Available online at www.sciencedirect.com
ScienceDirect
Brain and Cognition 64 (2007) 217–227
Brain and Cognition
www.elsevier.com/locate/b&c

Suitable stimuli to obtain (no) gender differences in the speed of cognitive processes involved in mental rotation ☆

Petra Jansen-Osmann, Martin Heil *

Institute of Experimental Psychology, Heinrich-Heine-University, 40225 Düsseldorf, Germany

Accepted 6 March 2007
Available online 12 April 2007

Abstract

Gender differences in speed of perceptual comparison, of picture-plane mental rotation, and in switching costs between trials that do and do not require mental rotation, were investigated as a function of stimulus material with a total sample size of $N = 360$. Alphanumeric characters, PMA symbols, animal drawings, polygons and 3D cube figures were used with an otherwise completely equivalent experimental design in which age and speed-based IQ were comparable across male and female groups. Small gender-related differences in speed of perceptual comparison were found with the magnitude as well as the direction depending upon the stimulus material. Polygons were the only material that produced substantial and reliable gender differences in mental rotation speed, and additionally revealed gender differences in switching costs. Thus, whereas gender differences in paper–pencil mental rotation tests constitute an empirical reality, the generalization that men outperform women in the speed of mental rotation was not supported in the present experiment.

Keywords: Mental rotation; Spatial cognition; Gender differences; Perceptual comparison; Maintaining readiness; Switching costs

1. Introduction

The cognitive process of imagining an object turning around is called mental rotation (Shepard and Metzler, 1971) and constitutes one important operation in the general class of mental transformations as well as a critical ingredient in spatial intelligence. Thanks to the many research areas dealing with mental rotation, a body of evidence has accumulated. Mental rotation seems to be a cognitive process implemented in the parietal cortex (e.g., Jordan, Heinze, Lutz, Kanowski, and Jäncke, 2001) and involves what Shepard and Chipman (1970) refer to as a "second order isomorphism" between the physics of real rotation and imagined rotation. In particular, rotations are continuous (e.g., Heil, Bajric, Rösler, and Hennighausen, 1997) and proceed through intermediate angles (Cooper, 1976).

Many authors claim, based on existing evidence that whereas females outperform males on e.g., measures of verbal fluency, males outperform females on certain tests of spatial ability (e.g., Halpern, 1992; Petrusic, Varro, and Jamieson, 1978). This male advantage is largest on mental rotation tasks (Linn and Petersen, 1985; Voyer, Voyer, and Bryden, 1995), where gender effects were usually investigated on the basis of paper–pencil tests. With the Vandenberg–Kuse Mental Rotation Test (MRT, Vandenberg and Kuse, 1978) that uses Shepard–Metzler 3D cube figures, the gender differences amount to one standard deviation (see, e.g., Voyer et al., 1995). The 2D Card Rotation Test (CRT, Ekstrom, French, and Harman, 1976), however, yielded a substantially smaller effect-size of 0.3, indicating the importance of the stimulus material used. Nevertheless, neither the underlying causes (e.g., Voyer et al., 1995) of

☆ This study was supported by the German Research Foundation. We thank Gabriela Gan, Brigitte Kobiela, Britta Porep, and Monika Renvert for their help during data acquisition. We gratefully appreciate Axel Buchner's discussions and his support regarding the power calculations. We also gratefully acknowledge the very helpful suggestions of Steven Christman and three anonymous reviewers.
* Corresponding author.
E-mail address: Petra.Jansen-Osmann@uni-duesseldorf.de (M. Heil).

0278-2626/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2007.03.002

nor even the performance mechanisms (e.g., Peters, 2005) responsible for this large effect-size are understood. In particular, it is still unclear whether the male advantage on mental rotation is caused by a speed superiority in the pure process of mentally rotating a stimulus, and if so, whether this superiority is dependent upon the type of stimuli to be mentally rotated or not.

Therefore, in the present paper we decompose the mental rotation task into information processing components to determine whether or not their durations are affected by gender. Therefore, the speed of three different components of information processing (a full description follows later) were investigated in detail: *Perceptual comparison* (i.e., the time to compare visual stimuli in a context where mental rotation is never required), *rotational uncertainty* (the cost of being in a context where some trials require mental rotation but where on the given trial no mental rotation is needed, see Ilan and Miller, 1994; Jansen-Osmann and Heil, 2006, in press) and *mental rotation speed* itself (i.e., the speed to mentally rotate a stimulus, expressed as °/s). Moreover, given that a wealth of evidence demonstrates that differences in stimulus materials can greatly affect the size (or even the sheer existence) of gender differences in mental rotation (see, e.g., Collins and Kimura, 1997), 5 different types of stimulus material were used for rotation in the picture plane: Alphanumeric characters, PMA symbols (Thurstone, 1958), animal drawings, polygons, and Shepard–Metzler 3D cube figures with otherwise completely equivalent experimental design.

1.1. Gender differences in paper–pencil mental rotation tests

The most common paper–pencil measure of mental rotation function is the Vandenberg and Kuse (1978; see Peters, Laeng, Latham, Jackson, Zaiyouna & Richardson, 1995) MRT, which uses depictions of 3-D cube figures designed by Shepard and Metzler (1971) that mentally are to be rotated in depth. The speed test consists of 24 items, and each item consists of a row of 1 standard cube figure and 4 comparison ones. Two comparison figures are correct matches rotated in depth; the remaining 2 are incorrect matches. Men typically outperform women on this task by as much as one standard deviation (Linn and Petersen, 1985; Voyer et al., 1995).

While the existence of the gender differences in the MRT is of no doubt, the causes and even the mechanisms responsible for differences in test performance are less well understood. Two broad classes of explanations are the "psycho-social" variety and the "biological-neuronal" variety. Examples of the "psycho-social" variety are stereotype threat (e.g., Shih, Pittinsky, and Ambady, 1999), sex role identification (e.g., Signorella and Jamison, 1986), or differential experience and socialization (e.g., Baenninger and Newcombe, 1989). Examples of the "biological-neuronal" variety are rate of maturation (Sanders and Soares, 1986), genetic complement (McGee, 1982), sex hormone level (e.g., Imperato-McGinley, Pichardo, Gautier, Voyer, and Bryden, 1991) and cerebral lateralization (e.g., McGlone, 1980).

Without doubt, the determination of the root causes for any gender effects observed is a challenging endeavor, and will require further research to uncover. Unfortunately, however, the performance mechanisms that yield the gender differences are also not understood yet (see, e.g., Voyer and Saunders, 2004). Goldstein, Haldane, and Mitchell (1990) reported findings that the gender difference on the MRT disappears when subjects were allowed sufficient time to attempt all items or when the scoring procedure controlled for the number of items attempted. In contrast to Goldstein et al., and in line with the majority of the published data (see, e.g., Delgado and Prieto, 1996; Resnick, 1993), Masters (1998) showed that the gender difference was not affected by performance factors, neither by the scoring method nor by the time limits used, a result which was also obtained by Peters (2005). Peters (2005) obtained evidence that although females attempted fewer items than males under standard timing condition, the magnitude of the gender difference did not change when subjects did the MRT with double the usual time allowed for the test. To sum up, the cognitive mechanisms that yield the gender differences are not understood yet.

Empirical evidence suggests that the dimensionality of the task (depth rotation versus picture plane rotation and 3D versus 2D objects) is not crucial with respect to the size of the male advantage (Collins and Kimura, 1997). Studies have failed to converge, however, on an unambiguous conclusion whether or not the magnitude of the gender difference may be a function of the difficulty of the test (with "difficulty" defined either as overall error rate or mean RT). The male advantage for the Spatial Relations subtest of the Primary Mental Abilities Battery (PMA; Thurstone, 1958), depicting picture plane rotations of 2D objects, is, on average, less than one half of that of the MRT (Voyer et al., 1995). Collins and Kimura (1997), however, introduced a test depicting picture plane rotations of 2D objects (some of them were PMA symbols) with different levels of task difficulty within this test. A male advantage (at least) as large as that seen on the MRT was found for the difficult version while for the easy version the male advantage missed significance. These data suggest that neither depth rotations nor 3-D objects are required to elicit substantial gender differences in mental rotation, but that the difficulty of the task might be crucial. Peters et al. (1995), however, compared the standard MRT with a very difficult version which required subjects to rotate the cube figures around two axes. Overall, performance on this more difficult version was about 30% lower than on the MRT, and gender differences were reduced by half compared to the MRT. This suggests that the magnitude of the gender difference might be a function of test difficulty for picture plane rotations only but not so for depth rotations, but more data are definitely needed.

1.2. Gender differences in mental rotation in reaction time studies

One approach to learn more about the mechanisms responsible for gender differences in mental rotation tasks is the chronometric approach based on RT (reaction time) measures. In chronometric studies of mental rotation, two stimuli are presented with varying angular disparity, and the RT is measured when participants decide whether these two do or do not match when mentally aligned. Cooper and Shepard (1973) assumed that RT in this situation is the sum of the times required for the following component processes: stimulus encoding, mental rotation of the stimuli, comparison of the stimuli once mentally aligned by rotation, and response selection and execution. The speed of mental rotation (expressed as degrees per second) is measured as the inverse of the slope of the regression line relating RT to angular disparity. The intercept of the RT function is regularly traced back to perceptual comparison time (assuming that response selection and execution do not differ between genders), although this conclusion is definitely not obligatory (Ilan and Miller, 1994; Jansen-Osmann and Heil, 2006, in press), see below.

Chronometric studies, however, often completely ignored the gender of the subjects as a potential factor. In cases were gender was included as a factor, the power to pick up any gender effects had often been limited by small sample sizes or by uncontrolled extraneous differences between the male and female participants (e.g., age, information processing speed, academic program, intelligence). Such studies have failed to converge on an unambiguous picture. Seven of the 15 chronometric studies that were entered into the meta-analysis of Voyer et al. (1995) showed no gender differences at all, resulting in a small to medium overall effect-size ($d = 0.37$) as defined by Cohen (1977). However, five of the seven non-significant studies included by Voyer et al. (1995) involved children precluding a generalization of the effect-size to adult subjects.

The conclusions to be drawn from chronometric studies are still unclear because the information processing locus of gender effects varies from experiment to experiment: Sometimes, gender differences were found in the slope of the function relating RT to orientation suggesting a higher speed of mental rotation for men (Kail, Carter, and Pellegrino, 1979). Sometimes, however, the intercepts differed either in addition to slope differences (Loring-Meier and Halpern, 1999) or without slope differences (Wiedenbauer, Schmid, and Jansen-Osmann, 2007) suggesting a male advantage in cognitive processes other than mental rotation itself that might explain (some of) the gender differences in the MRT. Similar inconsistencies plague cognitive neuroscience experiments in mental rotation (e.g., Epting and Overman, 1998; Ho, Gilger, and Brink, 1986; Hooven, Chabris, Ellison, and Kosslyn, 2004; Voyer and Bryden, 1990). In addition, little is known about gender differences in mental rotation speed as a function of the stimulus material.

1.3. An information processing approach decomposing cognitive processes involved during mental rotation

As stated above, gender related differences in the intercept of the RT function are regularly inferred to reflect gender differences in perceptual comparison time (see, e.g., Wiedenbauer et al., 2007), but this conclusion is definitely not obligatory as recent studies have revealed (Ilan and Miller, 1994; Jansen-Osmann and Heil, 2006, in press). Instead, the intercept (or the RT in a condition where the angular disparity is 0°) reflects the sum of perceptual comparison time and of the costs for maintaining readiness for mental rotation, as suggested by Ilan and Miller (1994). Put another way, in a typical mental rotation experiment, trials with 0° disparity are often outnumbered greatly by trials with >0° disparity. Thus, on a majority of the trials there is a need to rotate and on a minority of the trials, this process is unnecessary. Managing this difference between the task requirements for the two types of trials involves a cost. Ilan and Miller (1994) used the framework of Donders' (1868/1969) subtraction method to investigate whether or not the mental rotation process is purely inserted into another task, a same/different comparison task. The logic of the experiments of Ilan and Miller (1994) was as follows: When the stimulus pair is presented with an angular disparity of, e.g., 90°, then the above mentioned processing stages should occur: (1) perceptual processing including stimulus identification, (2) 90° mental rotation, (3) same vs. different judgment including response selection, (4) motor response. When both stimuli are presented upright, however, no mental rotation is needed, and (assuming the validity of Donders' assumption) only steps 1, 3, and 4 are to be executed. Thus, the upright condition should reflect the time participants would need in a perceptual comparison task. If the assumption of pure insertion is valid, then the RT on trials with upright stimuli (i.e., not requiring rotation) in the context of trials that do require rotation (SU-condition, i.e. Sometimes Upright) should not differ from the RT to upright stimuli embedded in upright stimuli only (AU-condition). This assumes that the 2 contexts (AU and SU) are otherwise experimentally equivalent, i.e., do not differ with respect to instructions, decisions, responses, as well as stimulus probabilities, a problem smartly solved by Ilan and Miller (1994). Ilan and Miller (1994) clearly revealed that participants took substantially longer to respond to upright characters in conditions containing rotated stimuli (SU) than in conditions containing only upright stimuli (AU). Furthermore, the authors showed that this rotational uncertainty effect was not caused by the need to determine stimulus orientation, and was separable from mental rotation itself (see also Jansen-Osmann and Heil, in press). In the studies that used this approach, no gender differences were obtained with alphanumeric stimuli (Ilan and Miller, 1994) and with animal drawings (Jansen-Osmann and Heil, 2006, in press).

Thus, the intercept of the RT function (as well as the RT for stimuli with 0° disparity in mental rotation blocks) in fact reflects the sum of perceptual comparison time (the RT in the AU condition) and the rotational uncertainty effect, i.e., the costs for maintaining readiness for mental rotation, outlined above (the increase in RT to upright stimuli in the SU condition compared to the AU condition). Gender-related differences in the intercept of the RT function could plausibly be ascribed to gender differences in a same/different comparison task and/or gender differences in the mixing costs, i.e. in the rotational uncertainty effect. In the former case, they would tell us something about gender effects in perceptual comparison time, in the latter case, they would tell us something about gender differences in the costs of switching (perhaps inhibiting mental rotation) between trials that do or do not require mental rotation. This distinction should be taken into account when gender differences in mental rotation tasks are investigated in detail.

1.4. The goal of the present experiments

The goal of the present study, therefore, was to investigate gender differences in different components of information processing involved in a mental rotation task and maintaining readiness for mental rotation as a function of stimulus material. In particular, we sought to accomplish this in an experiment with a sufficiently large sample size where age as well as cognitive speed based IQ was controlled to allow these effects a fair chance to be manifest. Given a sample size of $N = 72$ for each type of stimulus material, an effect-size $d = 0.50$ (that is, medium effects as defined by Cohen, 1977) could be detected with a level of $\alpha = .10$ and a probability of $1 - \beta = .80$. Based on the work of Collins and Kimura (1997), the five different types of stimuli were always rotated in the picture plane (see discussion). We used alphanumeric characters, PMA-symbols, animal drawings, polygons, and Shepard–Metzler 3D cube figures, with otherwise completely equivalent experimental design.

We derived three dependent variables to measure the different components of cognitive processing speed relevant in a mental rotation task:

1. Perceptual comparison time: The average RT in the AU condition was taken as an indicator of pure perceptual comparison time because these trials occur in a context where mental rotation is *never* required.[1]
2. Cost for maintaining readiness (rotational uncertainty effect): Given previous research, we expect that RT to upright stimuli embedded with stimuli at other orientations (SU 0°) is slower than RT to the same upright stimuli embedded in upright stimuli only (AU condition), see, e.g., Ilan and Miller (1994). These costs (i.e., the difference RT SU 0° minus AU 0°) reflect the need to maintain readiness for mental rotation even when it is not necessary at all for the upright stimuli in the SU condition. RT to upright stimuli in the SU condition (as well as the intercept of the regression line of RT as a function of angular disparity), therefore, reflect the sum of perceptual comparison time and the rotational uncertainty effect.[2]
3. Mental rotation speed was calculated as the inverse of the slope of the regression line, calculated separately for each subject, relating RT to angular disparity, expressed as degrees per second.[3]

2. Methods

2.1. Participants

In all, 360 adults participated, with 36 males and 36 females for each of the five different stimulus materials. With respect to age and cognitive speed-based intelligence (see Table 1),[4] we found neither an effect of gender or stimulus material group nor an interaction of these two factors (all F-values <1.0). Participants were paid € 6 per hour.

2.2. Apparatus and stimuli

The five stimulus material conditions were completely equivalent with the exception of the stimuli used. In condition 1, the stimuli consisted of 12 alphanumeric characters (F, P, R, 2, 4, 7, a, e, y, h, k, t. Font: Times New Roman). In condition 2, the stimuli consisted of 12 PMA symbols (Thurstone, 1958). In condition 3, the stimuli consisted of colored drawings of 12 different animals (camel, crocodile, dog, donkey, elephant, grizzly, lion, pig, rhino, sheep, turtle, and zebra, respectively, from Rossion and Pourtois, 2004). In condition 4, the stimuli consisted of 12 polygons with 5 or 6 vertices. In condition 5, the stimuli consisted of 12 perspective line drawings of three-dimensional forms (each one composed of 10 cubes) similar in construction to those used by Shepard and Metzler (1971). An example for each material is given in Fig. 1. In all cases, 2 versions

[1] All power calculations reported in this article were conducted using the GPower program (Erdfelder, Faul, and Buchner, 1996).

[2] The regression line estimated intercept revealed no new information. I.e., whenever perceptual comparison time and the rotational uncertainty effect did not differ between genders, the intercept (as an estimate of the sum of these two variables) also did not differ. In the cases where either perceptual comparison time or the rotational uncertainty effect differed between genders, the intercept followed the pattern of results. Therefore, intercept results are not presented.

[3] From the slope of the regression line, we calculated both estimates for the speed of mental rotation (expressed as degrees per second) and estimates for the slope of the RT function (expressed as milliseconds per degree). Owing to their non-linear relationship, the results may not be equivalent. Nevertheless, in the present cases, the analyses revealed comparable results for speed and slope. Therefore, only the findings with speed estimates are reported.

[4] Additionally, we also calculated analyses of covariance with age and estimated IQ as covariates. Since the pattern of results was not affected at all, these data are not reported.

Table 1
Mean age in years and mean estimated IQ (standard deviations in parentheses) per gender group for the five stimulus materials

Alphanumeric characters			
Females		Males	
Age	IQ	Age	IQ
23.3 (2.9)	112.2 (11.7)	24.6 (3.0)	109.4 (16.9)
PMA symbols			
Females		Males	
Age	IQ	Age	IQ
25.1 (5.3)	107.9 (16.0)	26.3 (5.8)	106.4 (14.2)
Animal drawings			
Females		Males	
Age	IQ	Age	IQ
25.9 (5.2)	108.4 (13.7)	26.1 (4.9)	109.1 (14.8)
Polygons			
Females		Males	
Age	IQ	Age	IQ
25.1 (5.3)	108.6 (16.6)	26.3 (5.7)	105.6 (13.9)
Cube figures			
Females		Males	
Age	IQ	Age	IQ
24.6 (5.0)	114.1 (17.2)	25.2 (3.5)	111.6 (17.8)

of the same stimulus were presented together. The right stimulus was either identical to the left or mirror-reversed. The two stimulus figures on the screen had a size of 4 by 4 cm, with a space of 2 cm between the stimuli. Subjects were free to choose the most comfortable viewing distance.

The 12 stimuli were divided into 3 sets of 4 stimuli each. One set always appeared with an angular disparity of 0° between the two stimuli in the AU as well as the SU condition. A second set appeared with 0° disparity in the AU condition but with 90° disparity in the SU condition. A third set of 4 stimuli also appeared with 0° disparity in the AU condition but with 180° disparity in the SU condition. Because all stimuli appeared equally often in both conditions, this arrangement equated stimulus as well as response probabilities. Which stimulus was used in which set was counterbalanced across participants. Participants responded "same" by pressing the left mouse button with their index finger and "different" by pressing the right mouse button with their middle finger. The experiments were run on a PC with a 17″ monitor (refresh rate: 75 Hz).

For cognitive speed-based IQ estimation, the *Number Connection Test* (Zahlenverbindungstest, ZVT, Oswald and Roth, 1987) was used, a test in which 4 sheets of papers have to be processed, On each sheet, the numbers from 1 to 90 are presented in random order in a matrix of 9 rows and 10 columns. Participants have to connect the numbers with lines in ascending order as fast as possible. The time needed is averaged across the 4 sheets. Normative data are available, and scores can be converted into IQ estimations. The correlation between the ZVT and standard IQ tests is about $r = .60$ to .80, the reliability (internal consistency as well as 6 month retest) of the ZVT is about .90 to .95. The ZVT is the equivalent to the Trail Making Test A

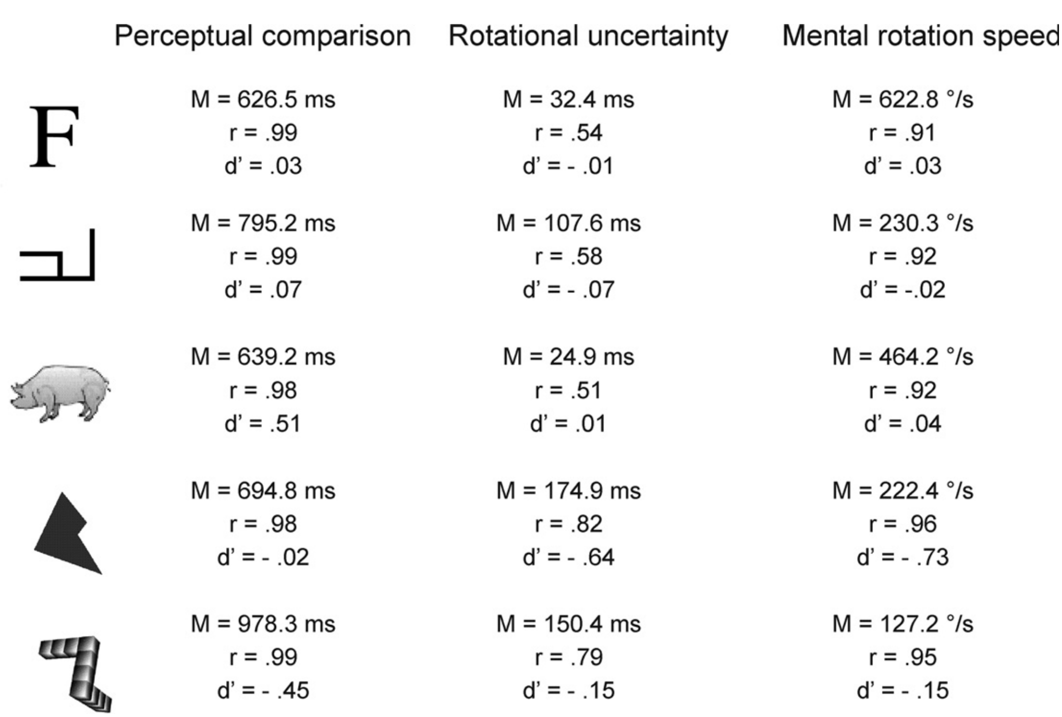

Fig. 1. Examples of the stimuli used and summary of the results, separately for perceptual comparison time, the rotational uncertainty effect, and mental rotation speed. Presented are the overall mean (M, averaged across both genders), the internal consistency (r) of the respective measure, and the effect-size (d) of the gender difference. Positive effect-sizes indicate situations where women outperformed men. Perceptual comparison time was based on RT in the AU condition, i.e., in a context where mental rotation is *never* required. The rotational uncertainty effect reflects the difference in RT to upright stimuli embedded with stimuli at other orientations (SU 0°) and RT to the same upright stimuli embedded in upright stimuli only (AU condition). Mental rotation speed was calculated as the inverse of the slope of the regression line relating RT to angular disparity, expressed as degrees per second.

(Reitan, 1956). The test administration including instructions and practice matrices, takes about 15 min.

2.3. Procedure

Individual test sessions lasted about 90 min and took place in a laboratory at the Heinrich–Heine-University of Duesseldorf. Participants were told to respond as quickly and as accurately as possible. Each session consisted of two conditions (AU and SU), the order of which was counterbalanced across participants, each one preceded by 48 corresponding practice trials.[5] In the AU-condition all stimuli were presented with 0° disparity, in the SU-condition some stimuli were presented with 0° disparity but others were rotated at 90° or 180°. Before each condition, participants were given instructions on the nature of the task required and they were told that stimuli would appear always upright (in the AU-condition) or sometimes upright (in the SU-condition). Between the two conditions, the Number Connection Test was administered.

Trials were presented in blocks of 48 trials each. The procedure was partially self-paced in that the participants initiated each block by pressing a key. Each trial began with a 500 ms background grey screen. Thereafter, the stimulus pair appeared and remained on until the subject responded. Feedback was given in the form of a "+" for correct responses or a "−" for an incorrect response presented for 500 ms in the centre of the screen. After 1500 ms the next trial began. Each combination of type of response (same versus mirror reversed), and stimulus (12) occurred 14 times resulting in 336 experimental trials for each of the two conditions (AU and SU).

3. Results[6]

Only trials with correct responses were used for RT analyses. Moreover, because angular disparity is not defined for "different" responses when cube figures are used as stimuli (see, e.g., Jolicoeur, Regehr, Smith, and Smith, 1985), the statistical analyses presented were restricted to "same" responses only. For the sake of completeness and to demonstrate that type of response did not influence the gender effects observed, results are presented separately for both types of responses in Table 3 and in Fig. 2. Prior to analyses, RT data were trimmed for outliers. RTs more than 2 SDs above or below the mean per condition and per subject were excluded. First, internal consistencies were determined by splitting the remaining raw data per subject according to odd versus event trial numbers, separately for (a) perceptual comparison time, (b) the rotational uncertainty effect, and (c) mental rotation speed (expressed as °/s). Second, analyses of variance (ANOVAs) were calculated with the between-subject factors "gender" and "stimulus material".[7]

Table 2
Mean error rate in percent (standard deviations in parentheses) as a function stimulus type, and gender in the always upright (AU) and the sometimes upright (SU) conditions

	AU	0°-SU	90°-SU	180°-SU
Alphanumeric characters				
Women	2.36 (1.8)	2.33 (2.1)	3.76 (2.9)	4.36 (4.2)
Men	2.18 (2.1)	2.26 (2.2)	4.03 (4.1)	4.18 (3.9)
PMA symbols				
Women	2.15 (2.3)	2.17 (2.1)	3.41 (2.2)	5.45 (4.1)
Men	2.56 (2.4)	1.66 (2.0)	3.59 (3.1)	5.78 (5.3)
Animal drawings				
Women	1.17 (2.0)	1.16 (2.3)	2.13 (2.6)	4.72 (4.4)
Men	1.48 (2.2)	1.91 (2.8)	2.60 (3.7)	4.51 (4.8)
Polygons				
Women	1.89 (2.1)	1.51 (1.9)	4.54 (3.5)	8.21 (5.6)
Men	2.09 (1.8)	1.22 (1.9)	4.53 (3.2)	6.93 (4.5)
Cube figures				
Women	1.62 (1.5)	1.52 (1.4)	3.80 (3.3)	7.25 (6.2)
Men	1.69 (1.3)	1.82 (2.1)	4.78 (4.4)	8.27 (5.6)

In the latter case, error rates are presented separately for the three angular disparities.

3.1. Internal consistency

Internal consistencies for perceptual comparison time were substantial for all five stimulus materials, amounting to either .98 or .99. For the rotational uncertainty effect, internal consistencies were moderately high to substantial for all five stimulus materials, ranging from .51 to .82. Internal consistencies for mental rotation speed were substantial for all five stimulus materials, ranging between .91 and .96, see Fig. 1.

3.2. Perceptual comparison time

The RT-ANOVA indicated a main effect of stimulus material ($F(4,350) = 58.42$, $p < .01$). Bonferroni post-hoc tests revealed that perceptual comparison time was slowest for cube figures, faster for PMA symbols, followed by

[5] From a chronometric point of view, 48 practice trials is a relatively low number whereas from a psychometric point of view, it seems quite a lot. However, the performance pattern does not change substantially with practice as long as a sufficient number of different stimuli are used (Heil, Rösler, Link, & Bajric, 1998).

[6] Additionally, error rates were also used as dependent variables. Error rates were low, and increased with increasing angular disparity. Error rates, however, were never affected by gender, neither as a main effect nor as an interaction term. Most importantly, we did not even observe a non-significant trend towards a speed accuracy trade-off. The only non-significant trend ($p = .1471$) present at all, pointed toward a somewhat larger error rate increase with increasing angular disparity for women when polygons were used. Error rates are presented in Table 2.

[7] Sequence of presentation (AU or SU condition first) did not interact with gender. Therefore, the results are presented collapsed across this factor.

Table 3
Mean perceptual comparison time, rotational uncertainty effect and mental rotation speed (with standard errors in parentheses) per gender group, separately for the two types of responses and the five stimulus types (significant gender differences in bold)

	Perceptual comparison (in ms)		Rotational uncertainty (in ms)		Mental rotation speed (in ms/°)	
	Same	Different	Same	Different	Same	Different
Alphanumeric characters						
Women	627.4 (14.4)	661.6 (15.2)	36.6 (10.2)	66.6 (12.2)	645.8(62.8)	765.1 (121.9)
Men	625.6 (17.7)	670.3 (20.7)	28.1 (10.0)	78.7 (12.6)	599.7 (82.5)	815.6 (126.9)
PMA symbols						
Women	791.6 (27.6)	875.0 (36.6)	117.4 (23.8)	170.1 (32.7)	233.2 (18.1)	266.1 (21.1)
Men	798.9 (36.2)	908.6 (64.3)	97.8 (26.8)	162.4 (58.6)	227.4(14.4)	282.0 (20.9)
Animal drawings						
Women	**618.2** (13.0)	**649.7** (14.2)	22.3 (8.5)	64.9 (8.3)	444.2 (31.0)	500.2 (30.4)
Men	**660.1** (16.0)	**700.9** (19.0)	29.6 (12.3)	64.8 (11.4)	484.1 (27.9)	531.2 (37.6)
Polygons						
Women	683.5 (25.1)	739.9 (27.8)	**218.7** (33.4)	**280.6**(45.6)	**177.4** (16.2)	**193.1** (16.7)
Men	706.0 (24.7)	711.9 (21.6)	**131.1** (26.7)	**194.6** (30.8)	**267.3** (35.4)	**275.1** (23.5)
Cube figures						
Women	1026.4 (42.8)	1238.8 (80.1)	111.5 (38.9)	—*	121.5 (10.3)	—*
Men	930.1 (33.5)	1101.0 (60.0)	189.2 (34.1)	—*	132.9 (15.4)	—*

Note. Perceptual comparison time was based on RT in the always upright condition, i.e., in a context where mental rotation was *never* required. The rotational uncertainty effect reflects the difference in RT to upright stimuli embedded with stimuli at other orientations and RT to the same upright stimuli embedded in upright stimuli only. Mental rotation speed was calculated as the inverse of the slope of the regression line relating RT to angular disparity, expressed as degrees per second.
* The rotational uncertainty effect and the mental rotation speed were not estimated for different pairs of cube figure because RT did not increase with angular disparity in the SU block, see Fig. 2 and text.

polygons. The fastest perceptual comparison times were found for characters and animal drawings, which did not differ from each other. Gender had no significant effect, neither as a main effect nor as an interaction. The interaction between gender and stimulus material, however, revealed a non-significant trend ($F(4,350) = 1.99$, $p < .0960$). The gender effect-sizes for characters, PMA symbols and polygons were close to 0, see Fig. 1. Animal drawings, however, revealed a female advantage in perceptual comparison time of $d = .51$ ($F(1,70) = 4.70$, $p < .05$) whereas cube figures revealed a trend towards a male advantage of $d = .36$ ($F(1,70) = 3.60$, $p < .0619$).

3.3. Rotational uncertainty effect

The ANOVA revealed a main effects of stimulus material ($F(4,350) = 55.41$, $p < .01$). Bonferroni post-hoc tests revealed that the rotational uncertainty effect was small although significantly different from zero for characters and animal drawings, medium sized for PMA symbols, and largest for cube figures and polygons.

Of main interest, however, was the significant interaction between gender and stimulus material ($F(4,350) = 2.91$, $p < .05$). The gender effect-sizes for characters, PMA symbols, animal drawings and cube figures, respectively, were low and not significantly different from 0, see Fig. 1. Polygons, however, revealed a male advantage in the rotational uncertainty effect of $d = .64$ ($F(1,70) = 7.36$, $p < .05$), i.e. a medium to large effect-size according to the definition of Cohen (1977).

3.4. Mental rotation

First, we calculated an ANOVA with RT itself as dependent variable with the between-subject factors "gender" and "stimulus material" and the within-subject factor "angular disparity". We obtained main effects of stimulus material ($F(4,350) = 118.77$, $p < .01$) and angular disparity ($F(2,700) = 708.38$, $p < .01$) as well as an interaction between stimulus material and angular disparity ($F(8,700) = 42.88$, $p < .01$). These effects are presented in Fig. 2, and are all in line with the literature. Most importantly, however, we obtained a three-way interaction between gender, stimulus material and angular disparity ($F(8,700) = 4.10$, $p < .01$) in addition to 2 two-way interactions between gender and stimulus material ($F(4,350) = 2.42$) and between gender and angular disparity ($F(2,700) = 3.12$, both $p < .05$). As can be seen from Fig. 2, when the data were analysed separately for the five stimulus materials, only with polygons an interaction between gender and angular disparity ($F(2,140) = 8.95$, $p < .01$) was present. No such effect was observable for the other four stimulus materials ($F(2,140) = 0.36$, 1.08, 1.09, and 0.27, respectively, all $p > .30$).

The ANOVA with mental rotation speed as dependent variable validated the results presented above with RT as dependent variable. We obtained a main effect of stimulus material ($F(4,350) = 89.32$, $p < .01$). Mental rotation speed was fastest with characters, slower with animal drawings, then slower with PMA symbols and polygons, which did not differ from each other, and finally slowest with cube figures (see Fig. 1).

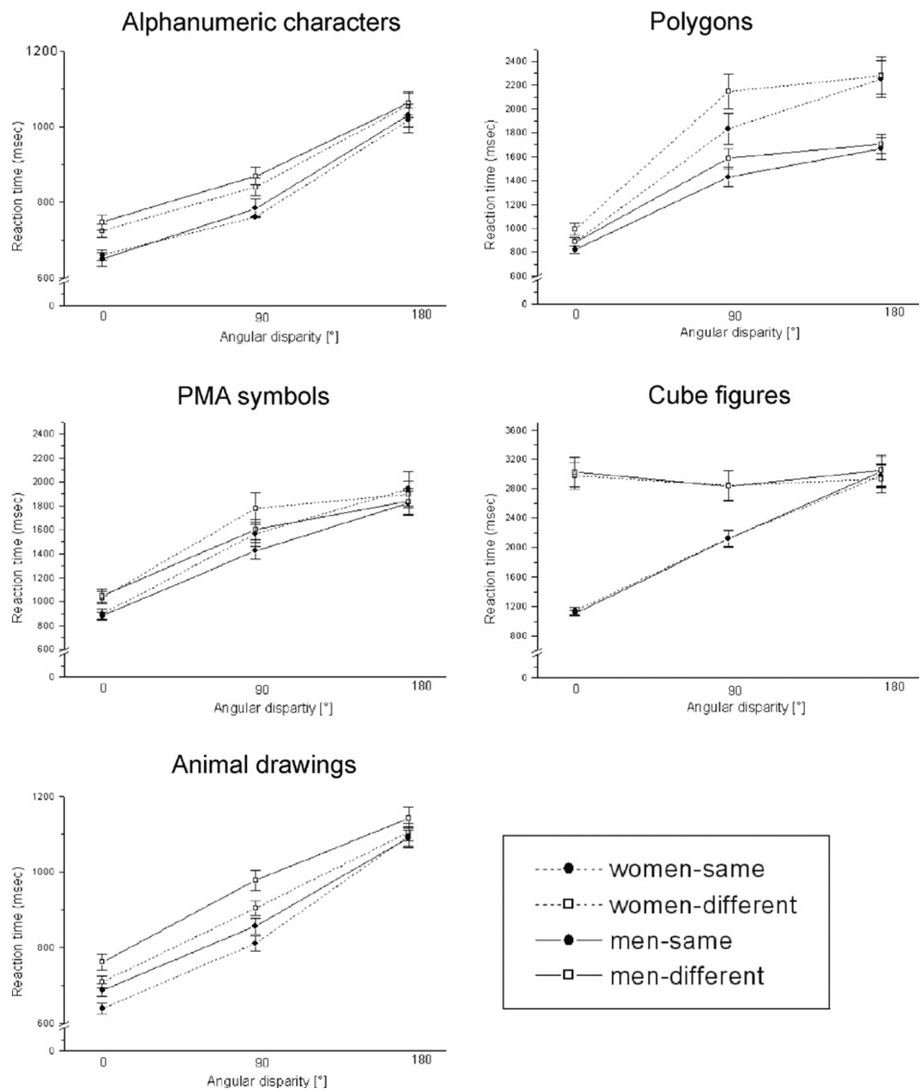

Fig. 2. Mean RT (in ms) as a function of angular disparity in the SU condition, separately for women and men and same and different response as a function of the stimulus material used. Error bars indicate standard errors.

Most importantly, a two-way interaction between gender and stimulus material was observed ($F(4, 350) = 3.18$, $p < .05$). The gender effect-sizes for characters, PMA symbols, animal drawings and cube figures, respectively, were low, never exceeded $d = .15$, and were not significantly different from 0, see Fig. 1. Polygons, however, revealed a male advantage in the speed of mental rotation of $d = 73$ ($F(1, 70) = 9.49$, $p < .05$), i.e. a large effect-size according to the definition of Cohen (1977).

4. General discussion

First of all, our results show that internal consistencies obtained with RT measures were substantial, irrespective of the stimulus material used. The internal consistency values for the perceptual comparison time (as a mean RT) as well as for the mental rotation speed (as the inverse of the slopes of the regression lines relating RT to angular disparity) ranged between .91 and .99, values comparable in magnitude to standard IQ test scores that allow individual diagnostics. Even the internal consistencies for the rotational uncertainty effects, ranging between .51 and .82, were satisfying, particularly in light of the fact that these scores reflect difference values. To sum up, the internal consistencies validated an information processing approach aimed at elucidating the mechanisms that account for gender differences in mental rotation performance.

Second, the results obtained in our study are straightforward:

4.1. Gender difference in visual comparison speed, maintaining readiness for mental rotation and mental rotation speed

In visual comparison speed, the gender differences were generally low, and the magnitude as well as the direction of the gender differences depended upon the stimulus material: No gender differences in perceptual comparison time at all were obtained with characters, PMA symbols and polygons, respectively. Whereas women outperformed men by about half a standard deviation in perceptual comparison time with animal drawings, men tended to outperform women by about a .45 of a standard deviation with cube figures.

The rotational uncertainty effect itself was significant in all five stimulus conditions, i.e., RT to upright stimuli embedded with non-upright stimuli was significantly longer than RT to the same upright stimuli in blocks containing only upright stimuli. We did not find gender differences in maintaining readiness for mental rotation using characters, PMA symbols, animal drawings, or cube figures as stimulus material, respectively. The maintaining readiness costs differed between genders only with polygons as stimuli. Men outperformed women by about 0.64 standard deviations. Thus, polygons resulted in equal perceptual comparison time for men and women but a smaller rotational uncertainty effect for men. Animal drawings, however, resulted in a faster perceptual comparison time for women than for men but in a comparable rotational uncertainty effect. A gender comparison of intercept differences, significant both for polygons and for animal drawings, therefore, is misleading if intercept differences are interpreted as differences in perceptual comparison time. Smaller intercepts in the mental rotation task can be a result of either a faster perceptual comparison time (for women with animal drawings) or smaller switching costs (for men with polygons).

The speed of mental rotation did not differ significantly between genders for most of the stimuli used, i.e., neither for characters, nor for PMA symbols, nor for animal drawings, nor for cube figures. This is a remarkable finding. In all of these cases, the effect-size of the non-significant gender differences never exceeded a value of $d = 0.2$. Given a desired level of $\alpha = .10$, and $1 - \beta = .80$, a completely unrealistic sample size of $N = 620$ would be needed to detect an effect of size $d = 0.20$. Thus, gender differences in the speed of mental rotation for characters, PMA symbols, animal drawings, or cube figures are small indeed. One might argue that the findings regarding characters, PMA symbols and animal drawings are in line with the literature (e.g., Jansen-Osmann and Heil, 2006; Waber, Carlson, and Mann, 1982) but that the findings regarding cube figures (the only 3D stimuli used) are at odds with the literature, but in fact, this is not the case. Peters (2005) using cube figures that, in contrast to the present study, were even rotated in depth, found no gender differences in the speed of mental rotation in a (selected) sample where men outperformed women in the MRT by as much as $d = 1.5$! Thus, gender differences in the speed of mental rotation do not account for gender differences in MRT scores.

4.2. Gender differences are significant with polygons as stimulus material

Polygons were the only material tested in our work that produced large effect-sizes according to the definition of Cohen (1977). Men outperformed women in speed of mental rotation with polygons by 73% of a standard deviation. It is not obvious, however, what the difference makes between polygons and the other stimulus material used in our experiments.

Moreover, the data do not support the hypothesis put forward by Stumpf (1993) and by Collins and Kimura (1997), that the size of the gender differences in mental rotation should be understood as a function of the difficulty of the mental rotation process. If the size of the gender difference is a function of task difficulty, then the largest gender differences should be found for 3D cube figures that produced mental rotation speeds about half as fast as those for polygons.

Polygons, in fact, were the only stimuli that revealed gender effects in mental rotation speed, and at the same time, were the only stimuli that revealed gender differences in the rotational uncertainty effect. Ilan and Miller (1994), however, as well as Jansen-Osmann and Heil (2006, in press) provided evidence that these two aspects are independent in that these authors were able to double-dissociate the two measures. We correlated the mental rotation speed with the rotational uncertainty effect, separately for men and women, as well as for the whole sample. For women, the correlation was negative ($r = -.40$, $p < .05$), i.e. the larger the interference the slower the mental rotation speed. The respective correlations were not significant for men, and also not significant for the whole sample (both $r < -.20$, $p > .10$). The reason for the special status of polygons with respect to both mental rotation speed and the rotational uncertainty effect is not clear at all at the moment, and further studies are needed.

One line of research might explore whether the effects observed in our experiments are or are not dependent upon the mental rotation *paradigm*, i.e., the simultaneous presentation paradigm used in this study, where two figures to be compared are presented side by side. Beside this, two other paradigms are used in the mental rotation literature, the successive presentation paradigm and the preparation rotation paradigm. In the successive presentation paradigm, the standard and the comparison figure are presented successively. The participants study the standard, and after a button press, the standard is replaced by the comparison figure which then has to be judged. In the preparation rotation paradigm, the standard figure is shown first, followed by an arrow specifying the orientation of the forthcoming comparison figure. Participants must mentally rotate the standard before they press a button, which triggers the replacement of the arrow by the comparison stimulus. The comparison is usually shown only for a short period of time to ensure that participants indeed execute the mental rotation process before requesting the comparison figure. Theoretically, mental rotation speeds estimated with the three different paradigms should not differ, but in reality, they do (see Cohen and Kubovy, 1993). Thus, one line of research should validate whether or not the presence of gender differences in mental rotation speed with polygons and their absence with the other stimulus materials is independent of the paradigm used.

When compared across the five types of stimuli, we found no evidence for an increase of gender differences in mental rotation speed with increasing task difficulty when task difficulty was defined as perceptual comparison time or as mental rotation speed. It is still possible however, that difficulty within one stimulus type might affect the size of the gender differences. Folk and Luce (1987) provided evidence that with polygons used as stimuli, mental rotation speed decreased with increasing complexity of the polygons. It would be interesting to evaluate whether or not gender differences in mental rotation speed (and, actually, also in the rotational uncertainty effect) would increase as a function of the complexity of the polygons, and whether such an effect would also be found for the other stimulus material.

Finally, in our study, the mental rotation task with the 3D cube figures was restricted to picture plane rotations. Thus, based on our own data, we cannot yet answer the question of whether or not depth rotation of 3D cube figures would have resulted in gender differences in mental rotation speed. As mentioned above, however, Peters (2005; see also Voyer, Butler, Cordero, Brake, Silbersweig, Stern & Imperato-McGinley, 2006, for comparable results) found no effect of gender on mental rotation speed for 3D cube figures rotated in depth.

4.3. Implications regarding the mechanisms responsible for gender differences in paper–pencil tests of mental rotation

Gender differences in paper–pencil mental rotation tests are substantial and reliable (see, e.g., Voyer et al., 1995), but neither the underlying causes nor even the performance mechanisms that yield these differences are understood. It is obvious that the data presented cannot provide conclusive answers as to which factor(s) are important as causal explanations for gender differences in mental rotation performance. Neither the more psychosocial nor the more biological-neuronal approaches can easily explain why in the present study, polygons evoked substantial gender differences in the speed of mental rotation, but alphanumeric characters, PMA symbols, animal drawings, and 3D-cube figures did not.

With respect to the mechanisms, however, the present data are in line with recent research suggesting that performance factors are not responsible (e.g., Masters, 1998; Peters, 2005). Allowing double the time for the MRT or abandoning time limits at all did not eliminate gender differences in these studies. Manipulating the time limits, however, constitutes an indirect way of examining the question whether or not the speed of mental rotation itself is responsible for the gender differences in test performance. In line with Peters (2005, Experiment 3), however, the present data clearly suggest that whatever the mechanisms yielding gender differences in test performance are, the speed of mental rotation itself might not be one of them.

5. Conclusions

On the basis of our present results, we can conclude (1) that gender differences in visual comparison speed are small and depend on the stimulus material both with respect to magnitude and with respect to direction,(2) that gender differences in mental rotation *speed* and in maintaining readiness for mental rotation are too small to deserve attention when characters, PMA symbols, animal drawings or cube figures are used as stimuli and (3) that gender differences of medium to large size according to Cohen (1977), however, were obtained with polygons for mental rotation speed *and* for the rotational uncertainty effect. Finally, we hope that our data will help to modify convictions: Substantial gender differences in the paper–pencil MRT (Vandenberg and Kuse, 1978, see Voyer et al., 1995) and the Collins and Kimura (1997) test constitute an empirical reality. The generalisation that men outperform women in the speed of mental rotation, however, was not supported in the present experiment.

References

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: a meta-analysis. *Sex Roles, 20*, 327–344.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Revised ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, D., & Kubovy, M. (1993). Mental rotation, mental representation, and flat slopes. *Cognitive Psychology, 25*, 351–382.

Collins, D. W., & Kimura, D. (1997). A large sex difference on a two-dimensional mental rotation task. *Behavioral Neuroscience, 111*, 845–849.

Cooper, L. A. (1976). Demonstration of a mental analog of an external rotation. *Perception and Psychophysics, 19*, 296–302.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing*. Oxford, England: Academic.

Delgado, A. R., & Prieto, G. (1996). Sex differences in visuo-spatial ability: do performance factors play such an important role? *Memory & Cognition, 24*, 504–510.

Donders, F. C. (1969). On the speed of psychological processes. *Acta Psychologica, 30*, 412–431, Original work published 1868.

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

Epting, L. K., & Overman, W. H. (1998). Sex-sensitive tasks in men and women: a search for performance fluctuations across the menstrual cycle. *Behavioral Neuroscience, 112*, 1304–1317.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: a general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*, 1–11.

Folk, M. D., & Luce, R. (1987). Effects of stimulus complexity on mental rotation rate of polygons. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 395–404.

Goldstein, D., Haldane, D., & Mitchell, C. (1990). Sex differences in visual-spatial ability: the role of performance factors. *Memory & Cognition, 18*, 546–550.

Halpern, D. F. (1992). *Sex differences in cognitive abilities*. Hillsdale, NJ: Erlbaum.

Heil, M., Bajric, J., Rösler, F., & Hennighausen, E. (1997). A rotation aftereffect changes both the speed and the preferred direction of mental rotation. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 681–692.

Heil, M., Rösler, F., Link, M., & Bajric, J. (1998). What is improved, if a mental rotation task is repeated—Efficiency of memory access or the speed of a transformation routine? *Psychological Research, 61*, 99–108.

Ho, H. Z., Gilger, J. W., & Brink, T. M. (1986). Effects of menstrual cycle on spatial information-processes. *Perceptual and Motor Skills, 63*, 743–751.

Hooven, C. K., Chabris, C. F., Ellison, P. T., & Kosslyn, S. M. (2004). The relationship of male testosterone to components of mental rotation. *Neuropsychologia, 42*, 782–790.

Ilan, A. B., & Miller, J. (1994). A violation of pure insertion: mental rotation and choice reaction time. *Journal of Experimental Psychology: Human Perception and Psychophysics, 20*, 520–536.

Imperato-McGinley, J., Pichardo, M., Gautier, T., Voyer, D., & Bryden, M. P. (1991). Cognitive abilities in androgen insensitive subjects - Comparison with control males and females from the same kindred. *Clinical Endocrinology, 34*, 341–347.

Jansen-Osmann, P., & Heil, M. (in press). Maintaining readiness for mental rotation interferes with perceptual processes in children but with response selection in adults. Acta Psychologica.

Jansen-Osmann, P., & Heil, M. (2006). Violation of pure insertion during mental rotation is independent of stimulus type, task, and subjects' age. *Acta Psychologica, 122*, 280–287.

Jolicoeur, P., Regehr, S., Smith, L. B. J. P., & Smith, G. N. (1985). Mental rotation of representations of two-dimensional and three-dimensional objects. *Canadian Journal of Psychology, 39*, 100–129.

Jordan, K., Heinze, H. J., Lutz, K., Kanowski, M., & Jäncke, L. (2001). Cortical activations during the mental rotation of different visual objects. *NeuroImage, 13*, 143–152.

Kail, R., Carter, P., & Pellegrino, J. W. (1979). The locus of sex differences in spatial ability. *Perception and Psychophysics, 26*, 182–186.

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

Loring-Meier, S., & Halpern, D. F. (1999). Sex differences in visuo-spatial working memory: components of cognitive processing. *Psychonomic Bulletin & Review, 6*, 464–471.

Masters, M. S. (1998). The gender difference on the mental rotations test is not due to performance factors. *Memory & Cognition, 26*, 444–448.

McGee, M. G. (1982). Spatial abilities: The influence of genetic factors. In M. Potegal (Ed.), *Spatial abilities: Development and physiological foundations* (pp. 199–222). New York: Academic Press.

McGlone, J. (1980). Sex differences in human brain asymmetry: a critical survey. *The Behavioral and Brain Sciences, 3*, 215–263.

Oswald, W. D., & Roth, E. (1987). *Der Zahlen-Verbindungs-test ZVT*. Göttingen: Hogrefe.

Peters, M. (2005). Sex differences and the factor of time in solving Vandenberg und Kuse mental rotation problems. *Brain & Cognition, 57*, 176–184.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg & Kuse Mental Rotations Tests: Different versions and factors that affect performance. *Brain and Cognition, 28*, 39–58.

Petrusic, W. M., Varro, L., & Jamieson, D. G. (1978). Mental rotation validation of two spatial ability tests. *Psychological Research, 40*, 139–148.

Reitan, R. M. (1956). *Trail making test. Manual for administration, scoring, and interpretation*. Indianapolis, IN: Indiana University Press.

Resnick, S. M. (1993). Sex differences in mental rotations: An effect of time limits? *Brain & sCognition, 21*, 71–79.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object Pictorial set: the role of surface detail in basic level object recognition. *Perception, 33*, 217–236.

Sanders, B., & Soares, M. P. (1986). Sexual maturation and spatial ability in college students. *Developmental Psychology, 22*, 199–203.

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: shapes of states. *Cognitive Psychology, 1*, 1–17.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychological Science, 10*, 80–83.

Signorella, M. L., & Jamison, W. (1986). Masculinity, femininity, androgyny, and cognitive performance: A meta-analysis. *Psychological Bulletin, 100*, 207–228.

Stumpf, H. (1993). Performance factors and gender-related differences in spatial ability: Another assessment. *Memory and Cognition, 21*, 828–836.

Thurstone, T. G. (1958). *Manual for the SRA primary mental abilities*. Chicago: Science Research Associates.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotation, a group test of three-dimensional spatial visualisation. *Perceptual and Motor Skills, 47*, 599–604.

Voyer, D., & Bryden, M. P. (1990). Gender, level of spatial ability, and lateralization of mental rotation. *Brain & Cognition, 13*, 18–29.

Voyer, D., Butler, T., Cordero, J., Brake, B., Silbersweig, D., Stern, E., et al. (2006). The relation between computerized and paper-and-pencil mental rotation tasks: a validation study. *Journal of Clinical and Experimental Neuropsychology, 28*, 928–939.

Voyer, D., & Saunders, K. A. (2004). Gender differences on the mental rotations test: a factor analysis. *Acta Psychologica, 117*, 79–94.

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

Waber, D. P., Carlson, D., & Mann, M. (1982). Developmental and differential aspects of mental rotation in early adolescence. *Child Development, 53*, 1614–1621.

Wiedenbauer, G., Schmid, J., & Jansen-Osmann, P. (2007). Manual training of mental rotation. *European Journal of Cognitive Psychology, 19*, 17–36.