Available online at www.sciencedirect.com
ScienceDirect
Intelligence 35 (2007) 211 – 223
INTELLIGENCE

Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity?

Scott Barry Kaufman *

University of Cambridge, England and Yale University, United States

Received 11 July 2005; received in revised form 30 June 2006; accepted 15 July 2006
Available online 11 September 2006

Abstract

Sex differences in spatial ability are well documented, but poorly understood. In order to see whether working memory is an important factor in these differences, 50 males and 50 females performed tests of three-dimensional mental rotation and spatial visualization, along with tests of spatial and verbal working memory. Substantial differences were found on all spatial ability and spatial working memory tests (that included both a spatial and verbal processing component). No significant differences were found in spatial short-term memory or verbal working memory. In addition, spatial working memory completely mediated the relationship between sex and spatial ability, but there was also a direct effect of sex on the unique variance in three-dimensional rotation ability, and this effect was not mediated by spatial working memory. Results are discussed in the context of research on working memory and intelligence in general, and sex differences in spatial ability more specifically.

Keywords: Intelligence; Working memory; Spatial working memory; Verbal working memory; Sex differences; Spatial visualization; Mental rotation

1. Introduction

Spatial abilities have long been thought to be an important component of intelligence. One of Thurstone's primary mental abilities was spatial visualization (Thurstone, 1938). Vernon postulated two broad factors below *g*, a verbal-educational and a spatial-mechanical factor (Vernon, 1971). Horn's expansion of the original *gf*–*gc* theory placed visual thinking in the second stratum (Horn, 1994). Carroll's extensive analysis led him to place a broad visual perception factor in Stratum II of his model, with more specific types of spatial abilities listed in Stratum I (Carroll, 1993).

Research has also confirmed Carroll's view that spatial ability is not a unitary process and can be decomposed into various distinct forms. Factor analytic studies of spatial ability tasks often point to two distinct spatial abilities, *visualization* and *orientation* (Hegarty & Waller, 2004; McGee, 1979). Visualization refers to the ability to mentally rotate and manipulate objects while orientation refers to the ability to retain spatial orientation with respect to one's body.

Other researchers have distinguished between mental rotation ability and spatial visualization ability. Linn and Peterson (1985) and Voyer, Voyer, and Bryden (1995) distinguished three categories of spatial ability based on the differing processes required to solve problems

* Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, United States.
E-mail address: scott.kaufman@yale.edu.

0160-2896/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2006.07.009

representative of each ability. The three categories they identified were *spatial perception*, *mental rotation*, and *spatial visualization*. Examples of tests in each category are (a) the water level test (Inhelder & Piaget, 1958), (b) the MRT (Vandenberg & Kruse, 1978), and (c) the space relations subtest of the DAT (The Psychological Corporation, 1995), respectively.

The importance of differentiating mental rotation ability from other types of spatial ability has been argued quite recently by Johnshon and Bouchard (2005). They posit that the fluid and crystallized intelligence stratum should be replaced with a stratum consisting of verbal, perceptual, and image rotation ability. Based on an analysis of 436 individuals who completed 42 mental ability tests, the authors compared the models of Vernon, Cattell, and Carroll using maximum likelihood confirmatory factor analysis. It was found that Vernon's model fit the data the best, with the need to extend Vernon's model by adding memory and higher-order image rotation factors to the model. The authors concluded that "the visualization processes involved in mental image rotation tasks have not been given the attention they deserve as important and relatively independent contributors to the manifestation of human intelligence" (p. 17).

That spatial ability can be fractionated is consistent with the large literature revealing sex differences on certain types of spatial tasks [note that there are no systematic differences in overall intelligence among males and females (see Loehlin, 2000)]. A meta-analysis of studies published before 1973 found an average difference of 0.45*d* in favor of males on tests of visuo-spatial ability (Hyde, 1981). Further analyses showed that the size of the sex difference varies considerably across different kinds of tests. Linn and Peterson (1985) and Voyer et al. (1995) conducted a large scale meta-analyses of over 50 years of research and reported significant differences favoring males in tasks requiring mental rotation and manipulation of mental images. Differences existed but were not as strong for spatial perception and were even smaller for spatial visualization. Masters and Sanders (1993) report a difference of 0.9*d* between males and females on tests of three-dimensional mental rotation ability. Even though some sex differences have seemed to decline from the years 1945–1995 (Voyer et al., 1995), there has been very little decrease in the size of the sex difference on mental rotation tasks during that time (Masters & Sanders, 1993; Voyer et al., 1995). On the 3D versions of mental rotation tasks, the difference is still close to one standard deviation (e.g., Mackintosh & Bennett, 2005).

Various distal hypotheses, biological (Kimura & Hampson, 1992; Lynn, 1994; Plomin, DeFries, & McClearn, 1990), environmental (Astin, Sax, Korn, & Mahoney, 1995; Harris, 1995; Lytton & Romney, 1991), and psychobiological (Halpern, 1997; Halpern & LaMay, 2000), have been put forward to explain sex differences in spatial ability. However, "research on sex differences has done remarkably little to elucidate the nature of the differences between various kinds of tests, or the differences in the psychological operations they engage" (Mackintosh, 1998, p. 191).

What could be a candidate proximal process that people use to solve the types of spatial tasks that demonstrate a sex difference? Solving a mental rotation task such as the Shepard-Metzler 3D rotation task or a spatial visualization task such as the space relations subtest of the Differential Aptitude Test (DAT) requires the ability to maintain an active representation of all the parts, and the interrelations of all the parts, while simultaneously rotating the image in the mind. This elaboration, involving both storage (holding the constituent parts in memory) and the simultaneous processing of spatial representations (the rotation component), fits closely with current conceptions of working memory (Miyake & Shah, 1999).

There is evidence to suggest that there is indeed a relationship between working memory (WM) capacity and mental rotation ability. One study conducted an analysis of performance on the cube comparison task (Just & Carpenter, 1985). Low and high spatial ability participants (all from a university) had to judge whether two views of children's alphabet blocks could represent the same block. Eye fixation analysis showed that low spatial ability participants sometimes had to rotate a particular cube face more than once, as though they had forgotten an intermediate representation, whereas high spatial ability participants rarely had to rotate the same face more than once. They also analyzed performance on the Shepard-Metzler Mental Rotation Task. They found that those performing poorly on the task rotated the figures at a slower rate than those performing well (see Lohman, 1986 for similar findings regarding sex differences). In addition, the low spatial group had more difficulty keeping track of their intermediate products, resulting in reinitializations of various processes. From these studies, Carpenter and Just concluded that "A general characterisation…is that low spatial subjects have difficulty maintaining a spatial representation while performing transformations" (Carpenter & Just, 1986, p. 236).

There is also experimental evidence for the link between WM capacity and spatial visualization ability. Salthouse, Babcock, Mitchell, Palmon, and Skovronek (1990) found that spatial visualization differences were more pronounced when some information was preserved while the same or other information was being processed. They further showed that individual differences in spatial visualization ability could not be attributed to variations in representational quality and transformational efficiency. It should be noted, however, that their sample consisted of 50 male students at a technical university, and their results may not generalize to a wider population.

Another study adds additional support to the link between spatial visualization ability and working memory (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). They showed that their spatial visualization factor (which consisted of the paper folding and DAT space relations test) had the highest degree of executive involvement (i.e., correlated highly with the Tower of Hanoi and Random Number Generation tests) and was a better predictor of performance on their spatial WM tasks (Letter Rotation and Dot Matrix tasks) than the two other spatial factors (spatial relations and visuo-spatial perceptual speed). It is important to note that Miyake et al. did not include a test of three-dimensional mental rotation ability in their study.

Taken together, these studies suggest that individual differences in mental rotation and spatial visualization ability may be accounted for, at least in part, by differences in working memory capacity.

There is, however, virtually no study that has attempted to use the current definition of working memory to elucidate sex differences in spatial ability. A handful of studies have found sex differences in spatial WM (Duff & Hampson, 2001; Geiger & Litwiller, 2005; Vecchi & Girelli, 1998), but they did not assess the relationship between WM and spatial ability.

Even though the experimental design used by Shah and Miyake (1996) would have lent itself nicely to an investigation of sex differences in spatial working memory and its relationship to spatial ability, they admit in a footnote (p. 8) that they did not collect any sex information so therefore could not do the proper analysis. The study conducted by Vecchi and Girelli (1998) would also have been relevant if they had included various tests of spatial ability, and had a larger sample of male and female participants. Lastly, even though Duff and Hampson (2001) used both spatial working memory and mental rotation tasks, they did not try to relate one to the other, and did not include a spatial working memory task that consisted of an explicit storage and processing component.

A major aim of the current study was to assess differences between males and females on various tests of spatial working memory and spatial ability in order to determine whether spatial working memory capacity is a critical factor in determining the male–female difference in mental rotation and spatial visualization ability. To assess this, multiple tests of spatial working memory were administered along with multiple tests of spatial ability. In addition, an increased working memory load modification was made to a well known test of spatial visualization to see if an increased working memory load on such a test would increase any sex difference.

A second goal of the study was to assess the particular characteristics of memory tasks that give rise to sex differences in spatial ability. By including a *verbal working memory* test (which consisted of a verbal storage and verbal processing component), a *spatial short term memory* task (which consisted of a spatial storage component and no processing component), and *two spatial working memory* tests (one test with a spatial processing component and the other with a verbal processing component), the experiment was able to assess the relationship between the characteristics of a working memory task and spatial ability.

The design also allowed for an assessment of the domain specificity of working memory. There is a whole research tradition demonstrating the extremely high correlations between the variance shared across a battery of working memory tests and a general factor of intelligence (Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Kane et al., 2004; Kyllonen, 1996). Therefore, at the most general level of analysis, there appears to be a strong relationship between executive functioning and general intelligence. However, there is also increasing evidence to suggest that working memory is not a unitary construct. Studies using an interference paradigm have shown that concurrent hand movement impairs a participant's performance on a visuo-spatial WM task, but not on a verbal WM task, whereas concurrent articulation impairs performance on a verbal span task but not a visuo-spatial task (e.g., Baddeley & Lieberman, 1980; Farmer, Berman, & Fletcher, 1986; Jurden, 1995; Logie, 1995). More recent studies have found that while performance on a verbal working WM task predicts verbal IQ it does not predict spatial IQ, while the reverse is true for performance on a spatial WM task (Mackintosh & Bennett, 2003; Shah & Miyake, 1996). Brain studies also lend support to the suggestion of separation by showing different patterns of activation when people are engaged in verbal WM (Petrides, Alivisatos, Evans, & Meyer, 1993), visual WM (Smith et al., 1995), and spatial WM tasks (Jonides et al., 1993). Therefore, the current study is also in a position to contribute to the growing literature on the domain specificity of working memory.

2. Methods

2.1. Participants

The 100 participants (50 males and 50 females), aged 16–18, were students taken from a slightly selective sixth form college (which takes in students who are in their last 2 years of secondary education) in Cambridge, United Kingdom. Admittance to the sixth form college is contingent upon adequate GCSE scores (equivalent to American SAT II subject test scores) and an overall B or better average in classes pertaining to the student's course of study.

An advertisement to participate in the study was sent to all members of college. The recruitment flyer mentioned that Cambridge University was looking for participants for a psychology experiment that would involve solving fun puzzles. There was no mention of spatial abilities in the flyer, and the sixth form college from which participants were recruited have roughly equal numbers of males and females. Therefore, to the best of the experimenter's control, there was no selection bias in recruitment of participants.

2.2. Procedure

Tests were administered in groups at desktop terminals during the course of two 1-h sessions in the following order:

Session 1
- DAT space relations test (standard version)
- DAT space relations test (increased working memory load version)
- Verbal working memory test

Session 2
- MRT-A
- DAT verbal reasoning test
- Simple block span
- Verification-block span
- Rotation-block span

The software for the experiment was programmed in Visual Basic and presented in a Windows 2000 environment. Each participant was required to participate in two testing sessions to earn £10. All 100 participants completed all the tests, except for the DAT verbal reasoning test ($n=72$, 36 males, 36 females), which began administration once the first few group of participants had already been tested.

2.3. Ability tests

2.3.1. DAT space relations test (The Psychological Corporation, 1995)

The DAT was designed for use with 8th through 12th graders for purposes of educational and vocational counselling (Bennett, Seahore, & Wesman, 1974). The DAT-SR test measures the ability to visualize a three-dimensional object from a two-dimensional pattern and to visualize how this object would look if rotated in space. Each problem shows one pattern, followed by four three-dimensional figures. Test takers are to choose the one figure that can be made from the pattern. Prior research demonstrates that the DAT-SR test loads primarily on the spatial visualization factor (Carroll, 1993); therefore, the test was used as a measure of *gv*. This test has demonstrated small sex differences but has not demonstrated as marked a sex difference as that found on the MRT (e.g., Linn & Peterson, 1985; Voyer et al., 1995).

A major purpose of the current study was to see whether sex differences in working memory capacity contribute to differences in spatial ability. As one test of this hypothesis, two versions of the DAT-SR test were presented to each participant: a standard version and one with an increased working memory load. If a difference in spatial working memory capacity is the driving force behind differences found on tests of mental rotation, then an increased working memory load manipulation to the DAT-SR would be expected to demonstrate a larger sex difference than that found on a standard version.

2.3.1.1. Standard version of the DAT-SR. This version consisted of 25 odd-numbered items taken from the original 50-item DAT-SR test booklet. A further change was introduced: based on data collected in a pilot study of 10 participants, the least selected answer option was deleted from each item, making the number of answer options for each item three instead of the usual four. This was done to keep the number of answer choices constant between the two versions of the test. A pilot study demonstrated that three answer options was the most appropriate number, since four made the second version of the test too hard (and two made it too easy). This first version of the test had a 15-min time limit.

2.3.1.2. DAT-SR with working memory manipulation. The 25 even numbered items of the original DAT-SR test were used for the second part of this test. This introduced an increased working memory load, by separating the target from the answer options. Participants were presented with the question target for 7 s, after which the target disappeared and participants were required to select the correct answer. The assumption was that this would increase working memory load, as the item would have to be stored in memory while each answer choice was being processed and considered. Participants had unlimited time to select an answer in this part of the test.

2.3.2. MRT-A mental rotations test

Set A of the MRT (Vandenberg & Kruse, 1978) which contains 24 problems was used as a measure of mental rotation ability. This test, which has consistently demonstrated the largest sex difference (Linn & Peterson, 1985), has also been suggested to be more difficult and involve different processes than mental rotation tasks in two dimensions (e.g., Rosser, 1980).

Each problem in the MRT-A shows a three-dimensional target figure paired with four choice figures, two of which are rotated versions of the target figure. To score a point, both correct answers must be given. After two practice items with feedback and an explanation, the first 12 problems were attempted in 4 min with a 2-min break before attempting the second 12 in another 4 min. The maximum score is 24.

2.3.3. DAT verbal reasoning test

The verbal reasoning section of the Differential Aptitudes Test was administered (The Psychological Corporation, 1995). This test was used to make sure the males and females in a sample were matched in general verbal reasoning ability.

Each problem had two words missing, and participants had to choose a pair of words from the answer stem that were related to the words in the sentence in some way. After two practice items, participants had 15 min to complete 40 problems. The maximum score is 40.

2.4. Working memory tests

2.4.1. Verbal working memory

For group testing, a technique similar to Baddeley, Logie, Nimmo-Smith, and Brereton's (1985) adaptation of Daneman and Carpenter's (1980) Reading Span Test was used.

12 groups of short sentences were presented on the screen. Each sentence contained a subject, verb, and object, e.g., "The teacher ate the orange." At the end of each group of sentences, participants were asked to recall either the subject of each sentence or the object of each in strict order. Half the sentences made sense and half were nonsensical, e.g., "The tomato kicked the ball." To ensure that participants were processing each entire sentence, they were asked, after the presentation of each sentence, to say whether or not it made sense. Therefore, a typical sequence of length two sentences would be:

1. "The book sang a song" (the participant would respond by clicking on the button "nonsense").
2. "The man ate the curry" (the participant would respond by clicking on the button "sense").

A screen would then follow, in which the participant would be asked to write the *object* of each sentence.

In this example, the correct answer would be "song" and "curry".

Participants were given one group of two sentences and one group of three sentences for practice. Two of these five sentences made sense, the remaining three were nonsense. These practice data were discarded.

The test successively comprised five groups of three sentences, five groups of four sentences and two groups of five sentences (23 sense, 22 nonsense randomized). Answer sheets clearly delineated the increase in length of groups, and participants had also been told of this increase.

After each group, they were given 15 (three sentences), 20 (four sentences), and 25 s (five sentences) to type the answers in the answer box. Performance was assessed in terms of total number of items recalled in the correct serial position, giving a maximum score of 45.

2.4.2. Spatial span tests

2.4.2.1. Simple block span. The *simple block span* was a modification of the *simple arrow span* task designed by Miyake and Shah (1999). The major difference was the replacement of arrows with Corsi blocks. This was thought to be more appropriate since the Corsi blocks task has been widely adopted in neuropsychological practice and in other areas of psychology (De Renzi & Nichelli, 1975; Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000), and has been claimed to be one of the few "pure" measures of the visuo-spatial sketchpad (Milner, 1971).

In the *simple block span*, test participants were required to recall the location of a series of blocks in the order in which they were presented.

Blocks were unevenly spaced. Participants received three trials of this type with a set size of 5, and three with a set size of 6. Participants were scored on the number of blocks they remembered correctly in the order in which they were presented, giving a maximum score of 33.

The *simple block span test* was used as a measure of spatial short-term memory, which has been shown to have different predictive powers from spatial working memory tests containing both a storage and processing component (e.g., Kane et al., 2004).

2.4.2.2. Complex spatial span measures. The verbal and spatial processing components were added to the storage component to assess the participant's capacity to simultaneously process and store information and thus make them actual tests of working memory capacity. In line with the Shah and Miyake (1996) study, I will refer to the spatial storage plus verbal processing task described below as the *verification-block span* test, and the spatial storage plus mental rotation processing task as the *rotation-block span* test.

2.4.2.2.1. Verification-block span. For the *verification-block span* test, participants were required to memorize the location of blocks while simultaneously processing verbal information. Ninety-eight sentences were prepared. The same set of sentences was initially randomized and presented in the same fixed order. There were three trials with set size 5 and three with a set size of 6. On each trial, the participant was shown a block location for approximately 1 s, and then a screen appeared with a sentence. The participant was asked to decide whether the sentence was sensible or not by pressing an appropriate button (either SENSE or NONSENSE) on a response box.

The reaction time and accuracy of the response was recorded by the computer. The sentence remained on the screen for a maximum of 2200 ms if no response was made. All participants were instructed to answer this secondary verbal task as quickly as possible. This maximum time limit was introduced in order to minimize the possibility that participants would keep rehearsing the location of the blocks. Approximately 50 ms after the participant's response, the next block location to be remembered appeared on the screen. Then, the next sentence appeared on the screen, and the same cycle was repeated until the end of the trial, when the recall grid appeared on the screen requesting the participant to select all the block locations that were presented during that trial in the order in which they were presented.

2.4.2.2.2. Rotation-block span. The *rotation-block span* test required the participant to simultaneously remember the location of blocks while judging whether each letter presented on the computer screen was normal or mirror-imaged.

A single trial consisted of the individual presentation of a small set of normal or mirror-image letters (F, J, L, P, and R). The same letter was presented within a set and presented in one of seven possible orientations (in 45° increments, not including upright). Each of the 70 possible combinations (letters × orientations × normal/mirror-image status) appeared only once in the task. The presentation of letters was also constrained such that the same orientation could appear only once in the same set. The letter remained on the screen for a maximum of 2200 ms if no response was made. Half of the letters were normal, and half were mirror-images. The participant indicated his or her judgment by pressing an appropriate button (either NORMAL or MIRROR) on a response box. The response time and accuracy of the response were recorded by the computer. The timing and methodology was the same as that of the *verification-block span* test. At the end of each trial, the recall grid appeared on the screen with instructions to recall the to-be-remembered blocks in the correct order in which they were presented. As in the *verification-block span* test, there were three trials each of five and six letters, presented in the order of increasing set size.

2.4.2.2.3. Practice trials. Prior to the administration of each of the three spatial span tasks, the participant received two practice trials with a set size of 2 that included exactly the same stimuli and format as the real task. Participants received feedback regarding their responses and were instructed to ask the experimenter any questions if the task was still not clear after going through the practice trials. When participants felt comfortable with the demands of the task, they could click a button to begin the trials.

3. Results

Table 1 lists the correlation matrix for all the spatial ability and span tests. DAT-SR reflects the average score of both versions (with and without the working memory manipulation) of the DAT space relations test. We collapsed the scores for three reasons. First, the increased working memory load version of the DAT space relations task was equally difficult for males and females. A repeated measures ANOVA with sex as the between-subjects factor and both versions of the DAT space relations test as the within-subjects factor revealed a main effect of DAT ($p<.001$), but the interaction was not significant ($p=.802$). Second, the increased working memory load version of the DAT space relations test did *not* display a higher correlation with any of the spatial working memory or ability tasks than did the

Table 1
Correlation matrix of all spatial ability and span tests ($N=100$)

Measure	1	2	3	4	5	6
1. DAT-SR	–	**.606	**.514	**.345	**.395	*.256
2. MRT	–	–	**.466	**.441	**.467	*.215
3. Simple block span	–	–	–	**.499	**.515	.117
4. Rotation-block span	–	–	–	–	**.625	.172
5. Verification-block span	–	–	–	–	–	.109
6. Verbal WM	–	–	–	–	–	–

*$p<.05$ (two-tailed), **$p<.01$ (two-tailed).

Table 2
Summary of descriptive statistics for sex differences

Measure	Female		Male		Difference	
	M	S.D.	*M*	S.D.	*t*	*d*
Ability measures						
MRT (out of 24)	12.08	5.00	17.28	3.83	***5.84	1.01
DAT-SR (out of 25)	14.88	4.19	16.58	3.78	*2.13	0.42
DAT-verbal (out of 40)	22.03	6.58	23.78	6.21	1.16	0.27
Spatial short-term memory measure						
Simple block span (out of 33)	27.34	4.42	28.26	3.91	1.10	0.22
Working memory measures						
Rotation-block span (out of 33)	15.14	6.65	19.84	7.17	**3.40	0.65
Verification-block span (out of 33)	18.54	5.57	22.20	5.90	**3.19	0.58
Verbal working memory (out of 45)	19.36	6.64	18.82	7.35	0.39	0.08

$N=50$ except for *DAT verbal reasoning test* (male=36, female=36), *$p<.05$, **$p<.01$, ***$p<.0001$.

standard version. Third, the two scores were significantly correlated with each other ($r=.617$, $p<.01$). Therefore, for parsimony, both versions were collapsed into a single score.

Table 1 reveals that all the spatial ability and spatial working memory tests were significantly correlated with each other. In addition, the verbal working memory test correlated with the spatial ability tests, but not the spatial working memory tests. Appendix A lists the correlation matrices for males and females listed separately. Even though at first sight it looks as though the two matrices are different, with an *N* of only 50 for each sex there was only one significant correlation difference between males and females (listed in bold and further discussed in Discussion). Apart from this one difference, the pattern of correlations between males and females is not significantly different from each other.

Table 2 shows the *mean* score and *standard deviation* for each test broken down by sex, and the *t* value and *effect size* for the difference between males and females.

On the ability measures, males scored significantly better than females on both tests of spatial ability, but there was no difference in general verbal reasoning. The spatial ability differences are consistent with the meta-analyses of Linn and Peterson (1985) and Voyer et al. (1995) showing male superiority on mental rotation and spatial visualization tests, with the difference much more pronounced on tests of mental rotation. That there was no significant different on the test of verbal reasoning suggests that all discussion of sex differences in this paper can be restricted to the spatial domain.

On the memory measures, there was no statistically significant difference between males and females on the simple block span test and verbal working memory test. However, both tests of spatial working memory (with both a spatial and verbal processing component) displayed a statistically significant difference (Fig. 1). To assess the separability between the simple spatial span and complex spatial span measures, a multivariate analysis of covariance (MANCOVA) was run looking at the difference between males and females on the complex span tests holding simple block span constant. Even with simple block span held constant, the difference on both *rotation-block span* and *verification-block span* remained significant [$F(1,97)=10.52$, $p<.01$, $F(1,97)=9.13$, $p<.01$, respectively]. This suggests at least some separability between the spatial short-term memory and spatial working memory tasks.

Taken together, the results of the memory measures suggest that males and females differ on a test of spatial working memory, with the nature of the processing component making little difference. Therefore, since males and females do not differ on a verbal working memory test (that consists of both a verbal storage and processing component), or a test that just has a spatial storage component, the sex difference in working memory seems to be limited to complex span measures with the defining characteristic being the spatial storage component.

It is important to note that there was no indication of participants' trading processing accuracy for better span scores. The correlations between the processing component and the storage component for all three working memory tasks were positive (verbal working memory, $r(100)=.269$, $p<.01$; rotation-block span, $r(100)=.266$, $p<.01$; verification-block span, $r(100)=.024$, *ns*). This result points to the absence of a strategic trade-off between processing and recall accuracy, since any trade-off between processing and storage would have implied a negative correlation between the two.

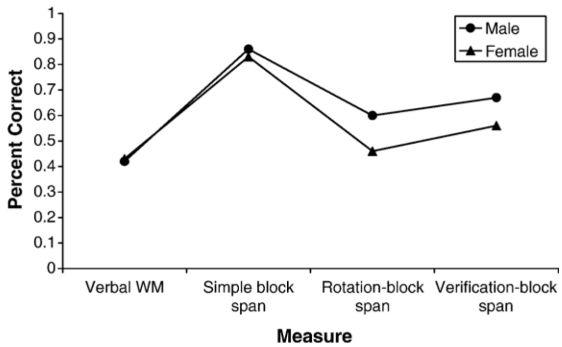

Fig. 1. Male–female differences on memory span measures as a function of percent correct.

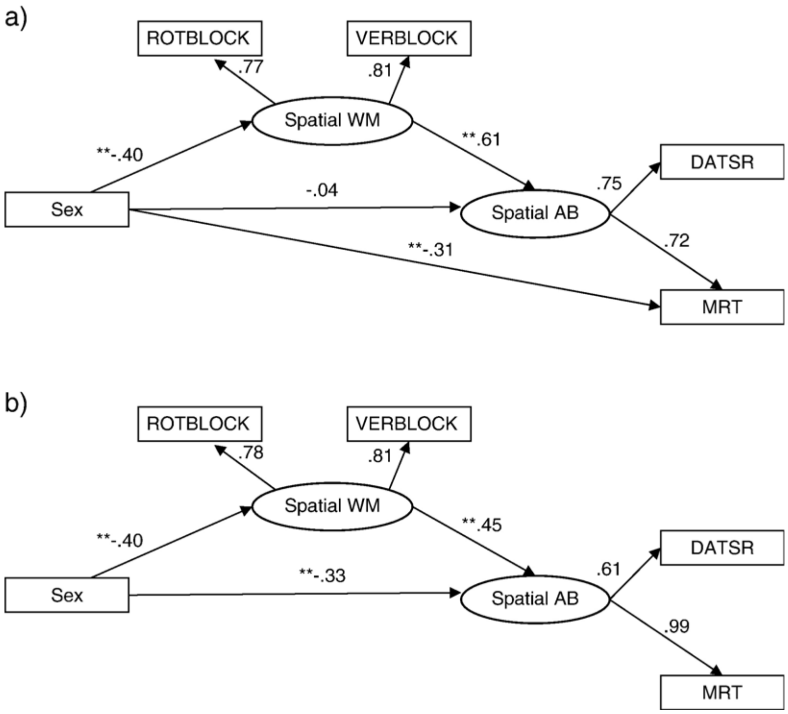

Fig. 2. Mediation of the effect of sex on spatial ability by spatial WM (a) with direct path from sex to MRT and (b) without direct path from sex to MRT. Sex is dummy-coded: male = 1, female = 2, *$p<.05$, **$p<.01$ (two-tailed). (a) Model fit: $\chi^2=.355$, $df=2$, $p=.837$, CFI = 1, RMSEA = .000. (b) Model fit: $\chi^2=5.54$, $df=3$, $p=.136$, CFI = .98, RMSEA = .093.

Because significant sex differences were found on all tests of spatial ability and spatial working memory tests, a mediation analysis, using structural equation modelling, was performed to see if spatial working memory mediated the sex difference in spatial ability (Fig. 2). In constructing the model, ROTBLOCK (rotation-block span) and VERBLOCK (verification-block span) were used as indicators of a latent spatial WM variable, while DATSR and MRT were used as indicators of a latent spatial ability variable. The model (Fig. 2a) provided an exceptionally good fit ($\chi^2=0.355$, $df=2$, $p=.837$, CFI = 1.00, RMSEA = .000). The significance of the indirect effect of sex on spatial ability, mediated by spatial WM, was calculated using the bootstrap method recommended by Shrout and Bolger (2002) to replace the more traditional but less accurate Sobel test. The direct effect of sex on spatial ability is near zero after controlling for spatial WM, and the indirect effect was significant ($p<.01$). This suggests a complete mediation.

Based on research showing greater sex differences on tests of three-dimensional mental rotation, we tested whether there is a direct sex effect on MRT that is not mediated by spatial WM. The direct effect of sex on MRT is significant ($\beta=-.31$, $p<.01$), controlling for spatial working memory (Fig. 2a). Furthermore, the model fit ($\chi^2=5.54$, $df=3$, $p=.136$, CFI = .98,

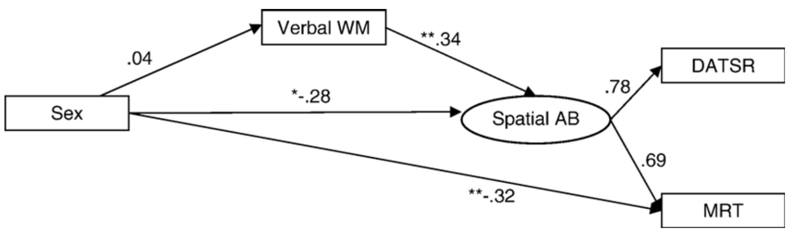

Fig. 3. Mediation of the effect of sex on spatial ability by verbal WM. Fit indices could not be computed because $df=0$. Sex is dummy-coded: male = 1, female = 2, *$p<.05$, **$p<.01$ (two-tailed).

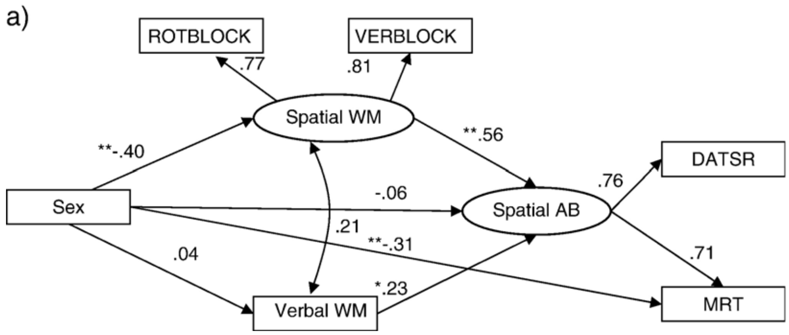

Fig. 4. The independent effect of verbal WM and spatial WM on spatial ability. Sex is dummy-coded: male = 1, female = 2, *$p<.05$, **$p<.01$ (two-tailed). (a) Model fit: $\chi^2=1.38$, $df=4$, $p=.847$, CFI = 1, RMSEA = .000.

RMSEA = .093) is worse when the direct path from sex to MRT is not taken into account (Fig. 2b). The RMSEA no longer falls in the acceptable range (<.08) and a chi-square-difference test reveals a significant decrement in fit (χ^2-difference = 5.19, $df=1$, $p=.023$).

This suggests that even though the MRT and DAT space relations test share an important source of variance that is related to sex differences and is accounted for by spatial WM, there is still a significant amount of variance unique to MRT (i.e. not shared with DAT space relations) that is accounted for by sex directly, rather than through sex's effect on spatial WM.

To test for the domain specificity of working memory, we ran a model to assess whether verbal working memory mediates the sex difference in spatial ability (Fig. 3), and a model that assessed the independent contribution of both spatial and verbal working memory on spatial ability (Fig. 4). The direct effect of sex on spatial ability is still significant ($\beta=-.28$, $p<.05$) after controlling for verbal WM, and the indirect effect of sex on spatial ability was not significant ($p=.64$, Fig. 3). Therefore, verbal WM does not mediate the sex difference in spatial ability. Even so, it is interesting to note that allowing spatial WM and verbal WM to correlate, both verbal WM ($\beta=.23$, $p<.05$) and spatial WM ($\beta=.56$, $p<.01$) independently predict spatial ability (Fig. 4). Therefore, both spatial WM and verbal WM seem to play a role in spatial ability performance, but only spatial WM accounts for the sex differences found in spatial ability.

4. Discussion

A central aim of the paper was to investigate whether spatial working memory capacity is the driving force determining the sex difference in mental rotation and spatial visualization ability. To achieve this aim, performance on various spatial ability and working memory measures was assessed. Then a mediation analysis using structural equation modelling was conducted to investigate whether the difference in spatial ability between males and females could be accounted for by measures of spatial working memory.

The results of the tests of spatial abilities showed that males and females differ significantly on both a test of spatial visualization and a test of three-dimensional mental rotation, with the former differences (measured by DAT-SR) being substantially smaller than the latter (measured by MRT). This is entirely consistent with the meta-analyses of Linn and Peterson (1985) and Voyer et al. (1995) showing male superiority on mental rotation and spatial visualization tests, with the difference being much more pronounced on tests of three-dimensional mental rotation.

The results of the spatial working memory tests also showed a significant difference between males and females. Both tests of spatial working memory displayed a statistically significant difference between males and females. This confirms and extends the work of Vecchi and Girelli (1998) who also found a sex difference in the active manipulation of spatial stimuli. However, it is not entirely consistent since they found no differences in the storage component. There is reason to believe that the current study was more valid. For one reason, the current study includes a larger sample of males and females. For another, the current study uses multiple measures of spatial working memory and explicitly set out to measure spatial working memory using the predominant methodology in the field of working memory.

The results of the mediation analysis showed that spatial working memory provided a complete mediation of the relationship between sex and spatial ability. In other words, when assessing the common variance across various tests of spatial ability, the answer to the question posed in the title of this paper is a resounding yes. Even so, there was still a significant amount of variance on the MRT that could not be explained by the male–female differences found on the spatial WM measures. These results are consistent with the research of Miyake et al. (2001) showing a strong relation between their latent spatial working memory factor and their latent factor of spatial visualization. However, the Miyake et al. study did not include the MRT in their battery of spatial visualization tests. Therefore, the current study is the first to show that spatial working memory is enough to explain the common variance across various tests of mental rotation, but is not enough to explain the unique variance on a test of three-dimensional mental rotation ability.

Further research should investigate the additional processes that are required in solving tests of three-dimensional mental rotation such as the MRT, since these are the tests that display the largest sex differences. The current study suggests that spatial working memory cannot be the whole picture, even though it is partly the picture. In addition, this study supports the suggestion that the type of mental rotation ability assessed by the MRT is an important aspect of spatial ability that should be studied separately from other tests of spatial visualization ability (Johnshon & Bouchard, 2005).

It should be noted that recent investigations in the social psychology domain have demonstrated significant effects of stereotype threat on mental rotation ability (McGlone & Aronson, in press) as well as working memory capacity (Schmader & Johns, 2003). In fact, the Schmader study found that a reduction in working memory capacity mediated the effect of stereotype threat on women's math performance. Therefore, it may be the case that stereotype threat affects spatial working memory capacity, which then influences spatial ability. In the current study, participants were asked their sex *before* participation. For females, this may have evoked a negative stereotype resulting in an inaccurate reflection of their true ability. Future research should investigate how stereotype threat affects sex differences in spatial working memory capacity and mental rotation ability.

A second goal of the paper was to assess the particular characteristics of memory tasks that give rise to sex differences. Males and females did not differ on a test of spatial short-term memory or a test of verbal working memory. They did however differ on a test of spatial working memory with either a spatial or a verbal processing component. Therefore, the difference between males and females in working memory is only found when the task is to maintain a representation of spatial stimuli while performing a concurrent processing task. The nature of that concurrent task is unimportant. This result is consistent with the results of Shah and Miyake (1996). Their rotation-arrow span and verification-arrow span tasks correlated with spatial ability .68 and .65, respectively.

The finding also parallels research in the verbal working memory domain. Conway and Engle (1996) administered a task in which participants had to do mental arithmetic (processing) while simultaneously remembering a single word (storage). All participants took a pre-test which determined their ability to solve mental arithmetic problems of various levels of difficulty. In the main study, participants received mental arithmetic problems that matched their specific ability level. Results showed that word-span scores correlated with scores on the verbal portion of the SAT between .50 and .60. Taken together, the previous research and the current results suggest that the correlation between working memory and abilities is probably not a result of differences in the ability to perform the processing component of a working memory task.

That females did not differ in spatial short term memory performance, but did differ once a processing component was added to the task is intriguing in light of the research looking at the different predictive powers of spatial short term memory and spatial working memory. Some studies have found evidence for a separation between spatial STM and spatial working memory, while others have found that spatial STM and WM are equally strong predictors of complex cognitive ability. Let us consider both sets of studies.

In the Miyake et al. (2001) study, a latent-variable model with a single spatial WM-STM factor fit the data as well as one that separated the WM and STM constructs (the latter were correlated at .86). In addition, this single spatial memory construct accounted for 35% of the variance in an executive-function construct measured by performance on the Tower of Hanoi problem and a task requiring the random generation of numbers. This study suggests that spatial WM and STM tasks measure a single construct related to executive attention.

At first blush, the results of the Miyake and Shah (1999) study also seem consistent with these findings. Spatial STM span tasks did correlate just as strongly with spatial ability as did their spatial WM span tasks. However, a partial-correlation analysis showed that spatial short-term memory and spatial working memory each contributed primarily unique, rather than shared, variance to spatial ability. In the Kane et al. (2004) study, spatial storage processes accounted for a substantial amount (30%) of the variance in gf over and above the variance accounted for by their executive-attention factor, suggesting that spatial STM and WM were each independently tapping different aspects of general ability.

The results of the current study may in fact be able to reconcile these discrepancies. With sex combined (see Table 1), the correlation between MRT and spatial short term memory was just as high (.466, $p<.01$) as the correlation between MRT and verification-block span (.467, $p<.01$) and rotation-block span (.441, $p<.01$). If no sex differences were taken into account, the result would seem to be consistent with the Miyake et al. (2001) study, by suggesting that spatial STM and spatial WM have the same predictive utility.

However, when correlation analyses were performed separately on males and females, a different picture emerged (see Appendix A). Spatial STM is more strongly related to spatial ability for females than males. None of the studies mentioned above looked at sex differences in their results. It is possible that at least some of the discrepancies between previous studies are a result of the fact that an analysis of the independent correlation of spatial STM and spatial WM with spatial ability may be dependent on the distribution of males and females in the sample. Since the current study only had an N of 50 for both males and females, this remains speculation. Nonetheless, this study suggests the potential importance for studies to conduct an analysis by sex.

A third goal of the paper was to assess the domain specificity of working memory. There are two sources of evidence from this study that suggest that working memory can be fractionated. First, males and females differed on all spatial working memory tests, but did not differ on the verbal working memory test. Second, the spatial working memory tests contributed to sex differences in spatial ability (accounting for most of the variance on the DAT space relations test), whereas the verbal working memory test did not. These results are consistent with other studies (e.g., Mackintosh & Bennett, 2003; Shah & Miyake, 1996) that have looked at the raw correlations (instead of the latent variable approach taken by working memory researchers such as Colom, Kane, and Engle) between tests and have found that spatial and verbal working memory have independent predictive utility. Therefore, when prediction of a domain specific ability is sought, it would appear that a domain specific view of working memory is useful.

An unexpected result was that the increased working memory load version of the DAT-SR test did not produce a larger sex difference than the standard version. This is puzzling since the increased working memory load version was carefully constructed to resemble the demands of a working memory test as much as possible. Also, the research of Gilhooly, Logie, Wetherick, and Wynn (1993) has demonstrated that the modification of working memory load on a test of ability affects the difficulty of the task. They varied the load on working memory by reading syllogisms out loud once only, so that participants had remember both premises while trying to work out their answer. This change of procedure had a significant impact on errors compared to a group who solved the syllogisms on a computer screen.

The results of the current study did show that the increased working memory load version of the test was harder for both males and females. Therefore, increasing the working memory load did increase the complexity of the task. It is also true that there was a significant difference between males and females on the increased memory load version of the test [$t(98)=1.16$, $p<.05$]. However, the sex difference was not greater than what was found for the standard version.

One possible explanation for this finding is that the increased working memory load version simply did not add enough extra to increase the magnitude of the sex difference already generated by the standard version. In other words, the standard version of the DAT-SR test already imposed a significant load on spatial WM—and it was this that caused the sex difference on the standard version. This view is supported by the fact that the sex differences on the standard version evaporated once both measures of spatial working memory were controlled for ($F<1$).

A limitation of the study is that only two tests were used to form a spatial ability factor and a spatial working memory factor. Further studies should use more than two tests to measure each latent variable. Also, only one test of verbal working memory was administered, so conclusions regarding the distinction between spatial and verbal working memory are merely suggestive. Even with this said, the tests that were used in this study have a very large research database to support their reliability and validity, and the verbal working memory test quite starkly did *not* account for the sex differences in spatial ability, *even though* it independently predicted spatial ability. Nonetheless, further studies investigating sex differences should include more than one measure of verbal working memory so as to better ascertain the independent contribution of verbal and spatial working memory in accounting for sex differences in domain specific abilities.

The phenomenon of sex differences in mental rotation and spatial visualization ability is well documented. The proximal causes of these differences are not as well documented, however. Hopefully, the methods and aims of this paper will open up a fruitful line of research that will further an understanding of sex differences in spatial ability specifically, and the nature of working memory and intelligence more generally.

Acknowledgments

The author would like to thank Nicholas Mackintosh for helpful comments and suggestions on an earlier draft. The author would also like to thank Sheila Bennett and Jim Blair for their kind assistance in recruiting participants, and Nicholas Turk-Browne and Colin DeYoung for assisting him with the data analysis. This paper is based on a thesis completed in partial fulfilment of the M. Phil in Experimental Psychology at University of Cambridge, and received support from the Gates Cambridge Scholarship. The author is currently at Yale University.

Appendix A. Correlation matrices for males and females separated

Table A1
Correlation matrix of all ability and span tests for males (n = 50)

Measure	1	2	3	4	5	6
1. DAT-SR	–	**.522	*.358	*.353	.218	.246
2. MRT-A	–	–	*.341	**.456	*.517	.227
3. Simple block span	–	–	–	**.528	**.537	.217
4. Rotation-block span	–	–	–	–	**.585	.190
5. Verification-block span	–	–	–	–	–	.157
6. Verbal working memory	–	–	–	–	–	–

*p < .05 (two-tailed), **p < .01 (two-tailed).

Table A2
Correlation matrix of all ability and span tests for females (n = 50)

Measure	1	2	3	4	5	6
1. DAT-SR	–	**.646	**.622	.250	**.489	*.298
2. MRT	–	–	**.573	.252	.276	*.320
3. Simple block span	–	–	–	**.464	**.486	.031
4. Rotation-block span	–	–	–	–	**.583	.202
5. Verification-block span	–	–	–	–	–	.093
6. Verbal WM	–	–	–	–	–	–

*p < .05 (two-tailed), **p < .01 (two-tailed).

References

Astin, A., Sax, L., Korn, W., & Mahoney, K. (1995). *The American freshman: National norms for fall 1995.* Los Angeles: Higher education Research Institute.

Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), *Attention and performance*, *Vol. VIII*521–539. Hillsdale, NJ: Erlbaum.

Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, *23*, 119–131.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1974). *Manual for the Differential Aptitude Tests* (5th ed.). New York: Psychological Corporation.

Carpenter, P. A., & Just, M. A. (1986). Spatial ability: An information processing approach to psychometrics. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*, *Vol. 3*. Hillsdale, NJ: Erlbaum.

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

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence*, *32*, 277–296.

Conway, AR. A., & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, *4*, 577–590.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

De Renzi, E., & Nichelli, P. (1975). Verbal and non-verbal short-term memory impairment following hemispheric damage. *Cortex*, *11*, 341–354.

Duff, S. J., & Hampson, E. (2001). A sex difference on a novel spatial working memory task in humans. *Brain and Cognition*, *47*, 470–493.

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial sketch pad in working memory. *Quarterly Journal of Experimental Psychology*, *38A*, 675–688.

Geiger, J. F., & Litwiller, R. M. (2005). Spatial working memory and gender differences in science. *Journal of Instructional Psychology*, *32*(1), 49–57.

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic reasoning tasks. *Memory and Cognition*, *21*, 115–124.

Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist*, *52*, 1091–1102.

Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, *12*(2), 229–246.

Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review*, *102*, 458–489.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*, 175–191.

Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 443–451). New York: MacMillan.

Hyde, J. S. (1981). How large are cognitive gender differences? *American Psychologist*, *36*, 892–901.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic.

Johnshon, W., & Bouchard Jr., T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature*, *363*, 623–625.

Jurden, F. H. (1995). Individual differences in working memory and complex cognition. *Journal of Educational Psychology*, *87*, 93–102.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, *92*, 137–172.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). Generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Janal of Experimental Psychology: General*, *133*(2), 189−217.

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi Block-Tapping task: Standardisation and normative data. *Applied Neuropsychology*, *7*, 252−258.

Kimura, D., & Hampson, E. (1992). Neural and hormonal mechanisms mediating sex differences in cognition. In P. A. Vernon (Ed.), *Biological approaches to human intelligence.* Norwood, NJ: Ablex.

Kyllonen, P. C. (1996). Is working memory capacity Spearman's *g*? In I. Dennis, & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 49−75). Mahwah, NJ: Erlbaum.

Linn, M. C., & Peterson, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479−1498.

Loehlin, J. C. (2000). Group differences in intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 176−193). New York: Cambridge University Press.

Logie, R. H. (1995). *Visuo-spatial working memory.* Hove: Lawrence Erlbaum.

Lohman, D. F. (1986). The effect of speed-accuracy tradeoff on sex differences in mental rotation. *Perception and Psychophysics*, *39* (6), 427−436.

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences*, *17*(2), 257−271.

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin*, *109*, 267−296.

Mackintosh, N. J. (1998). *IQ and human intelligence.* New York, NY: Oxford University Press.

Mackintosh, N. J., & Bennett, E. S. (2003). The fractionation of working memory maps onto different components of intelligence. *Intelligence*, *31*, 519−531.

Mackintosh, N. J., & Bennett, E. S. (2005). What do Raven's matrices measure? An analysis in terms of sex differences. *Intelligence*, *33*, 663−674.

Masters, M. S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics*, *23*, 337−341.

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

McGlone, M. S., & Aronson, J. (in press). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*.

Milner, B. (1971). Interhemispheric differences in the localisation of psychological processes in man. *British Medical Bulletin*, *27*, 272−277.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621−640.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.

Petrides, M., Alivisatos, B., Evans, A. C., & Meyer, E. (1993). Functional activation of the human frontal-cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of the United States of America*, *90*, 878−882.

Plomin, R., DeFries, J. C., & McClearn, G. E. (1990). *Behavior genetics: A primer.* New York: Freeman.

Rosser, R. (1980). *Acquisition of spatial concepts in relation to age and sex (Final report on Grant No. NIE-6-79-0091 from the National Institute of Education, Department of Education).* Tucson: University of Arizona.

Salthouse, T. A., Babcock, R. L., Mitchell, D. R. D., Palmon, R., & Skovronek, E. (1990). Sources of individual differences in spatial visualization ability. *Intelligence*, *14*, 187−230.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440−452.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4−27.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422−445.

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial versus object working-memory-PET investigations. *Journal of Cognitive Neuroscience*, *7*, 337−356.

The Psychological Corporation. (1995). *DAT for selection-verbal reasoning (European adaptation).* Harcourt Brace and Company.

Thurstone, L. L. (1938). *Primary mental abilities.* Chicago, IL: University of Chicago Press.

Vandenberg, S. G., & Kruse, A. R. (1978). Mental rotations: Group tests of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*, 599−604.

Vecchi, T., & Girelli, L. (1998). Gender differences in visuo-spatial processing: The importance of distinguishing between passive storage and active manipulation. *Acta Psychologica*, *99*, 1−16.

Vernon, P. E. (1971). *The structure of human abilities.* London: Methuen.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial ability: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250−270.