ASE

Different Perspectives: Spatial Ability Influences Where Individuals Look on a Timed Spatial Test

Victoria A. Roach,[1] Graham M. Fraser,[2] James H. Kryklywy,[3] Derek G.V. Mitchell,[3,4] Timothy D. Wilson[4]*

[1]Department of Biomedical Sciences, William Beaumont School of Medicine, Oakland University, Rochester, Michigan

[2]Department of Medical Biophysics, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada

[3]Department of Psychiatry, Brain and Mind Institute, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada

[4]Department of Anatomy and Cell Biology, Corps for Research of Instructional and Perceptual Technologies, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada

Learning in anatomy can be both spatially and visually complex. Pedagogical investigations have begun exploration as to how spatial ability may mitigate learning. Emerging hypotheses suggests individuals with higher spatial reasoning may attend to images differently than those who are lacking. To elucidate attentional patterns associated with different spatial ability, eye movements were measured in individuals completing a timed electronic mental rotation test (EMRT). The EMRT was based on the line drawings of Shepherd and Metzler. Individuals deduced whether image pairs were rotations (same) or mirror images (different). It was hypothesized that individuals with high spatial ability (HSA) would demonstrate shorter average fixation durations during problem solving and attend to different features of the EMRT than low spatial ability (LSA) counterparts. Moreover, question response accuracy would be associated with fewer fixations and shorter average response times, regardless of spatial reasoning ability. Average fixation duration in the HSA group was shorter than LSA ($F(1,8) = 7.99$; $P = 0.022$). Importantly, HSA and LSA individuals looked to different regions of the EMRT images (Fisher Exact Test: 12.47; $P = 0.018$); attending to the same locations only 34% of the time. Correctly answered questions were characterized by fewer fixations per question ($F(1, 8) = 18.12$; $P = 0.003$) and shorter average response times ($F(1, 8) = 23.89$; $P = 0.001$). The results indicate that spatial ability may influence visual attention to salient areas of images and this may be key to problem solving processes for low spatial individuals. Anat Sci Educ 00: 000–000. © 2016 American Association of Anatomists.

Key words: gross anatomy education; spatial reasoning; mental rotation; spatial ability; eye tracking; gaze patterns; visual salience; timed tests

INTRODUCTION

The term "spatial ability" is often used to describe an individual's aptitude for interpreting three-dimensional relationships in space (Lohman, 1996). A prevalent topic in cognitive psychology for decades, work has sought to explore not only spatial ability itself, but its related sub-skills, spatial visualization, spatial orientation, and spatial relations (Lohman, 1988). Spatial relations, often referred to more generally as mental rotation ability (MRA), (Linn and Petersen, 1985; Voyer et al., 1995), is the ability to visualize the translation of an object about an axis, while recognizing that it is the same from any perspective (Shepard and Metzler, 1971).

It is accepted across the literature that mental rotation, like many other complex cognitive processes in humans, requires multiple brain areas for success (Cohen et al., 1996; Booth et al., 2000). In a recent meta-analysis of cortical activation during mental rotation, it was concluded that the superior parietal, frontal, and inferotemporal cortices were

*Correspondence to: Dr. Timothy D. Wilson; Department of Anatomy and Cell Biology, MSB 490, CRIPT Lab, The University of Western Ontario, London, Ontario N6A5C1, Canada. E-mail: tim.wilson@uwo.ca

Received 29 April 2016; Revised 2 August 2016; Accepted 3 September 2016.

Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/ase.1654

consistently activated during mental rotation (Zacks, 2008). The activity was observed bilaterally in most areas; however, activity in the parietal cortex was more consistently observed in the right hemisphere, while activity in the frontal cortex was more consistently observed in the left hemisphere (Zacks, 2008).

More specifically, Zacks (2008) suggests that the foci of cortical activation found in the superior parietal, frontal and inferotemporal cortices contribute to a large focal activation area surrounding the intraparietal sulcus, and roughly equates to Brodmann's areas 7 (superior parietal lobule), 19 (secondary visual cortex), 39 (angular gyrus), and 40 (inferior parietal lobule) (Zacks, 2008). These findings are well supported by neuropsychological data obtained by Ratcliff, who found consistent activation of the superior parietal cortex during mental rotation tasks (Ratcliff, 1979). Moreover, the posterior parietal cortex (and the brain regions extending into the superior posterior occipital cortex) is consistently activated during mental rotation across a range of tasks, imaging modalities, and statistical analysis strategies (Zacks, 2008). As a result, it is reasonable to suggest that this region may implement the transformation-specific computations required to complete mental rotation. This finding aligns with the work of Farah, who conducted neuropsychological studies on mental rotation (Farah, 1989) and the work in transcranial magnetic stimulation conducted by Harris and Miniussi (2003).

Spatial ability, and mental rotation ability have drawn particular attention in anatomical and medical education circles because of their numerous linkages to success in surgical skill acquisition (Wanzel et al., 2002, 2003; Brandt and Davies, 2006), and anatomical knowledge acquisition (Lufler et al., 2012; Nguyen et al., 2012; Zumwalt et al., 2015). These linkages are supported by the fact that success in both anatomy, and surgical skill mastery is reliant on a firm understanding of the interactions of three-dimensional structures in the visually complicated environment of the human body (Lufler et al., 2012; Nguyen et al., 2014; Zumwalt et al., 2015). Recently Nguyen has also suggested that spatial ability may serve as a robust predictor for success in anatomy comprehension (Nguyen et al., 2012), and recommends that pedagogical efforts be taken to train spatial ability in those who lack aptitude for the cognitive skill, so that they may interpret spatially intricate anatomical and surgical information with less difficulty.

Research in cognitive psychology suggests that mental rotation may be intrinsically linked to the movements of the eye, as fixations (the maintenance of visual gaze on a single location for a defined period of time) (Carpenter, 1988) are intimately involved in our ability to visually encode spatially pertinent information (Just and Carpenter, 1976; Shepard and Cooper, 1986). Thus, eye fixations may serve as a representation of the cognitive stages (i.e., search, transformation and comparison, and confirmation) that occur during processing of visually complex anatomical information (Just and Carpenter, 1976).

In a pioneering study, Just and Carpenter (1985) explored how eye movements related to spatial reasoning using a simple, yet spatially complex untimed task. Under their paradigm, significant differences in the eye movement metrics of individuals of high mental rotation ability (HMRA) and low mental rotation ability (LMRA) were identified using simulations (Just and Carpenter, 1985), derived from samples of participants completing a test of cube comparisons (Ekstrom et al., 1976). On average, LMRA individuals exhibited longer trial response times, and conducted more fixations per trial. However, the results obtained by Just and Carpenter (1985) were in contrast with the results observed recently by Roach et al. (2016), where a similar test of mental rotation ability (the electronic mental rotation test, EMRT), was employed without a time restriction. Roach and colleagues found little relationship between most chronological (average response time) and numeric (number of fixations per question) measures of eye movement and MRA; but upon conducting an analysis of visual salience, (i.e., where on the images participants devoted visual attention), found wide differences between the HMRA and LMRA groups (Roach et al., 2016).

Visual salience can be defined as a quality, or a set of characteristics, held by a region of an image that is perceived by the viewer as being conspicuous relative to its surroundings (Itti et al., 1998). In the aforementioned EMRT experiment, the authors refer to regions of the images as being "spatially salient." These "salient domains" of the presented stimuli conveyed depth and positional information, and commanded visual attention, which may have facilitated accurate, and quick completion of the untimed EMRT (Roach et al., 2016). Despite studies evaluating the performance of HMRA and LMRA individuals in untimed environments, little is known about how individual eye movement behaviors differ in time-restricted, or "speeded" testing environments. This is of particular importance for large university level courses, such as anatomy and histology, where strict time limits are routinely applied in the context of evaluation. In anatomy, assessments often occur as timed bell-ringers, where students are queried orally or in writing on the identity, function, innervation, clinical relevance, or vascular supply of various pre-arranged, immobilized anatomical structures. This method of evaluation, designed to expedite large volumes of students through the evaluation process, and streamline the marking and feedback processes, poses a significant problem. Little consideration is given to the impact that these restrictions pose on the students' ability to answer effectively, and to the impact that the immobilization of the specimens may pose on low mental rotation ability individuals—who may struggle with orienting anatomical structures positioned in an unfamiliar position. With this in mind, while time-restricted assessment tools may be commonplace, these constrained evaluation environments may place another layer of perceptual challenge onto the low-ability student, that is outside the objectives of the test itself.

Like bell-ringer examinations, mental rotation ability is typically measured by performance on timed tests; including "gold standard" test of mental rotation: the paper-and-pencil based Vandenberg and Kuse mental rotations test (VKMRT) (Vandenberg and Kuse, 1978; Peters and Battista, 2008). Tests of this nature are commonly employed to stratify individuals as either high, intermediate or low MRA based on their individual score (Geiser et al., 2006) and have served to facilitate comparisons between groups according to their underlying spatial ability (Just and Carpenter, 1985; Nguyen et al., 2012). Investigations into participant accuracy on timed tests of spatial reasoning have reported that individuals who score higher on tests of mental rotation ability do so in less time, and with greater accuracy than those with lower spatial ability (Nguyen et al., 2014). With this considered, if accuracy and response time are implicated differently across levels of MRA in speeded testing environments, how might these findings manifest in the movements of the eyes during the completion of a speeded test of MRA?

The current study explores how specific eye movements relate to mental rotation ability (MRA) during the completion of a timed electronic test of mental rotations (timed EMRT). More specifically, the study aims to investigate the chrononumeric and visual apprehension patterns associated with accurately completing mental rotation tasks. In this instance, the term "chrononumeric" refers to a group of eye-related, and performance-related measures collected during the completion of the timed EMRT, (including the average number of fixations per question, average response time, and average fixation duration that were selected as indices of performance that represent individual behavior during completion of the test), wherein the term "fixation" refers to the maintenance of gaze on a single point for a period of time exceeding 200 milliseconds. As this overarching aim is composed of two distinct, yet related components (chrononumeric and visual apprehension measures), they will be addressed separately in the interest of clarity.

Relating Chrononumeric Patterns to Mental Rotation Ability

Through analysis of the chrononumeric data collected from participants during the completion of the timed EMRT, the current study aims to identify how the chrononumeric metrics may relate to MRA. Specifically, the current study aims to ascertain how average fixation duration, average response time and average fixations per question each relate to MRA score. Additionally, the current study also seeks to identify how response accuracy is related to average fixation duration, average response time, and average fixations per question. It is predicted that each chrononumeric metric relates to MRA in a different way and varies according to the accuracy of an individual's answer. It is hypothesized that average fixation duration will be shorter for HMRA individuals than for LMRA individuals, and it is expected that average fixation duration will be equivalent across both correct and incorrect answers. Like average fixation duration, it is expected that average response time will also be shorter for HMRA individuals than for LMRA individuals, but average response time will differ according to accuracy; being shorter on correct answers, than on incorrect answers. Finally, it is predicted that the average fixations per question will be consistent across both MRA groups, and equivalent across both correct and incorrect answers.

Relating Visual Apprehension Patterns to Mental Rotations Ability

As video-based corneal reflection eye tracking also yields spatiotemporal data during the performance of the Timed EMRT, the current study aims to ascertain if the attention directed to particular regions of a presented stimulus is contingent on MRA. That is, do individuals of different MRA look at different areas of images while problem solving? It is predicted that different regions of the presented stimuli will convey salience between the two groups, as measured by the Fisher Exact Test, illustrating that the two groups attend to different structures during spatial problem solving.

In short, it is predicted that each of the selected eye-tracking derived chrononumeric indices reflect MRA in a different way and will vary between correct and incorrect answers on the EMRT. Further, it is hypothesized that HMRA and LMRA individuals will view the EMRT images differently, which will be observable as differences in salience distribution patterns.

MATERIALS AND METHODS

Participants

Volunteer graduate students in the allied health sciences and anatomy and cell biology at The University of Western Ontario with normal, or corrected to normal vision by way of contact lenses, were invited to participate in this exploratory study, under approval from the institution's Research Ethics Board. Individuals ($n = 10$; 5 male and 5 female) first completed a standardized electronic Mental Rotations Ability test (EMRT) and the Vandenberg and Kuse Mental Rotations Test (VKMRT) (Vandenberg and Kuse, 1978; Peters and Battista, 2008). While traditionally the VKMRT is administered using a paper-and-pencil approach, an electronic iteration of the test, developed by Peters (Peters and Battista, 2008) has been used commonly in the literature (Nguyen et al., 2012; Roach et al., 2012; Nguyen et al., 2014; Roach et al., 2016). The electronic version of the test was identical in presentation to the original paper-and-pencil test. It employed the same scoring and timing parameters as the original (Vandenberg and Kuse, 1978), the same number of questions, and question stimuli as the original, and differed only as participants viewed the stimuli on a computer screen, and responded using a mouse click, rather than a pencil (Peters and Battista, 2008). More specifically, the VKMRT consisted of 24 questions, each worth 1 point per question. Participants were required to correctly identify both rotations of the target block figure in order to score a point. The time limit for this test was 6 minutes total.

In total, 26 individuals were prescreened using the VKMRT. Of the 26 pre-screened individuals, (mean VKMRT score: 8.3 ± 2.9), individuals with VKMRT scores that exceeded one standard deviation above the sample mean were considered to be HMRA ($n = 5$; 1 female and 4 males), and those with VKMRT scores less than one standard deviation below the sample mean were considered to be LMRA ($n = 5$; 4 females and 1 male). The individuals of these two groups thus represent the two extreme ends of the mental rotation ability distribution in the population. All other individuals ($n = 16$, mean VKMRT score: 8.25 ± 1.65; 7 male and 9 female) who demonstrated scores within one standard deviation of the sample mean in either direction were considered to have intermediate MRA, and were not included in this study. The division into HMRA and LMRA groups was adopted, rather than a median split, to exacerbate the distinction between HMRA and LMRA individuals (Just and Carpenter, 1985; Kozhevnikov et al., 2007; Roach et al., 2016).

Experimental Design

Following assortment into either HMRA or LMRA groups via the electronic VKMRT, participants completed the electronic Timed EMRT while monocular (right eye) gaze was monitored. Measurements of gaze were obtained according to the specifications detailed in Roach et al. (2016). All measurements were collected at a rate of 1,000 Hz using EyeLink 1000 eye-tracking equipment (SR Research Ltd., Mississauga, Ontario, Canada). Chrononumeric metrics consisting of average fixation duration, average number of fixations per question, and average response time were collected, along with the region of highest salience. Target images were viewed from a distance of 40 cm, such that each figure subtended approximately 10 degrees of visual angle, and the center-to-center distance between the two figures subtended approximately 15 degrees. Ambient light conditions were kept constant in the testing room at all times.

Target Images

The target images presented to the participants constituted a timed electronic Mental Rotations Test (timed EMRT) based on the original line drawings of Shepherd and Metzler, and used previously in Roach et al. (2016). This test required participants to view two 3D block figures (a "block pair"), and indicate if the pair was the same, or different by responding using two keys on the keyboard as quickly, and accurately as possible. A button-press of "1" indicated a "same" pair, while a "2" indicated a "different" pair. This test was composed of 16 questions, presently randomly in triplicate, to constitute 48 individual trials. Participants were allotted 6 seconds to answer each question, totaling an overall exposure of 4.8 minutes. In instances where the participant was unable to respond within the 6 seconds time-frame, the test automatically advanced to the next question. Correctly answered questions were awarded 1 point each, for a maximum total score of 48.

The use of eye tracking enabled the quantification of eye movements during the presentation of each question's block pair. By tracking the position of a participant's eye during EMRT performance, information pertaining to the participant's locus of attention on the image could be recorded. This information was compiled as a fixation map, where each point of fixation was overlaid onto the presented image. Each point of fixation was then overlaid with a Gaussian distribution to represent visual acuity (Lee et al., 2011), and scaled according to the duration of the fixation. This process effectively converted the positional "fixation map" into a salience map, which represented both the location and duration of attention for each individual, or salience, for each question. Each salience map was then normalized in magnitude based on the region of highest salience (the location with the most fixations, with the highest durations). Normalized maps for each image for each participant were then combined to yield mean group salience maps for comparison between groups.

The regions of highest salience were classified into six location-based categories based on the target stimuli (ICC: 0.88) by three observers (V.A.R., G.M.F., and T.D.W.) (Fig. 1). This identification system served to classify which regions of the blocks conveyed the greatest visual salience during the problem-solving process. The areas colored red represent the most attended region of the image, and convey the highest salience across the group.

Data Analysis

Chrononumeric analysis. The relationship between MRA score and accuracy (correct vs. incorrect answers) with average fixation duration, question response time, and number of fixations per question was analyzed via a 2 × 2 (MRA: HMRA or LMRA) × (Accuracy: Correct or Incorrect) Mixed ANOVA for each metric. Three Mixed ANOVAs were implemented because of the presence of both a between- (MRA) and within- (Accuracy) subjects variable of interest, and because the ANOVA is known to be robust to deviations in distribution normality secondary to small sample size (Schmider et al., 2010).

Analysis of salience. The location-based categorization of salience facilitated a between- group comparison across each category by frequency using the Fisher Exact Test (McDonald, 2014). This test was employed as the Fisher Exact Test is robust to small sample sizes, and is specific to categorical data (Fisher, 1922, 1932; McDonald, 2014), such as the six locations employed in the measurement of salience, illustrated in Figure 1. An additional comparison of question-by-question agreement was then conducted using Cohen's kappa to determine how often the HMRA and LMRA visually attended to the same location on a given question (McHugh, 2012).

The Electronic Mental Rotations Test (EMRT) and the Inclusion of the Low Mental Rotation Ability (LMRA) Group

Often in studies of performance, individuals of the highest performance ability are studied and used as exemplars for the behavior (Just and Carpenter, 1985). However, as the goal of this study, and of many other studies of mental rotation ability (Just and Carpenter, 1985; Gages, 1994; Mayer and Sims, 1994; Wanzel et al., 2002; Geiser et al., 2006; Kozhevnikov et al., 2007; Nguyen et al., 2012), is to distinguish between the characteristic differences in ability, it is necessary to include lower functioning individuals in analyses.

The EMRT is classified as a 2-Alternate Forced Choice (2-AFC) test, where the participant must make a decision of "same" or "different" when presented with a question (Fechner, 1860/1966). As some individuals in the LMRA group demonstrated EMRT scores approaching 50%, analysis to ensure that the individuals were actively engaged in the task, and not guessing (or scoring at the level of chance) was required. This was carried out through the use of the Binomial Test (Howell, 2007). The binomial test is an exact test of the significance of differences between an observed (actual) and expected (chance) set of observations, and effectively quantifies how closely the respondents' answers mirror the pattern expected by chance, or guessing (Howell, 2007). The binomial test revealed that the proportion of correct answers obtained by the low group (0.58) was higher than that expected by chance (0.5), $P = 0.025$ (1-sided). As such, it was found that the individuals in the LMRA group were performing statistically higher than that expected by chance, and confirmed that the LMRA group was not guessing as they completed the test. As a result, this finding re-affirmed the inclusion of the data derived from the LMRA group.

For all analysis, a value of less than $P = 0.05$ was considered to be statistically significant unless otherwise stipulated.

RESULTS

To ensure that the Timed EMRT was a valid measure of MRA, the scores on the Timed EMRT were correlated with scores on the VKMRT. The relationship between Timed EMRT scores and VKMRT scores was significantly

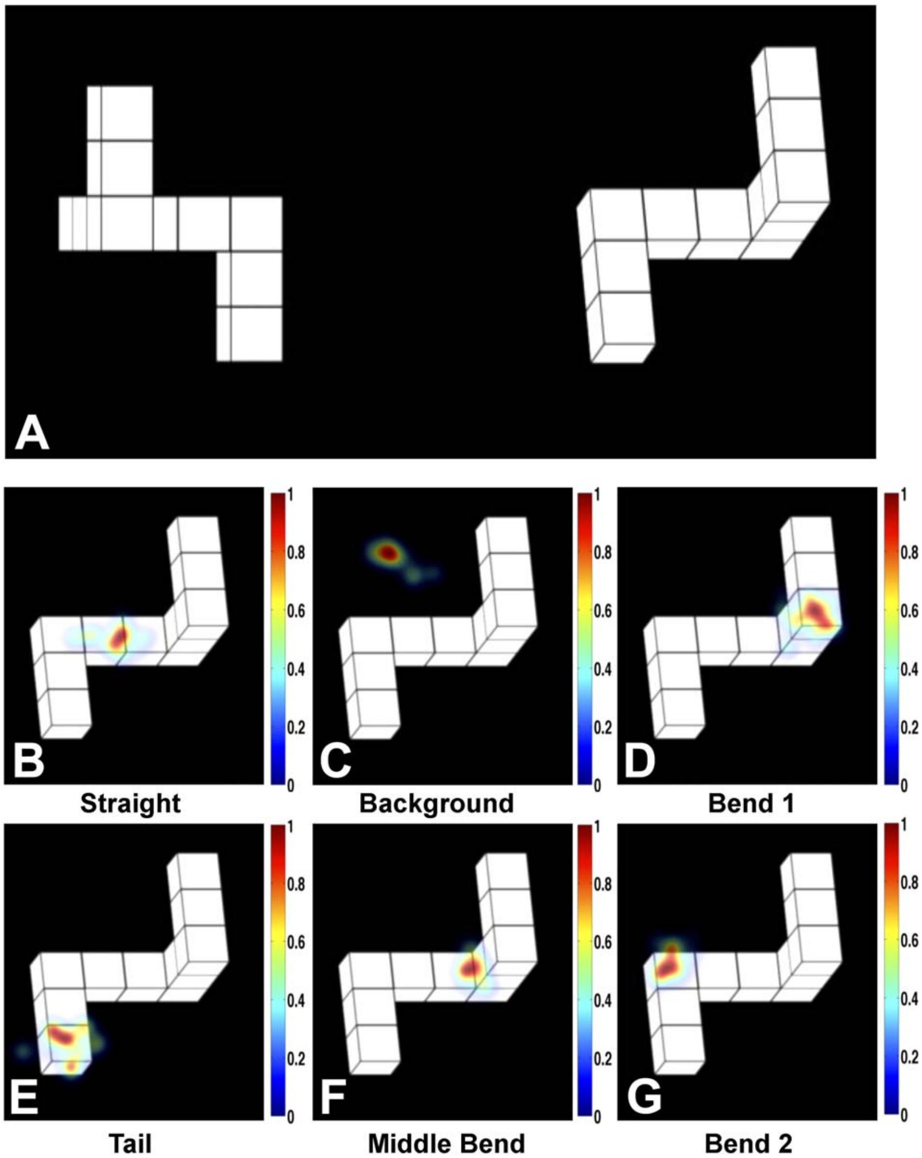

Figure 1.

Visual representations of test images. A, An exemplar image of a typical timed electronic mental rotation test (EMRT). Test image, designed using the line drawings of Shepard and Metzler (1971), adapted and digitized by Peters and Battista (2008), where participants must judge if the two block figures are the same (rotations) or different (reflections) within a span of 6 seconds; B–G, Heat maps depicting the six possible categorizations for the location of highest salience. The areas colored red represent the most attended region of the image, and are indicative of the highest salience across the group. Reprinted with permission from Roach et al. (2016).

positive, as tested by a Pearson Correlation ($r = 0.77$, $n = 10$, $P = 0.009$).

Relating Chrononumeric Patterns to Mental Rotation Ability

The average fixation duration (mean ± SD) of the HMRA group (223.24 ± 19.89 ms) was significantly shorter than the LMRA group (288.93 ± 46.18 ms) $F(1,8) = 7.99$ ($P = 0.022$) (Fig. 2A).

In addition to eye-related performance differences occurring between HMRA and LMRA groups, the within-group differences occurring in average fixation duration when questions were answered correctly, or incorrectly were also of interest. In this case, average fixation duration did not differ significantly according to question accuracy, for either group $F(1,8) = 0.011$ ($P = 0.918$) (Partial η^2: 0.001) (Fig. 2B). The average fixation duration (mean ± 95% confidence interval) for HMRA on correctly answered questions (226.56 ± 4.91 ms), and incorrect questions (219.92 ± 4.91 ms), compared with the LMRA group's correct (285.10 ± 7.99 ms) and incorrect answers (292.75 ± 7.99 ms).

Average response time was also analyzed in the same manner as average fixation duration, via 2 × 2 (MRA × accuracy) Mixed ANOVA to discern both group and participant

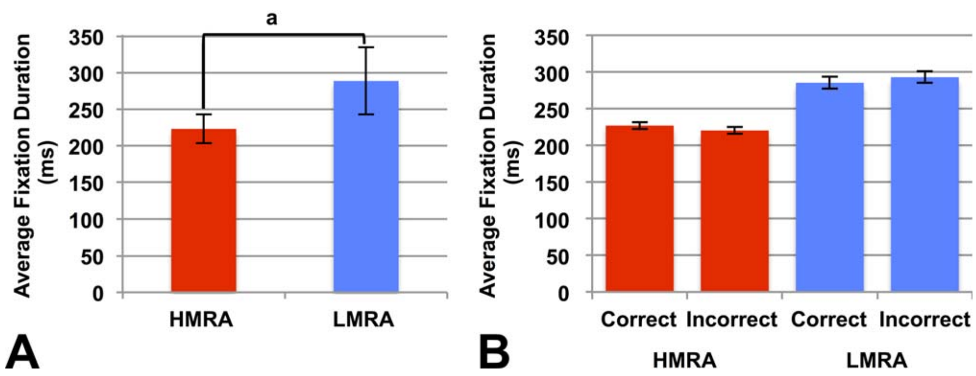

Figure 2.

Average fixation durations by group and by answer accuracy. A, High mental rotation ability (HMRA) individuals conduct fixations that are quicker than those completed by low mental rotation ability (LMRA) individuals, on average, during spatial problem solving. Error bars indicate one standard deviation ([a]$P = 0.022$); B, Average fixation duration is constant across different levels of accuracy for both mental rotation ability groups. Error bars indicate 95% confidence interval for within-group comparison.

level differences. Both groups responded with approximately equivalent response times (mean ± SD), HMRA: 3706.46 ± 752.26 ms and LMRA: 4669.57 ± 872.82 ms, ($F(1,8) = 4.82$; $P = 0.059$) (partial η^2: 0.376) (Fig. 3A).

The within-group analysis of the average response times for both correct and incorrectly answered questions query if answer accuracy is reflected in shorter average response time. Correctly answered questions (mean ± 95% confidence interval) were found to be significantly briefer (HMRA: 3357.17 ± 163.66 ms, LMRA: 4224.83 ± 274.11 ms) than those questions answered incorrectly (HMRA: 4055.74 ± 163.66 ms, LMRA: 5114.31 ± 274.11 ms); $F(1,8) = 23.89$ ($P = 0.001$) (Partial η^2: 0.749) (Fig. 3B).

The final eye movement related performance metric, average number of fixations per question, was also analyzed in the same manner as the previous two: via 2 × 2 (MRA × Accuracy) Mixed ANOVA to discern both group and participant level differences. Analysis suggests both MRA groups demonstrated equivalent average fixations per question (mean ± SD), (12.72 ± 2.99 and 14.27 ± 2.91) for HMRA and LMRA respectively ($F(1.8) = 0.853$; $P = 0.383$) (Partial η^2: 0.096) (Fig. 4A).

With regard to the role that answer accuracy imparts on the average number of fixations per question, within-participant analysis was carried out to address the hypothesis that the average number of fixations per question is equivalent across both correct and incorrectly answered questions. However, analyses found that the average number of fixations per question of correctly answered questions (HMRA: 11.50 ± 0.65 and LMRA: 13.01 ± 0.94) was significantly lower than incorrectly answered questions (mean ± 95% confidence interval) (HMRA: 13.95 ± 0.65 and LMRA: 15.53 ± 0.94), $F(1.8) = 18.12$ ($P = 0.003$) (Partial η^2: 0.682) (Fig. 4B). The current results align with the aforementioned average response time results for accuracy, as with briefer

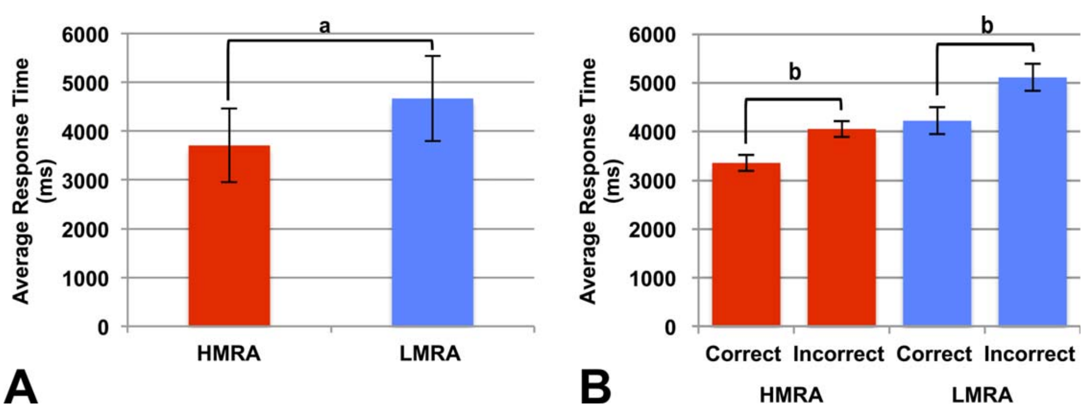

Figure 3.

Average response times by group, and by answer accuracy. A, High mental rotation ability (HMRA) and low mental rotation ability (LMRA) individuals exhibit equivalent response times on average ([a]$P = 0.059$). Error bars represent one standard deviation; B, High and low mental rotation ability individuals show different response times based on accuracy ([b]$P = 0.001$). Error bars represent 95% confidence interval for within group comparison.

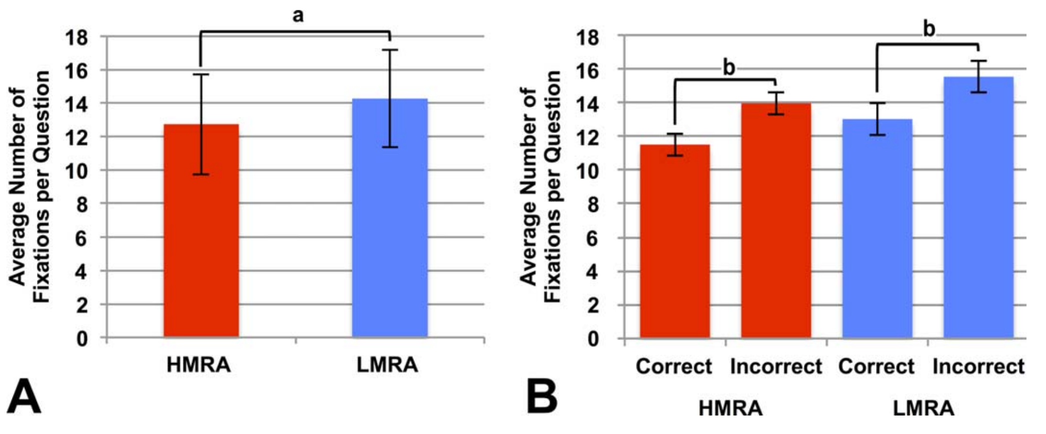

Figure 4.

Average number of fixations per question by group, and by answer accuracy. A, High mental rotation ability (HMRA) and low mental rotation ability (LMRA) individuals are equivalent in terms of average number of fixations per question ($^{a}P = 0.383$). Error bars represent one standard deviation; B, Correctly answered questions exhibit significantly fewer fixations per question than incorrectly answered questions ($^{b}P = 0.003$). Error bars represent 95% confidence interval for within group comparison.

overall response times, it is logical that fewer fixations may occur in a shorter time frame.

Relating Salience Patterns to Mental Rotations Ability

In addition to chrononumeric measures, the current experiment also sought to determine the level of attentional agreement for different regions of the target images according to spatial ability. A Fisher Exact Test distinguished the two groups based on the distribution across the six categories of salience (Fisher Exact Test: 12.47; $P = 0.018$) (Fig. 5). The test suggests that HMRA and LMRA individuals attended to different regions at different frequencies. High MRA individuals attend preferentially to "Straight" regions, while LMRA individuals attend predominantly to "Bend 1" during problem solving.

Further analysis was undertaken to establish a level of attentional agreement between HMRA and LMRA individuals on a question-by-question basis. Through application of Cohen's Kappa for Agreement, it was observed HMRA and LMRA individuals attended to the same region of salience in only 34% of questions ($\kappa = 0.20$) illustrating a poor agreement between the two groups (McHugh, 2012).

DISCUSSION

Through correlational analyses, it was noted that a significant and strong positive relationship existed between individuals' timed EMRT and VKMRT scores; suggesting that scores obtained on the timed EMRT are an accurate representation of mental rotations ability. This finding is not surprising, as both tests employ similar timing, and level of difficulty. As a result, for applications requiring eye movement recording, it

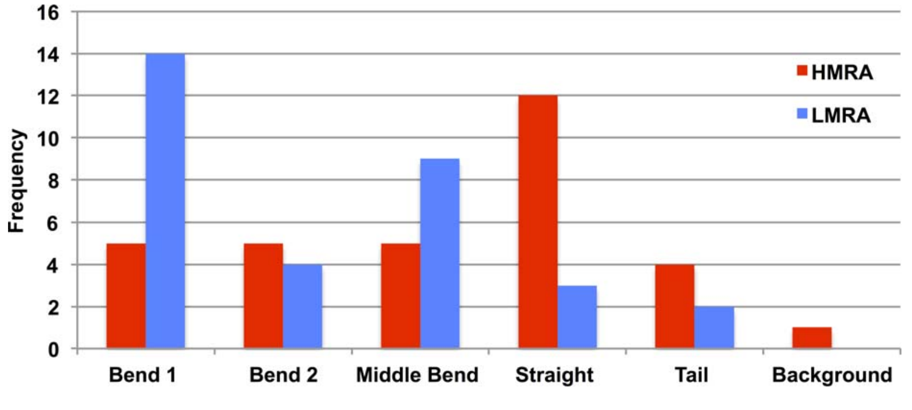

Figure 5.

Differing areas of attention on electronic mental rotations test (EMRT) represents dichotomy between areas of visual salience represented by frequencies fixation across the six categories by high mental rotation ability (HMRA) and low mental rotation ability (LMRA) individuals.

can be concluded that the timed EMRT provides a valid mechanism for the testing of MRA.

As previous research suggests that HMRA individuals complete MRA tests more quickly than LMRA individuals (Just and Carpenter, 1976; Roach et al., 2016), it was hypothesized that HMRA individuals would demonstrate shorter average fixation durations during problem solving. The average fixation duration of HMRA individuals was found to be significantly shorter than for LMRA individuals, suggesting that HMRA individuals are able to encode visual information more quickly during tasks than LMRA. This finding is supported by literature that suggests that individuals of HMRA have a higher spatial working memory capacity, and are thus better equipped to hold spatial representations in their minds eye for comparison during problem solving (Kyllonen and Christal, 1990; Kyllonen, 1996; Engle et al., 1999; Oberauer et al., 2000; Miyake et al., 2001). Higher spatial working memory may enable quicker comparisons between images and allow for a more consistently accurate problem solving process (Wilhelm and Schulze, 2002). In terms of accuracy, average fixation duration was found to be consistent within both MRA groups regardless of answer accuracy, and in line with the previous findings in an untimed environment (Roach et al., 2016). As a result, average fixation duration may be somewhat constant within individuals regardless of answer accuracy.

As average fixation duration differs significantly across MRA groups, one could infer that this difference would also be manifest in the average response times and average number of fixations per question of the two groups. However in this experiment, no difference in average response time, or average number of fixations per question was noted between HMRA and LMRA. This finding lies in contrast with available literature on other tests of spatial reasoning that suggest that apt individuals tend to perform spatial tasks with greater speed than those who struggle with spatial reasoning (Nguyen et al., 2014). However, the same pattern of consistent average response time and average number of fixations per question across MRA groups was also presented in previous findings (Roach et al., 2016). When the same testing parameters (EMRT) were employed without a time limit, both groups responded at approximately the same speed, and with the same number of fixations during problem solving, yet achieved different scores.

Further, the dichotomy of scores between the HMRA and LMRA individuals is not revealed when considering the accuracy of the question when evaluating the average response time and average number of fixations per question of individuals. For both groups, correctly answered questions were associated with significantly shorter response times, and significantly fewer fixations per question overall. Essentially, what can be inferred from these findings is that both groups complete the task in the same way. Questions that are answered correctly are answered quickly; while those that are not solved are puzzled over; directly supporting the findings observed by Roach et al. under the untimed condition (Roach et al., 2016). Ultimately, the frequency with which the question is solved accurately represents the aptitude for spatial reasoning. If individuals differ fundamentally in average fixation duration, but not in terms of average response time and number of fixations, further inspection into the location of these fixations is warranted. Therefore, it is possible that where individuals direct their attention on spatial tasks, this may be more critical to success in problem solving, than the speed with which an individual views the stimulus.

It was hypothesized that HMRA and LMRA individuals would attend to different features of the EMRT while problem solving; potentially contributing to varied success on the task. It was observed that HMRA and LMRA individuals attended to different salient regions of the EMRT images. On a question-by-question basis, the two groups attended to the same location only 34% of the time. This finding suggests that although groups share similar average response times and fixation times within questions, differences in gaze location exist between the two groups as they solve EMRT questions. This observation suggests that there may be specific, task-relevant regions of spatially complex images, and that skill in identifying and attending to these areas is key to successful problem solving (Lowe and Schnotz, 2008; Kaakinen et al., 2011). Furthermore, if specific regions are important to spatial reasoning, and are identifiable through salience maps, they may be translated as visual cues. These visual cues could be applied to train LMRA individuals, direct their attention, and potentially improve their spatial reasoning skills through guidance.

Limitations of the Study

This study may have been limited by the use of repeated image pairs. By presenting each image pair in triplicate, there is a possibility that with each subsequent exposure, the participants could have experienced a familiarity with the stimulus, and refer to previous conclusions. As a result, it is possible that individuals may have been reliant on short-term memory, or on implicit learning subsequent to contextual cueing (Chun and Jiang, 1998) to solve the question, than on active spatial reasoning. If this were true, it may have influenced the findings pertaining to average response time and average number of fixations per question, and reflect the individual's spatial working memory capacity more than their MRA. However, as the test battery consisted of 16 image pairs, which were each visually similar (one single block, differing only by degree of rotation or reflection) and were assorted randomly to each participant, it is unlikely that memorization or recognition was employed. Further, anecdotal inspection of the data did not convey any evidence of memorization in participant responses.

This study may also have been limited by a lack of individual factor evaluation, including but not limited to working memory, visual attention, conflict monitoring, attention deficit disorder, or the participant's presence on the autism spectrum. Each of these factors may have contributed to the observed results, but are not directly measured here. Future studies should seek to employ these metrics as covariates in their analysis, to gain better insight into the cognitive processes underlying mental rotation.

Moreover, it is known that question difficulty increases with increasing angular disparity. As the Timed EMRT employed increments of angular disparity only up to 80 degrees, it is possible that this test was not challenging enough to exacerbate substantial chrononumeric differences (Shepard and Metzler, 1971). If greater degrees of disparity were employed across the test, it is possible that HMRA individuals would have solved the problems more quickly, and revealed group-wise differences on the test.

Additionally, it is widely known that a sex difference that favors males exists in mental rotation literature. It appears that males consistently, and significantly outperform their female peers on tests of mental rotation. The rationale for the difference in performance is well disputed, and may be the result of the strategy employed by the test-taker. Specifically, most males tend to answer mental rotation test items instinctively, placing more value in speed than accuracy; while most females employ greater scrutiny and "double-check" their answers—placing more value in accuracy than speed (Lunneborg, 1982; Lunneborg and Lunneborg, 1984; Glück and Fabrizii, 2010). This dichotomy results in more answer opportunities for males, than females, and subsequently more opportunities for higher scores than females. With this in mind, it is possible that the patterns of eye movement observed in this study are not exclusively the result of the individual being high or low MRA, but may also reflect the strategic approach adopted by the individual.

Finally, this study may also have been limited by a small sample size. With each group consisting of five individuals, it is possible that the lack of differences in average response time and average number of fixations per question may be the result of reduced statistical power. As a result, future studies that explore this chrononumeric relationship alone should seek to include a larger sample. This being said, other studies in the literature that include a larger pre-sample often exclude the results, or even further testing of intermediate scoring participants that lay about the mean plus or minus one standard deviation (Just and Carpenter, 1985; Kozhevnikov et al., 2007; Nguyen et al., 2012). By studying the extremes of the spatial ability behavior a hypothesis regarding spatial ability in general can be formulated.

When considering the process of spatial problem solving, it appears that there are fundamental differences in the eye movements of HMRA and LMRA individuals. High MRA individuals are able to identify areas of the stimuli that are most important to accurately solving the problem quickly, while LMRA individuals look elsewhere, fixating on different locations for long periods of time, and ultimately reach incorrect answers more often than the HMRA individuals.

Similar findings have been obtained in the field of laparoscopy, in which the eye movements of expert and novice laparoscopists were contrasted during the performance of standardized laparoscopic tasks (Law et al., 2004; Wilson et al., 2010, 2011a, b). Significant differences in the patterns of eye movements between the novices and experts were observed, and successful approaches held by the experts were defined (Wilson et al., 2010, 2011a, b). This finding drove further inquiry to determine if novice individuals are directed to expert-defined task-relevant regions, would they have the cognitive capacity to interpret the information with greater facility and improve performance scores on tests of laparoscopic skill? In effort to answer this question, Wilson et al. created a gaze-directed training tool based on the eye movements of expert laparoscopists to train novice where to look on a specific laparoscopic technique (Wilson et al., 2011a, b). When evaluated, the gaze-directed approach proved to be an effective protocol to guide the attention of novices, and improve their performance on a specific laparoscopic task (Wilson et al., 2011a, b). If one considers the dichotomy of expert to novice eye movements in laparoscopy as a parallel to that of HMRA and LMRA individuals, is it then possible that gaze training may serve direct LMRA individuals toward improvements on an array of spatially complex tasks?

If a fundamental difference exists in how high and low individuals examine images, and presumably how they process the stimuli, this finding could have broader implications for disciplines requiring an understanding of complex visual imagery, such as anatomy.

CONCLUSIONS

It appears that some individuals who lack the ability to reason spatially may be handicapped to a degree by complex visual imagery, and may require additional direction toward key regions of images to ensure understanding and minimize distraction. This concept is supported widely by the cognitive theory of multimedia learning (Mayer and Moreno, 2003), adapted from Cognitive Load Theory (Sweller, 1988, 1989; Kirschner, 2002). Sweller's model of Cognitive Load Theory suggests that learning is facilitated by instruction that directs a human's finite working memory resources toward information that is important to learning, rather than to extraneous material that impedes, or slows cognitive processing, and distracts from learning (Sweller, 1988, 1989).

Similarly, Mayer and Moreno's cognitive theory of multimedia learning suggests that individuals have a finite allotment for interpreting visual stimuli, and that this allotment may be exceeded by poorly designed or complex, detailed images, including those with spatial complexity (Mayer and Moreno, 2003; Wilson, 2015). In instances where the images cannot be simplified (such as in anatomical imagery, or in spatially complex test images) the cognitive theory for multimedia learning recommends that "cognitive load can be reduced by providing cues to the learner about how to select and organize the material—a technique called signaling" where visual indicators are used to identify features of complex images or content that are particularly informative, or salient (Meyer, 1975; Lorch, 1989), and that "signaling seems to help in the process of selecting and organizing relevant information" (Mayer and Moreno, 2003). With this in mind, it can be hypothesized that visual signaling through gaze directed to salient areas *within* images may mitigate deficiencies in spatial ability, reduce cognitive load, and thus increase learning across all levels of learners except those with the highest spatial ability.

With the concept of signaling, and gaze-directed training in mind, future studies should seek to construct a method of visual cueing that is based on the eye movement patterns of HMRA individuals. Such an intervention could in theory, direct or signal, the attention of LMRA individuals to task-relevant locations on images, and potentially improve their performance in spatial reasoning, and consequently, in spatially complex fields such as anatomy and other science, technology, engineering and mathematics (STEM) disciplines as a whole.

NOTES ON CONTRIBUTORS

VICTORIA A. ROACH, Ph.D., is an assistant professor in the Department of Biomedical Sciences at Oakland University William Beaumont School of Medicine, in Rochester, Michigan. She teaches medical gross anatomy and is investigating the role that spatial reasoning plays in the acquisition of anatomical and surgical skill.

GRAHAM M. FRASER, Ph.D., is an adjunct assistant professor in the Department of Medical Biophysics at Schulich School of Medicine and Dentistry, The University of

Western Ontario, London, Ontario, Canada. He teaches biophysics, problem solving, and gross anatomy and his research interest resides in microvascular transport and regulation.

JAMES H. KRYKLYWY, M.Sc., is a doctoral candidate in the Department of Psychiatry, at the Brain and Mind Institute, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada. He investigates how emotions affect perception and attention using MRI, eye-tracking, and audio simulated environments.

DEREK G.V. MITCHELL, Ph.D., is an assistant professor in the Departments of Psychiatry, and Anatomy and Cell Biology and a member of the Brain and Mind Institute at Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada. His research is principally aimed at determining how dissociable neural systems integrate emotion with cognition and behavior. His research also aims to determine how the human brain bias attention toward the most behaviorally relevant information.

TIMOTHY D. WILSON, Ph.D., is an associate professor in the Department of Anatomy and Cell Biology at The University of Western Ontario, London, Ontario, Canada. He founded and directs the Corps for Research of Instructional and Perceptual Technologies (CRIPT) laboratory investigating digital anatomy development, deployment, and efficacy in pedagogy. His research aims to determine how attention, spatial ability, and student success are linked.

LITERATURE CITED

Booth JR, MacWhinney B, Thulborn KR, Sacco K, Voyvodic JT, Feldman HM. 2000. Developmental and lesion effects in brain activation during sentence comprehension and mental rotation. Dev Neuropsychol 18:139–169.

Brandt MG, Davies ET. 2006. Visual-spatial ability, learning modality and surgical knot tying. Can J Surg 49:412–416.

Carpenter RHS. 1988. Movements of the Eyes. 2nd Ed. London, UK: Pion Ltd. 593 p.

Chun MM, Jiang Y. 1998. Contextual cueing: Implicit learning and memory of visual context guides spatial attention. Cognit Psychol 36:28–71.

Cohen MS, Kosslyn SM, Breiter HC, DiGirolamo GJ, Thompson WL, Anderson AK, Brookheimer SY, Rosen BR, Belliveau JW. 1996. Changes in cortical activity during mental rotation: A mapping study using functional MRI. Brain 119:89–100.

Ekstrom RB, French JW, Harman HH, Dermen D. 1976. Manual for Kit Factor-Referenced Cognitive Tests. 1st Ed. Princeton, NJ: Educational Testing Service. 224 p.

Engle RW, Tuholski SW, Laughlin JE, Conway ARA. 1999. Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. J Exp Psychol Gen 128:309–331.

Farah MJ. 1989. The neural basis of mental imagery. Trends Neurosci 12:395–399.

Fechner GT. 1860/1966. Elements of Psychophysic. Volume 1. In: Howes DH, Boring EG (Editors). Elements of Psychophysic. 1st Ed. New York, NY: Holt, Rinehart, and Winston Inc. 286 p.

Fisher RA. 1922. On the interpretation of $\chi 2$ from contingency tables, and the calculation of P. J Roy Stat Soc 85:87–94.

Fisher RA. 1932. Statistical Methods for Research Workers. 4th Ed. Edinburgh, UK: Oliver and Boyd. 307 p.

Gages TT. 1994. The Interrelationship Among Spatial Ability, Strategy Used, and Learning Style for Visualization Problems. 1st Ed. Columbus, OH: Ohio State University. 302 p.

Geiser C, Lehmann W, Eid M. 2006. Separating "rotators" from "nonrotators" in the mental rotations test: A multigroup latent class analysis. Multivariate Behav Res 41:261–293.

Glück J, Fabrizii C. 2010. Gender differences in the mental rotations test are partly explained by response format. J Indiv Differ 31:106–109.

Harris IM, Miniussi C. 2003. Parietal lobe contribution to mental rotation demonstrated with rTMS. J Cogn Neurosci 15:315–323.

Howell DC. 2007. Statistical Methods for Psychology. 6th Ed. Belmont, CA: Wadsworth Publishing. 739 p.

Itti L, Koch C, Niebur E. 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell 20:1254–1259.

Just MA, Carpenter PA. 1976. Eye fixations and cognitive processes. Cogn Psychol 8:441–480.

Just MA, Carpenter PA. 1985. Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. Psychol Rev 92:137–172.

Kaakinen JK, Hyönä J, Viljanen M. 2011. Influence of a psychological perspective on scene viewing and memory for scenes. Q J Exp Psychol 64:1372–1387.

Kirschner PA. 2002. Cognitive load theory: Implications of cognitive load theory on the design of learning. Learn Instruct 12:1−12.

Kozhevnikov M, Motes MA, Hegarty M. 2007. Spatial visualization in physics problem solving. Cognit Sci 31:549–579.

Kyllonen PC. 1996. Is working memory capacity Spearman's g? In: Dennis I, Tapsfield P (Editors). Human Abilities: Their Nature and Measurement. 1st Ed. Mahwah, NJ. Lawrence Erlbaum Associates Inc. p 49–75.

Kyllonen PC, Christal RE. 1990. Reasoning ability is (little more than) working-memory capacity?!. Intelligence 14:389–433.

Law B, Atkins MS, Kirkpatrick AE, Lomax AJ, Mackenzie CL. 2004. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In: Duchowski AT, Vertegaal R (Editors). Proceedings of the 2004 Symposium on Eye Tracking Research and Application (ETRA 2004); San Antonio, TX, 2004 March 22–24. p 41–47. Association for Computing Machinery, New York, NY.

Lee WF, Huang TH, Yeh SL, Chen HH. 2011. Learning-based prediction of visual attention for video signals. IEEE Trans Image Process 20:3028–3038.

Linn MC, Petersen AC. 1985. Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Dev 56:1479–1498.

Lohman DF. 1988. Spatial abilities as traits, processes, and knowledge. In: Sternberg RJ (Editor). Advances in the Psychology of Human Intelligence. Volume 4. 1st Ed. Hillsdale, NJ: Lawrence Erlbaum Associates. p 181–248.

Lohman DF. 1996. Spatial Ability and g. In: Dennis I, Tapsfield P (Editors). Human Abilities: Their Nature and Measurement. 1st Ed. Mahwah, NJ. Lawrence Erlbaum Associates Inc. p 97–116.

Lorch RF. Jr. 1989. Text-signaling devices and their effects on reading and memory processes. Educ Psychol Rev 1:209–234.

Lowe R, Schnotz W (Editors). 2008. Learning with Animation: Research Implications for Design. 1st Ed. Cambridge, NY: Cambridge University Press. 402 p.

Lufler RS, Zumwalt AC, Romney CA, Hoagland TM. 2012. Effect of visual–spatial ability on medical students' performance in a gross anatomy course. Anat Sci Educ 5:3–9.

Lunneborg PW. 1982. Sex differences in self-assessed everyday spatial abilities. Percept Mot Skills 55:200–202.

Lunneborg CE, Lunneborg PW. 1984. Contribution of sex-differentiated experiences to spatial and mechanical reasoning abilities. Percept Mot Skills 59: 107–113.

Mayer RE, Moreno R. 2003. Nine ways to reduce cognitive load in multimedia learning. Educ Psychol 38:43–52.

Mayer RE, Sims VK. 1994. For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. J Educ Psychol 86: 389–401.

McDonald JH. 2014. Fisher's exact test of independence. In: McDonald JH (Editor). Handbook of Biological Statistics. 3rd Ed. Baltimore, MD: Sparky House Publishing. p 77–85.

McHugh ML. 2012. Interrater reliability: The kappa statistic. Biochem Med 22:276–282.

Meyer BJF. 1975. The Organization of Prose and its Effects on Memory: North-Holland Studies in Theoretical Poetics. Volumne 1. 1st Ed. Amsterdam, The Netherlands: North-Holland American Elsevier Publishing Co. 249 p.

Miyake A, Friedman NP, Rettinger DA, Shah P, Hegarty M. 2001. How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. J Exp Psychol Gen 130:621–640.

Nguyen N, Mulla A, Nelson AJ, Wilson TD. 2014. Visuospatial anatomy comprehension: The role of spatial visualization ability and problem-solving strategies. Anat Sci Educ 7:280–288.

Nguyen N, Nelson AJ, Wilson TD. 2012. Computer visualizations: Factors that influence spatial anatomy comprehension. Anat Sci Educ 5:98–108.

Oberauer K, Süß HM, Schulze R, Wilhelm O, Wittmann WW. 2000. Working memory capacity — Facets of a cognitive ability construct. Pers Indiv Differ 29:1017–1045.

Peters M, Battista C. 2008. Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. Brain Cogn 66:260–264.

Ratcliff G. 1979. Spatial thought, mental rotation and the right cerebral hemisphere. Neuropsychologia 17:49–54.

Roach VA, Brandt MG, Moore CC, Wilson TD. 2012. Is three-dimensional videography the cutting edge of surgical skill acquisition?. Anat Sci Educ 5: 138–145.

Roach VA, Fraser GM, Kryklyw JH, Mitchell DG, Wilson TD. 2016. The eye of the beholder: Can patterns in eye movement reveal aptitudes for spatial reasoning?. Anat Sci Educ 9:357–366.

Schmider E, Zeigler M, Danay E, Beyer L, Bühner M. 2010. Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. Meth Eur J Res Meth Behav Soc Sci 6:147–151.

Shepard RN, Cooper LA. 1986. Mental Images and Their Transformations. 1st Ed. Cambridge, MA: The MIT Press. 376 p.

Shepard RN, Metzler J. 1971. Mental rotation of three-dimensional objects. Science 171:701–703.

Sweller J. 1988. Cognitive load during problem solving: Effects on learning. Cognit. Sci 12:257–285.

Sweller J. 1989. Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. J Educ Psychol 81:457–466.

Vandenberg SG, Kuse AR. 1978. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills 47:599–604.

Voyer D, Voyer S, Bryden MP. 1995. Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. Psychol Bull 117:250–270.

Wanzel KR, Hamstra SJ, Anastakis DJ, Matsumoto ED, Cusimano MD. 2002. Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet 359:230–231.

Wanzel KR, Hamstra SJ, Caminiti MF, Anastakis DJ, Grober ED, Reznick RK. 2003. Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. Surgery 134:750–757.

Wilhelm O, Schulze R. 2002. The relation of speeded and unspeeded reasoning with mental speed. Intelligence 30:537–554.

Wilson TD. 2015. Role of image and cognitive load in anatomical multimedia. In: Chan LK, Pawlina W (Editors). Teaching Anatomy: A Practical Guide. 1st Ed. New York, NY: Springer International Publishing. p 237–246.

Wilson M, McGrath J, Vine S, Brewer J, Defriend D, Masters R. 2010. Psychomotor control in a virtual laparoscopic surgery training environment: Gaze control parameters differentiate novices from experts. Surg Endosc 24:2458–2464.

Wilson MR, McGrath JS, Vine SJ, Brewer J, Defriend D, Masters RS. 2011a. Perceptual impairment and psychomotor control in virtual laparoscopic surgery. Surg Endosc 25:2268–2274.

Wilson MR, Vine SJ, Bright E, Masters RS, Defriend D, McGrath JS. 2011b. Gaze training enhances laparoscopic technical skill acquisition and multi-tasking performance: A randomized, controlled study. Surg Endosc 25:3731–3739.

Zacks JM. 2008. Neuroimaging studies of mental rotation: A meta-analysis and review. J Cogn Neurosci 20:1–19.

Zumwalt AC, Iyer A, Ghebremichael A, Frustace BS, Flannery S. 2015. Gaze patterns of gross anatomy students change with classroom learning. Anat Sci Educ 8:230–241.