuospatial working memory only, used a composite of spatial ability measures, or has focused only on single object-based manipulation (intrinsic-dynamic spatial skills). Furthermore, no research to date has used and compared a cross-sectional sample to determine if this relationship varies across development.

The aim of the current study was to examine the relationship between 7 and 11-year-olds' performance in a range of spatial ability measures, based on the Uttal et al. (2013) model, and their performance in a science assessment covering aspects of biology, chemistry and physics.

Methods

Participants

The initial sample consisted of 127 participants who were recruited from a large, ethnically diverse London primary school. Three pupils did not go on to complete the study

because they were unsuitable for the study due to having a special educational need or an insufficient level of English. Due to missing data, four further participants (one participant per year group) did not have a full set of scores. Three of those participants were missing data from one task only, and so their missing scores were estimated by calculating the mean for their respective year group. The fourth participant, who was missing several variables, was excluded from the analysis. Thus, four participants were excluded in total. The final sample consisted of 123 participants in UK Years 3-6: Year 3 (N=32, mean (s.d.) age=8.0 (0.28) years), Year 4 (N=31, mean (s.d.) age=9.0 (0.32) years), Year 5 (N= 31, mean (s.d.) age=10.0 (0.33) years), Year 6 (N= 30, mean (s.d.) age=11.0 (0.30) years). Parental consent was obtained for all participants.

Measures

Intrinsic-Dynamic Spatial: Mental Rotation

In this mental rotation task (based on Broadbent, Farran & Tolmie, 2014), children were shown two upright cartoon monkeys, above a horizontal line, on a computer screen, and one monkey below a line which was rotated by varying degrees (0°, 45°, 90°, 135°, 180°). One monkey above the horizontal line had a blue left hand and a red right hand, and the other monkey had the reverse pattern and was a mirror image of the other. Children were asked which of the two monkeys at the top of the screen matched the monkey at the bottom of the screen, which had been rotated. This task began with six practice items, in which the monkey below was not rotated (0° trials) and then progressed to 36 experimental trials (4 x 0° trials, 8 x 45° trials, 8 x 90° trials, 8 x 135° trials and 8 x 180° trials).

Intrinsic-Dynamic Spatial: Mental Folding

This mental folding task (Harris et al., 2013) required children to imagine folds made to a piece of paper, without physical representation of the folding action itself (see Figure 1). Children were shown a shape at the top of a computer screen which contained a dotted line and an arrow. The dotted line represented the imaginary fold line and the arrow indicated where the paper should be folded to. Beneath this item on the screen, children were shown four images of how the item at the top might look after being folded at the dotted lines, only one of which was correct. Children first completed two practice items. Answers to practice questions were checked by the researcher, and if a child had an incorrect answer, they were given one further attempt of that trial. The experimental trials then began, where children had 14 novel items to work through. The test progressed automatically as the child clicked one of the four images at the bottom of the screen. Accuracy and response time were recorded.

Intrinsic-Static Spatial: Embedded Figures Task

The Children's Embedded Figures Task (CEFT: Witkin et al., 1971) consists of complex figures in which a simple form is embedded (see Figure 1). Children were shown an

image constructed of colourful geometric shapes and asked to locate either a simple house or tent shape 'hidden' within the image. Children were shown the house or tent shape as a cardboard form; it was kept by the child for the first three trials of each block and hidden thereafter. For the first part of the test (11 items) children located a triangular tent shape within each image, the simpler of the two shapes, and for the other half of the test (14 items) children located a house shape. Children were given a score of 1 for correctly locating the shape hidden within the figure. Accuracy and response times were recorded using a laptop computer.

Extrinsic-Static Spatial: Scaling

Our novel spatial scaling task was based on a similar task by Frick and Newcombe (2012). Children were required to find equivalent corresponding locations on two grids, when one was varied in size relative to the other by a predetermined scale factor (see Figure 1). The task was presented to children as a game which involved pirates' treasure maps. Treasure maps were printed on yellow paper and mounted in a large ring bound pad. Each page contained one yellow map with a grid printed in black. Nine (out of 18) items contained grids which separated the map into 6 x 6 sections, whereas the other nine items contained grids which separated the map into 10 x 10 sections. For each trial, one target section of the printed grid map was coloured in black (the treasure); the target section varied across trials. Participants were also presented with four maps on a touch screen computer, which each had one black square coloured. One computer map contained a black square at a location which corresponded to the printed map; the locations of the black squares on the other three incorrect computer options were systematically chosen. The larger printed maps were either unscaled (1:1), or scaled to either 1:4 or 1:8, relative to the maps on the computer screen. Participants first completed two practice items after which they completed the main 18 trials of the test. Six items were presented at each scale factor.

Extrinsic-Dynamic Spatial: Perspective Taking Task

This task was identical to one developed by Frick et al. (2014) which involved spatial perspective taking in which children were required to visualise what photographs would look like when taken from cameras placed at different positions and angles relative to their viewpoint. The child first completed four practice questions involving actual Play Mobil characters, and then one practice question on a computer, which showed a character taking a photograph of two shapes from the same perspective as the child. The child selected the correct option, of four, from below the main image, which showed what the photograph would look like by pressing a touch screen computer. If a child made an error on the practice items, they were given a second attempt. On passing the practice questions, the task continued with the experimental items, where they again chose the correct image out of four options. These varied per the number of objects in the layout (one, two or three)

and the angular difference between the photographer's and the child's perspective (0°, 90° or 180°).

Science Assessment

The science assessment consisted of two paper-based tests, which children completed in class groups in two separate sessions under the supervision of the researcher. The assessment was curriculum-based and all questions were taken from an online database of past science UK standardised test questions ("Test Base", 2017, January 31st) designed to assess the science curriculum in this age range. The test included approximately equal amounts of biology, chemistry and physics content from a selection of topics appropriate to this curriculum stage.

Each paper had a total possible score of 50 marks leading to a total science mark of 100. The assessment included questions which varied in difficulty, which again mapped onto curriculum descriptors. Paper one contained questions of low to medium cognitive demand, and paper two contained questions of high cognitive demand. Questions focused on one conceptual sub-topic, e.g., magnetism or changing state. Questions were further sub-divided into items that were either: factual/recall items (e.g. label a diagram; recall a function); problem solving items, which drew on conceptual knowledge; or items that were in the context of hypothetical experiments, thus drawing on procedural skills.

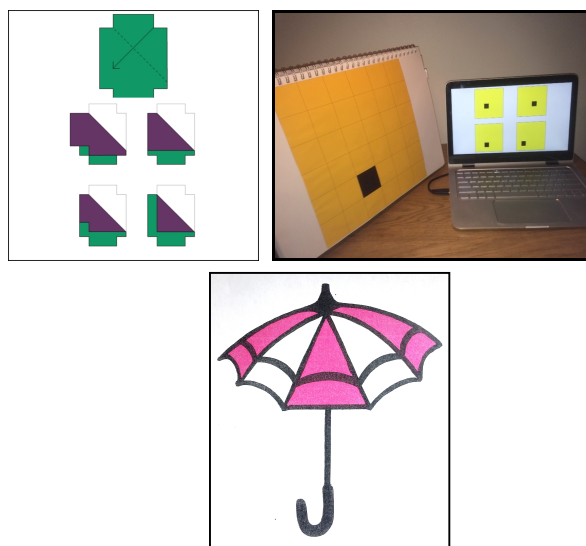


Figure 1: significant spatial predictors (left to right): mental folding task, spatial scaling task (1:8, 6x6 trial), embedded figures task (locate a triangular 'tent' trial).

Control Variables

The British Picture Vocabulary Scale-III (BPVS-III; Dunn, Dunn, Styles, & Sewell, 2009) was included as a measure of verbal ability.

Procedure

All children first completed the two paper-based science assessments, in two sessions administered by the researcher in class groups, within the child's own classroom. Spatial ability was then assessed within two separate sessions. Children were first tested in a computer-based group session of no more than 10 children where they completed the mental folding task and the monkey mental rotation task, in a counterbalanced order. The BPVS, CEFT, spatial perspective taking task and scaling task were then completed in an individual testing session with the researcher, which lasted approximately 30 minutes per child. The order of testing in the individual sessions was also counterbalanced to control for fatigue and order effects.

Results

A series of multiple regression analyses were conducted to determine the amount of variance in science scores that was accounted for by each of the spatial ability measures. There were no significant gender differences ($p > .05$ for all); therefore, participants were treated as one group in the subsequent regression analysis. A separate analysis was run for total science score and for each area of science (physics, biology, chemistry). In each of the models, the covariates (age, BPVS raw score) were added first (steps 1-2). All spatial measures were subsequently entered in a single block, using forward step-wise entry, to determine the best model of spatial ability predictors. Beta values refer to the final model with all variables entered.

Entered in the first step of each model, age in months significantly predicted each science total. Age was most strongly predictive of physics sub-score, with this initial step accounting for 27% of the variance, $\Delta F(1,121) = 45.52$, $p < .001$. Age accounted for 21% of the variance in total score, $\Delta F(1,121) = 31.27$, $p < .001$, 13% of the chemistry scores, $\Delta F(1,121) = 17.75$, $p < .001$ and 8% of biology scores, $\Delta F(1,121) = 11.03$, $p < .001$. In the final overall models, after additional variables were entered, age remained as a significant predictor of total score ($\beta = .122$, $t = 2.03$, $p = .044$) and physics score ($\beta = .320$, $t = 4.24$, $p < .001$), but not biology score ($\beta = .010$, $t = .10$, $p = .916$) or chemistry score ($\beta = .102$, $t = 1.23$, $p = .223$).

BPVS raw score was entered in the second step of each model and was a significant predictor. BPVS scores were most strongly predictive of total science score, accounting for an additional 37.6% of the score variance, $\Delta F(1,120) = 106.16$, $p < .001$, 25% of the biology scores, $\Delta F(1,120) = 46$, $p < .001$, 21% of the chemistry scores, $\Delta F(1,120) = 36.93$, $p < .001$ and 15% of the physics scores $\Delta F(1,120) = 29.78$, $p < .001$. In the final model, BPVS scores remained significant predictors of total science score ($\beta = .567$, $t = 8.9$, $p < .001$), biology score ($\beta = .443$, $t = 5.38$, $p < .001$), chemistry score ($\beta = .485$, $t = 5.93$, $p < .001$) and physics score ($\beta = .375$, $t = 4.85$, $p < .001$).

The step-wise analysis of spatial measures to predict total science score resulted in mental folding being entered in

step three and spatial scaling being entered in step four of the regression model. Mental folding accounted for an additional 6% of the variance in total science score, above the previously entered covariates $\Delta F(1,119) = 20.62$, $p < .001$, $\beta = .211$, $t=3.54$, $p = <.005$. The scaling task accounted for an additional 2% of the variance in total science scores, $\Delta F(1,118) = 6.8$, $p < .010$, $\beta = .162$, $t = 2.60$, $p = .010$).

Step-wise entry of spatial measures predicting biology scores retained mental folding and the scaling task only, in steps three and four. Mental folding accounted for an additional 6% of the variance in biology scores, $\Delta F(1,119) = 11.65$, $p < .001$, $\beta = .195$, $t = 2.52$, $p = .013$. Spatial scaling accounted for an additional 3% of the variance in biology scores, above the previously entered covariates $\Delta F(1,118) = 5.50$, $p < .021$, $\beta = .190$, $t = 2.34$, $p = .021$).

Only the mental folding task was retained as predictor of the physics score ($\beta = .198$, $t = 2.8$, $p = .006$) following step-wise analysis, and was entered in step three of the model. It accounted for an additional 4% of the variance in physics scores $\Delta F(1,119) = 7.82$, $p = .006$).

The CEFT was the only emerging predictor of chemistry scores ($\beta = .173$, $t = 2.27$, $p = .025$). It accounted for an additional 3% of the variance in chemistry scores $\Delta F(1,119) = 5.130$, $p = .025$). Any remaining spatial tasks not reported did not significantly predict any additional unique variance in science ability beyond those spatial measures included in the above models ($p > .05$ for all).

To determine if age interacted with any of the spatial ability measures, a further four models were constructed in which the covariates were again entered in step 1, followed by the spatial ability measures found to be significant for that science measure, followed by interaction terms (age in months*spatial measure). No significant interactions with age were found ($p > .05$ for all).

Discussion

This study revealed that spatial ability predicted children's performance in a curriculum-based science assessment. That is, after controlling for age, gender and verbal ability, spatial ability accounted for 8% additional variance in total science scores. This builds upon longitudinal research linking spatial ability to STEM outcomes in adults (Wai et al., 2009) as well as correlational research that has associated spatial ability with various aspects of science learning in adults (e.g., Kozhevnikov et al., 2006). It also expands upon the existing findings from child data (Tracy, 1990; Harris, 2014) by investigating a broader range of spatial measures and science topic areas and, also, comparing a wider age range of children within one study.

Mental folding, an intrinsic-dynamic spatial skill, accounted for the most variance in total overall science scores, relative to the other spatial tasks. This is likely to be due to it being a predictor of both physics and biology topics, whereas spatial scaling was not. It is likely that the relationship with physics scores is driven, in part, by

questions on topics such as magnetism, forces and motion, which required visualisation of how objects move and interact; this suggests that children who are more skilled at visualising paper folds are also better at predicting the direction of various types of forces acting on objects.

The role of mental folding in relation to biology scores is less likely to be directly related to the visualisation skill, as in physics, discussed above. One possibility is that the ability to flexibly maintain and manipulate spatial information, as measured through the folding task, may also be related to mental model construction and use. The mental models children possess for the conceptual topics within biology may be spatially-based. For example, when recalling the function of roots, children may recall a spatial mental model of a plant, which is integrated with verbal/propositional information.

Although mental rotation falls into the same category (Uttal et al., 2013) as mental folding (which was a strong predictor), mental rotation accuracy did not feature in any of the final regression models. The two measures correlated only moderately ($r = .294$). This may be because there are differences between folding and rotation, despite them both being intrinsic/dynamic measures. For example, rotation is a rigid, intrinsic transformation and folding is a non-rigid, intrinsic transformation (Atit, Shipley & Tikoff, 2013). Further research is needed to investigate this distinction.

Spatial scaling, an extrinsic/static skill, also emerged as a predictor of total science scores, although it contributed less to this model than folding because it was significant only for biology questions. One interpretation of the role of scaling is that it predicted performance because children who perform well on this task are also more able to determine the correspondence between representations of scientific concepts at different scales. Children may, for example, in the classroom, need to determine the correspondence between: an actual plant; scaled-up versions of plant diagrams on an interactive whiteboard; or scaled-down, abstract printed diagrams.

The CEFT, an intrinsic-static spatial skill, was a significant predictor of chemistry scores, but did not feature in the other final regression models. Intrinsic-static spatial skills relate to the processing of objects without further transformation: the arrangements of parts of the object (sub-parts) as well as the size and orientation objects. This could relate to chemistry items including diagrams which require inspection and discrimination between sub-parts of objects (e.g. 3 beakers with ice cubes, which either have 1, 2 or 3 layers of insulation).

An analysis of age-based interactions revealed that the relationships described were steady across development. We had predicted that spatial skills would contribute more to science performance for younger children, than older children, based on prior research (e.g., Hambrick et al., 2012), which would suggest that as domain-specific knowledge increases, spatial abilities play less of a role in science. It may be that, although older children are more experienced in science, the knowledge they have available

for immediate recall may be restricted to topics that they are currently covering.

One limitation of the study was that we only included the BPVS as a control for general level of ability. One might argue that the spatial tasks are also capturing a more general problem solving ability. However, the differential role of various spatial skills revealed in the analyses demonstrate that spatial ability is having an impact on science performance versus it being a general problem solving ability proxy. However, further research should include other measures of general cognitive ability. Second, the nature of the questions, drawn from standardised assessments, meant that it is difficult to determine if the relationships observed relate to scientific knowledge/understanding, or application of knowledge in scientific problem solving. Further studies are planned to systematically include a range of question types for comparison across categories of items.

In summary, the current study provides evidence for a distinctive role for mental folding (intrinsic-dynamic spatial ability), spatial scaling (extrinsic-static spatial ability) and the CEFT (intrinsic-static spatial ability) in children's science learning. The findings have implications for how we can move forward to support children in the science classroom through spatial training and interventions.

Acknowledgements

This work was supported by an Economic and Social Research Council PhD studentship awarded to the first author.

References

- Atit, K., Shipley, T. F., & Tikoff, B. (2013). Twisting space: are rigid and non-rigid mental transformations separate spatial skills? *Cognitive processing*, 14(2), 163-173.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: University Press.
- Broadbent, H. J., Farran, E. K., & Tolmie, A. (2014). Object-based mental rotation and visual perspective-taking in typical development and Williams Syndrome. *Developmental Neuropsychology*, 39(3), 205-225.
- Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *The British Picture Vocabulary Scale - 3rd Edition*. London: GL Assessment.
- Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds. *Frontiers in Psychology*, 5, 386.
- Frick, A., & Newcombe, N. S. (2012). Getting the big picture: Development of spatial scaling abilities. *Cognitive Development*, 27(3), 270-282.
- Garg, A., Norman, G. R., Spero, L., & Maheshwari, P. (1999). Do virtual computer models hinder anatomy learning? *Academic Medicine*, 74(10), S87-89.
- Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., LaDue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *Journal of Experimental Psychology: General*, 141(3), 397.
- Harris, J. (2014). *Where will it go? Concepts of motion in complex events*. (Unpublished doctoral thesis). Temple University, Philadelphia, Pennsylvania.
- Harris, J., Newcombe, N. S., & Hirsh-Pasek, K. (2013). A new twist on studying the development of dynamic spatial transformations: Mental paper folding in young children. *Mind, Brain, and Education*, 7(1), 49-55.
- Hegarty, M. (2014). Spatial thinking in undergraduate science education. *Spatial Cognition & Computation*, 14(2), 142-167.
- Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review. *Educational Psychology Review*, 22(3), 245-269.
- Jarvis, H. L., & Gathercole, E. E. (2003). Verbal and non-verbal memory and achievements on national curriculum tests at 11 and 14 years of age. *Educational and Child Psychology*, 20(3), 123-140.
- Kozhevnikov, M., & Thornton, R. (2006). Real-time data display, spatial visualization ability, and learning force and motion concepts. *Journal of Science Education and Technology*, 15(1), 111-132.
- Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology: General*, 145, 1206-1227.
- Newcombe, N. S. (2010). Picture this: increasing math and science learning by improving spatial thinking. *American Educator*, 34(2), 29.
- Stull, A. T., Hegarty, M., Dixon, B., & Stieff, M. (2012). Representational translation with concrete models in organic chemistry. *Cognition and Instruction*, 30(4), 404-434.
- Test Base. (2016, January 31st). Retrieved from <http://www.testbase.co.uk/sec/index.php>.
- Tracy, D. M. (1990). Toy-playing behavior, sex-role orientation, spatial ability, and science achievement. *Journal of Research in Science Teaching*, 27(7), 637-649.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial Ability for STEM Domains: Aligning Over 50 Years of Cumulative Psychological Knowledge Solidifies Its Importance. *Journal of Educational Psychology*, 101, 817-835.
- Witkin, H. A. (1971). *Children's Embedded Figures Test*, Consulting Psychologists Press, Palo Alto, CA.