

Where are the Gender Differences? Male Priming Boosts Spatial Skills in Women

Tuulia M. Ortner · Monika Sieverding

Published online: 1 May 2008
© Springer Science + Business Media, LLC 2008

Abstract The effects of gender stereotype activation by priming on performance in a spatial task were investigated among a mixed adult sample (including students) of 161 men and women (mean age=31.90) from Austria (Europe). They were assigned to one of four experimental groups according to gender and stereotype activation condition. After a male or female gender stereotype activating task, participants worked on a test assessing mental rotation (three-dimensional cube test, Gittler 1990). A significant main effect of priming on the performance in the mental rotation task emerged. Cohen's d showed a pronounced gender difference emerging only in the *female* priming condition ($d=.59$), whereas it disappeared in the *male* priming condition ($d=.01$).

Keywords Gender priming · Gender differences · Spatial skills · Mental rotation · Test fairness

Introduction

Gender stereotypes exist in many domains, including personality factors and intellectual performance (Costa et

al. 2001; Furnham et al. 1999; Halpern 1992). With regard to personality attributes, men are usually seen as more independent, assertive, and competitive, while women are seen as more sensitive, expressive, and tender-minded (Deaux and LaFrance 1998; Feingold 1994; Williams and Best 1990). These gender-related personal attributes mirror and reproduce the typical social roles as the male breadwinner and the female caregiver (Eagly 1987; Eagly et al. 2000). With regard to intellectual performance, for most aspects, the male gender stereotype has more of an advantage as compared to the female stereotype. Men are usually seen as more intelligent by others, and generally estimate themselves as more intelligent than women in most domains (Bornholt et al. 1994; Furnham et al. 1999; Rammstedt and Rammsayer 2000). Stereotypes have been discussed as resulting from a gender-related division of labor in former times, when people lived under different social conditions.

Today, stereotypes are being used to rationalize and legitimize existing gender differences in role and status (Hoffman and Hurst 1990). In addition to this, within social interaction, stereotypes can also serve as a basis for self-fulfilling prophecies (Deaux and Major 1987). Stereotypes can therefore not only affect one person's expectations and thoughts but also his or her behavior — for example, achievement test performance (Deaux and Major 1987; Dijksterhuis and van Knippenberg 1998; Steele 1997). However, gender stereotypes are not universal throughout societies and ethnicities; they differ across cultures (Best 2001; Durik et al. 2006; Gibbons et al. 1991; Harris 1994; Kaufman 2006). Variations of gender stereotypes have been shown to be affected by variables such as religion and urbanization, as well as by the magnitudes of the existing gender gap and power distance between genders found in the societies (Désert and Leyens 2006; Eriksson and

T. M. Ortner (✉)
Department of Psychology,
Division for Psychological Assessment,
Humboldt University Berlin,
Unter den Linden 6,
10099 Berlin, Germany
e-mail: tuulia.ortner@psychologie.hu-berlin.de

M. Sieverding
Department of Psychology,
Division for Differential and Gender Psychology,
University of Heidelberg,
Heidelberg, Germany

Lindholm 2007; Frieze et al. 2003; Williams and Best 1990). A recent study also indicated that positive economic conditions are related to more egalitarian gender role attitudes (Olson et al. 2007). Gender differences in self-construals have also generally been shown to be stronger in Western cultures. This has been interpreted as a product of self-stereotyping, that occurs when between-gender social comparisons are made, which are more likely, and exert a greater impact in Western nations (Guimond et al. 2007).

The purpose of this study was to investigate the effects of gender stereotype activation on intellectual performance. We investigated the effects of the activation of a female stereotype on spatial performance of men and women and compared these effects to those of male stereotype activation. We expect new results with reference to the question how activation of stereotypes influences intellectual performance. We also seek to determine the manner in which culture-specific stereotypes influence discrepancies between the performances of men and women.

The present study focuses on performance in the domain of intelligence—spatial skills—which is one of the best-investigated constructs in this area, and the one with the largest gender differences. The male advantage in some types of spatial skills has been established in two meta-analyses (Linn and Petersen 1985; Voyer et al. 1995). Gender differences in spatial skills have been observed in numerous studies in different countries — for example, in the USA and China (Geary and DeSoto 2001); in the UK (Peters et al. 2007); in Israel (Birenbaum et al. 1994); in France (Guillot et al. 2007); as well as in Austria (Weiss et al. 2003). Although gender differences were found in different countries, intercultural studies raised the question of the stability of gender differences in spatial abilities throughout different cultures (e.g. Flaherty 2005).

It has mostly been proposed that spatial skills can be classified into three main categories, which have been shown to differ with regard to the size of gender differences (Linn and Petersen 1985): (1) *Spatial visualization* involves the ability to process spatial information (as understanding relationships between objects in space). Gender differences do not always appear here; if they do, they are small (effect sizes in the meta-analyses of Linn and Petersen: $d=.13$, Voyer et al.: $d=.19$). (2) *Spatial perception* includes the ability to identify and locate judgments of horizontal and vertical orientation in space when distracting information is present. This has disclosed large but rather inconsistent gender differences (Halpern 1992; Signorella and Jamison 1986; Voyer et al. 1995. Effect sizes: $d=.44$ in both meta-analyses). (3) *Mental rotation* describes the ability to imagine the transformation of a multidimensional figure as rotated in space and has revealed the largest and most consistent pattern of gender differences (effect sizes Linn and Peterson 1985: $d=.73$; Voyer et al. 1995: $d=.56$). Men

were mostly identified as being able to rotate objects more accurately (Luehring and Altman 2000; Masters 1998; Oosthuizen 1991; Resnick 1993) and faster (Kail et al. 1979; Petrusic et al. 1978) than women. This often-replicated male advantage is of particular relevance, as the use of spatial abilities, and especially mental rotation, is seen as an important precondition to achieving success in lucrative, high-status professions, such as architecture or engineering (Rasanen 1991). However, these effects were not shown to be stable across cultures as often assumed. Besides a number of biological factors (Kimura 1992; Newcombe and Bandura 1983; Sanders and Soares 1986), a number of factors strongly related to the social situation, as well as cultural preconditions, have been proposed as possible determinants of these gender differences (Voyer et al. 2004). The choice of activities and training (Baenninger and Newcombe 1989), as well as gender-role identification (Hamilton 1995; Signorella and Jamison 1986), and different problem-solving strategies (Lawton 1994), were identified as relevant factors. All of these sociocultural or educational factors are strongly associated with existing cultural circumstances. The fact that gender differences have decreased in magnitude over the preceding decades has also been explained by changes in gender roles and related changes in educational practices (Feingold 1988; Voyer et al. 1995). Several researchers have emphasized the importance of the social and cultural context of assessment in increasing or decreasing psychological gender differences (see Hyde 2005 for a review).

Two related psychological mechanisms have recently been explored as influencing gender differences in intellectual performance: *stereotype threat* and *stereotype priming*. Situations where stereotypes negatively affect the target person may lead to stereotype threat, an impairment of task performance (Spencer et al. 1999; Steele 1997). Some studies investigating American college students found that an instruction that deemphasizes the spatial character of the task can reduce or even eliminate gender differences in performing mental image rotation tasks (Scali et al. 2000; Sharpset al. 1994). On the other hand, there is stereotype priming, which can be seen as the combination of two possible processes, depending on group membership and contextual factors: First, it leads to a “cold” cognitive pathway process for both nontarget and target persons, resulting in behavioral patterns in line with a given stereotype. Therefore, stereotype-related priming has been shown to have an impact on intellectual performance, even if persons did not belong to the group primed on (Dijksterhuis and van Knippenberg 1998). Priming may additionally cause “hot” processes in targets, which enhance the negative effect (Wheeler and Petty 2001) related to stereotype threat. These behavioral effects are not necessarily provoked by the *active* activation of stereotypic

traits, they can also be caused by subliminal presentation (Bargh and Chartrand 1996).

The effects of stereotype priming have been shown for several cognitive domains in samples including individuals from different western societies — e.g., for math performance in a sample of American students (Wheeler et al. 2001) or memory (Dijksterhuis et al. 2000) and general knowledge testing in a sample of undergraduate students in the Netherlands (Dijksterhuis and van Knippenberg 1998). However, stereotype priming does not necessarily cause a decrease in the individual's performance. Some studies testing students in the USA indicate that performance can also be *shifted* as an effect of priming (Shih et al. 1999). Asian-American women primed with their Asian identity performed better on a math test, while participants primed with their female identity showed a decrease in performance, compared with women in a control group (Shih et al. 2002).

In comparison with other domains such as mathematical or verbal abilities, very few studies deal with the effects of stereotype activation on spatial abilities: a study of McGlone and Aronson (2006) conducted in the USA tested multiple relevant identities and their impact within a priming condition on spatial reasoning. They found that women who were primed with their identity as students at a selective private college performed better in a mental rotation task than those who were primed with their gender or a test-irrelevant identity. Although the impact of negative stereotype effect decreased within the priming conditions not explicitly referring to gender, the differences within the conditions between men and women were still given in considerable amounts (gender condition: $d=3.24$; private college student condition: $d=1.00$; control condition: $d=1.42$). This may be attributed to the fact that the identity as a student may have implied facets of gender: For women, the identity as a college student is not necessarily neutral with regard to gender.

Our research question was whether the priming of women with a male stereotype (quite distinct from self) can have even more clear effects on their performance in a mental rotation test. We therefore focused on the impact of a clear male and female stereotype priming on spatial reasoning applied to men and women. The study was conducted in a country which shows a social situation with particular strong gender role allocation: In Austria, in academic fields, only 9.5% of total Grade A professorships are held by women (European Commission 2006). In the economic sector, only 5% of the members of directorates of the largest publicly quoted companies are women, which gives Austria a position at the bottom area of an overall listing within the European Union (e.g. compared to other European countries: Norway: 34%; Germany 11%; France: 9%; European Commission 2007). The strong traditional

understanding of gender roles is also demonstrated by a recent survey, which revealed that 67% of the Austrian population are in the opinion that women are supposed to stay at home, care for the children, and not be employed (Dittmann and Scheuer 2007).

We chose an experimental design with two independent variables and one dependent variable: men and women assigned to one of two implicit gender-stereotype activating tasks (male or female) established the independent variables. As the dependent measure, we used the score gained by a mental rotation task, which was applied after the gender-stereotype activating task. According to the *ideomotor perspective* of priming (Dijksterhuis and Bargh 2001), an activated stereotype schema includes behavioral representations and can have behavioral effects due to an assumed *perception-behavior link*. This is supposed to be based on our innate propensity to imitate others.

Our general assumption was that gender stereotype priming has a higher impact than gender (which means: being male or female) in predicting performance in a mental rotation test (main effect). As a first hypothesis, we therefore expect that advantageous male gender-stereotype activation leads to a shift in spatial performance, while the disadvantageous female stereotype association leads to lower performance.

According to additional “hot processes” in individuals, who are affected as potential targets of stereotype threat (Wheeler and Petty 2001), we expect stronger effects of lowered performance in the female gender-stereotype activation condition on women's performance than on men's (second hypothesis; interaction effect of gender and priming condition).

In a third step, we compared the resulting gender differences in performance within two different conditions. Due to the theory and results described above (Dijksterhuis and Bargh 2001; Wheeler and Petty 2001), we expect as a third hypothesis a gender difference (d) in the comparison of men's performance in the male stereotype-activation condition with women's performance in the female stereotype activating condition (both genders confronted with the consistent stereotype). Comparing women's and men's performance within the *same* stereotype activating condition, we expect the gender difference (d) to be diminishing in the male priming condition for men and women.

Method

Overview

In an experimental design, men and women were primed either with a typical male or a typical female stereotype before they underwent a mental rotation test. All tests were

conducted within 5 days in mixed gender group settings in the same large conference room situated in the center of Vienna. Attention was drawn to the fact that all groups consisted largely of the same amount of men and women. Every test was conducted by two out of five female students who gave standardized instructions (one instructed and one assisted).

Participants

One hundred and sixty-one voluntary participants (87 men aged 18 to 57, $M=30.44$, $SD=10.43$, and 74 women aged 18 to 60, $M=33.34$, $SD=12.63$) were recruited by five students of Psychology. The requirement was that the participants were not themselves studying Psychology. According to their educational level, 38.5% of the participants had no school leaving certification (40.2% of the men and 36.5% of the women), 61.5% had gained general qualification for entrance to a university. Some of the participants were students. The participants were assigned to one of two stereotype-activation task-related experimental groups by drawing lots: Group 1 consisted of women given a male gender-stereotype activating task ($n=39$); the Group 2 women were given a female gender-stereotype activating task ($n=35$). Men given a male gender-stereotype activating task formed Group 3 ($n=45$) and men given a female gender-stereotype activating task formed Group 4 ($n=42$).

Materials

Gender-stereotype activating-task. For stereotype activation, two texts about a day in the life of a man and a woman were designed, including typical gender stereotypes. Based on the cross-cultural studies from Williams and Best (1990), adjectives and characteristics typically seen for each gender were used and embedded into the two different texts covering a half standard page. Only positive adjectives were used to avoid negative feelings and perceptions of the persons described within the texts. The text of the female gender-stereotype activating-task dealt with a day in the life of a “typical woman:” she is described as someone who takes care of her family; works part time; shows herself to be insightful, helpful, and agreeable; and meets friends for a chat in a café and often day-dreams. The male gender-stereotype activating-task dealt with characteristics that are typically associated with men: the protagonist is, for example, described as a self-confident and tough-minded person, who drives a motorbike, works in a leading position and reprimands assistants, takes risks, and does weight training after work. The texts were handed out as “a task to measure the ability to put oneself in someone else’s position.” The question posed after the text was: “If you were the person

described in the text, which adjectives would you use to describe yourself?”.

Mental rotation test. A short form of the three-dimensional cube test (3DC; Gittler 1990) consisting of the first nine items (out of 17 in the long form) was used. The correlation of the short form with results gained from the long form is sufficient ($r=.94$). The 3DC is a paper pencil test and has been described as one of the best investigated tests in general and has been praised for its thoroughly investigated validity and scaling (Fay 1996). In the test, the subjects have to mentally rotate/transform six trial cubes, which have different patterns on their surfaces and decide if one of them matches the target cube. (To avoid disadvantageous effects due to guessing, there is additionally the choice of selecting none of the cubes as right or choosing the option “I do not know the solution”.) No time limitations are set. An item example is given in Fig. 1. The 3DC obtains an advantageous scaling: since the 3DC fulfills the criteria of the Rasch Model (Rasch 1960), it has the property of unidimensionality and therefore measures the same latent cognitive dimension (ability) in different groups of subjects. The raw scores are calculated as the total amount of solved items. Results gained by different long and short versions of the test can be compared through the estimation of person parameters on the basis of the difficulty parameters of the presented items. The test has already been used for measurement of mental rotation within different international published studies (Weiss et al. 2003; Windischberger et al. 2003).

Gittler (1990) reports consistent and robust gender differences for the probability of finding the solution for every given item within the 3DC, which conforms with the theoretical implications of the scaling. Effect sizes of gender differences calculated on the basis of the given norms beyond the age of 18 were $d=.64$ ($n=432$) for general student norms.

For comparison with a sample without experimental treatment, we compared our results to one of the samples described by Gittler (1990). We chose the sample which showed the highest similarity with our sample (220 males,

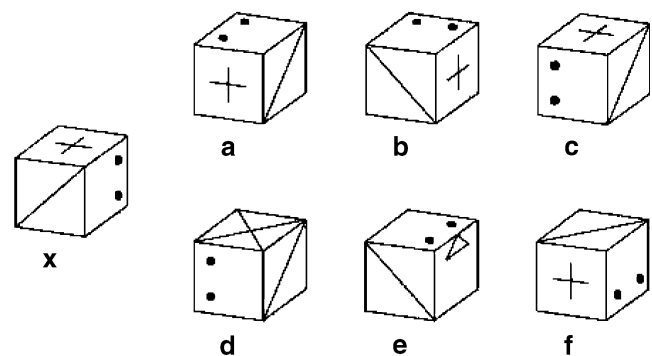


Fig. 1 Item example from the 3DC (Gittler 1990).

age $M=22.48$, $SD=4.05$, and 212 females, aged $M=22.46$, $SD=4.44$, all college students). The average raw scores of the version used in our study and the version of Gittler cannot directly be compared since we used a short form. According to the model from Rasch, an option is to calculate ability parameters independent from the amount of the presented items, if item parameters are given. We therefore compared the average performance on the level of person parameters.

Procedure

Persons were welcomed and general information was given about the duration and content of the study. It was explained that two tasks would have to be completed: the first would address the ability to put oneself in someone else's position and the second would measure a component of aptitude. Persons were not informed about the existence of different conditions. After the persons were assigned to one of two experimental groups by drawing lots, work sheets were handed out on the basis of group affiliation. The work sheet started with one of the two gender-stereotype activating-tasks describing the stimulus person with adjectives. Participants were instructed to start the mental rotation test by themselves once they had finished the first task by moving to the next page. The test instruction for the mental rotation test started with task examples. Participants were also instructed to leave the test room once they had finished. An assistant waited outside the test room and informed them about the aim of the study.

Statistical Analysis

For the calculation of group differences, ANOVA was performed, taking gender and primed gender into the analyses as fixed factors. For an additional evaluation of differences between the groups, simple main effects and Cohen's d were calculated to compare the magnitude of gender differences in the different conditions.

Results

Within the four groups, the following raw scores were obtained for the 3DC: women primed female $M(SD)=3.9$ (2.85); women primed male $M(SD)=5.5$ (2.94); men primed female $M(SD)=5.1$ (3.29); and men primed male $M(SD)=5.5$ (2.80).

A 2 (gender) \times 2 (primed gender) ANOVA revealed a main effect for primed gender, $F(1, 157)=4.59$, $p=.03$. This corroborates our first hypothesis that the advantageous male gender-stereotype activation leads to a higher spatial performance, while the disadvantageous female stereotype association leads to lower performance. Neither the main

effect for gender nor the interaction (gender \times priming) was significant (the effect for gender was $F(1, 157)=2.00$, $p=.16$ and, for the interaction of gender and primed gender, it was $F(1, 157)=1.73$, $p=.19$). Therefore, the assumption of stronger effects of lowered performance in the female gender-stereotype activation condition on women's performance than on men's (second hypothesis) cannot be sustained as this has not been shown statistically. Table 1 shows the characteristics of the four groups. The mean performance score in the mental rotation task was the lowest for women primed with the female gender-stereotype activating task.

The lack of a significant interaction indicates that the general main effect for gender was not significantly different across both stereotype activation conditions. To more closely examine whether the overall effect is larger and more significant for the groups confronted with consistent stereotype than for the same male stereotype priming condition, we additionally calculated simple main effects between these as relevant hypothesized groups. We calculated this as more conservative tests to show whether the hypothesized differences in significance within the groups of interest are given. Additionally, we calculated Cohen's d as an indicator of the size of differences. So, taking into account the most *natural* condition with consistent stereotype activation, women primed female and men primed male, the effect size was $d=.59$ with $t(78)=2.64$, $p=.01$ favoring men. Referring to men and women tested under the *same* male condition, the following effects emerged: the effect size results as $d=.01$ with $t(82)=.07$, $p=.94$. According to the third hypothesis, the gender difference in performance diminished and nearly vanished, including the performance of both men and women under the male stereotype activation condition compared with the other condition. According to our presumptions, gender differences in performance are given for women primed female and men primed male and even vanish in the same male stereotype activation condition for men and women.

Comparing the obtained results to a sample without experimental treatment of Gittler (1990), Table 2 shows the

Table 1 Raw scores obtained for mental rotation performance as a function of gender and priming condition.

Group	M	SD
Women		
Primed female	3.86	2.85
Primed male	5.49	2.94
Men		
Primed female	5.14	3.29
Primed male	5.53	2.80

Means and standard deviations of raw scores (possible values range from 0 to 9).

Table 2 Average person parameters for mental rotation performance gained within the groups.

Source	Average Person Parameter
Women	
Primed female	-.3
Primed male	.5
(Gittler 1990)	(.1)
Men	
Primed female	.3
Primed male	.5
(Gittler 1990)	(.8)

Zero-sum standardized person parameter values have a mean of 0 and usually range from -3.0 to 3.0.

mean person parameters established for the mean scores and the (long form) scores within all groups: The results of women without experimental treatment therefore lie between the two priming conditions used in our experiment. The male results within our study were both lower than the results of Gittler (1990). This can be explained by the composition of Gittler's sample: while the student sample is composed of students from technical universities ($n=75$) and other universities, 90% of the technical sample are male. As the average person parameter of the technical university sample constitutes a value of 1.1, this most likely leads to a significant shift of the males within this unbalanced population sample by Gittler.

Discussion

The aim of the present study was to analyze the effects of gender and gender stereotype priming on performance of a mental rotation task. Women and men who were either primed with a traditional feminine or masculine stereotype worked on a short form of the three-dimensional cube test (3DC; Gittler 1990). Concordant with our main hypothesis, the effect of gender priming on performance was significant, whereas the main effect of gender was not significant. The second hypothesis, which assumed that the effect of gender priming would be stronger in women than in men, was not supported by our results. Gender priming applied to men and women showed the following effects on spatial abilities: persons of both sexes were affected by the negative female stereotype and showed lower task performance. The study showed that gender differences disappeared in male priming condition for women, whereas it had a pronounced effect on men primed with male content and on women primed with female content.

The findings of our study support the results of the research done by McGlone and Aronson (2006) and Sharps et al. (1994) — that gender differences in mental rotation tasks can be significantly reduced by the way in which the

task is embedded and presented. In many studies, stereotype activation was not manipulated explicitly: McGlone and Aronson (2006) primed persons on multiple identities by giving them a short questionnaire designed to make a particular identity more salient (e.g., “List three reasons why one might attend a private liberal arts college”) and tested them straight after. In the study by Sharps et al. (1994), men performed at higher levels than women if the spatial nature was mentioned within the instruction, but no sex differences were obtained with non-spatial instructions. The results of these studies as well as those of our study support the hypothesis that, in “normal” test situations addressing mental rotation abilities, a subtle or even obvious stereotype threat might be operating and particularly decrease the performance of women.

In contrast to the study of McGlone and Aronson (2006), the gender differences in mental rotation in our study not only decreased but vanished in one of the conditions (the male priming condition). An explanation for the poorer performance of women in the standard test situation might be caused by the instruction of the 3DC: it emphasizes the spatial nature of the task and might therefore evoke gender-stereotype-specific associations as in the study of Sharps et al. (1994). These effects might have been diminished within the male stereotype condition in our study.

However, the absence of a control group can be seen as a limitation of the present study: since we did not test a group without experimental treatment, we cannot make definitive conclusions on the effects of achievement under a gender-priming condition compared with a standard situation.

In our study, a new type of mental rotation test, the 3DC (Gittler 1990) was used with advantageous psychometric properties. Under standard conditions, this test shows consistent gender differences, although they are not as large in terms of effect size as the ones obtained from the Vandenberg Mental Rotation Test (MRT, Vandenberg 1971). For generalization of our results, future research should investigate whether the priming instructions used in this study can also decrease gender differences in other mental rotation tests, for example, the MRT.

Our results suggest that, in accordance to the social role theory (Eagly 1987) spatial skills of men and women depend on whether they act in accordance with their social roles, which are related to cultural conditions. If men empathize with women's part (or even live it), their spatial skills worsen. What about the consequences of the results in practice? A fair assessment of spatial performance may not be given under *standard* test-taking conditions: actual cues, thoughts, or environmental factors can be seen as triggers that influence the magnitude of gender differences in spatial performance due to their impact on the actual performance of females. The results of our study support existing evidence (McGlone and Aronson 2006; Sharps et al.

1994) that the disadvantages of women eventually can be overcome. Within this study, women were clearly able to perform the same level of mental rotation as men with low effort. However, for practical purposes, it does not seem possible at this point to give test takers a certain text about persons to read before taking a test: as there are ethical standards on informing persons about the procedure and purpose of tests, a priming condition, as applied here as a subtle strategy, which is not made transparent, would fail. Nevertheless, it might be a future strategy for coaching and training to inform female applicants about the enhancing effects of certain behaviors in order to enable them to use techniques, which could improve their performance. Generally, to improve the fairness of assessment procedures in the future, all persons should be given a better opportunity to show the abilities they possess.

Acknowledgement We would like to thank Melanie Fiegl, Michaela Kacvincka, Sandra Peyerl, Nicole Pritz and Sandra Raffetseder for their help with the data collection and two anonymous reviewers for their thoughtful comments.

References

- Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: a meta-analysis. *Sex Roles, 20*, 327–344.
- Bargh, J. A., & Chartrand, T. L. (1996). The unbearable automaticity of being. *American Psychologist, 54*, 462–476.
- Best, D. L. (2001). Gender concepts: convergence in cross-cultural research and methodologies. Cross-cultural research. *The Journal of Comparative Social Science, 35*, 23–43.
- Birenbaum, M., Kelley, A. E., & Levi-Keren, M. (1994). Stimulus features and sex differences in mental rotation test performance. *Intelligence, 19*, 51–64.
- Bornholt, L. J., Goodnow, J. J., & Cooney, G. H. (1994). Influences of gender stereotypes on adolescents' perceptions of their own achievement. *American Educational Research Journal, 31*, 675–692.
- Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of Personality and Social Psychology, 81*, 322–331.
- Deaux, K., & LaFrance, M. (1998). Gender. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology*. New York: Random House.
- Deaux, K., & Major, B. (1987). Putting gender into context: an interactive model of gender-related behavior. *Psychological Review, 94*, 369–389.
- Désert, M., & Leyens, J.-P. (2006). Social comparison across cultures I: Gender stereotypes in high and low power distance cultures. In S. Guimond (Ed.), *Social comparison and social psychology: Understanding cognition, intergroup relations and culture* (pp. 303–317). Cambridge, UK: Cambridge University Press.
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & Van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology, 36*, 531–544.
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–40). San Diego, CA: Academic Press.
- Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology, 74*, 865–877.
- Dittmann, J., & Scheuer, A. (2007). Berufstätigkeit von Müttern bleibt kontrovers [Occupation of mothers remains controversial]. *Informationsdienst Soziale Indikatoren, 38*, 1–5.
- Durik, A. M., Hyde, J. S., Marks, A. C., Roy, A. L., Anaya, D., & Schultz, G. (2006). Ethnicity and gender stereotypes of emotion. *Sex Roles, 54*, 429–445.
- Eagly, A. H. (1987). *Sex differences in social behaviour: A social role interpretation*. Hillsdale, NJ: Lawrence Erlbaum.
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: a current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Erlbaum.
- Eriksson, K., & Lindholm, T. (2007). Making gender matter: the role of gender-based expectancies and gender identification on women's and men's math performance in Sweden. *Scandinavian Journal of Psychology, 48*, 329–338.
- European Commission (2006). She figures 2006, women and science, statistics and indicators: European Communities.
- European Commission (2007). Database—social and economic domain: decision-making in the largest publicly quoted companies.
- Fay, E. (1996). *Tests unter der Lupe. Aktuelle Leistungstests-kritisch betrachtet [A closer examination of current performance tests from a critical point of view] (Band 1. Testrezension 3DW (S. 21–27))*. Heidelberg: Asanger.
- Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43*, 95–103.
- Feingold, A. (1994). Gender differences in personality: a meta-analysis. *Psychological Bulletin, 116*, 429–456.
- Flaherty, M. (2005). Gender differences in mental rotation ability in three cultures: Ireland, Ecuador and Japan. *Psychologia, 48*, 31–38.
- Frieze, I. H., Ferligoj, A., Kogovsek, T., Renner, T., Horvat, J., & Sarlija, N. (2003). Gender-role attitudes in university students in the United States, Slovenia and Croatia. *Psychology of Women Quarterly, 27*, 256–261.
- Furnham, A., Fong, G., & Martin, F. (1999). Sex and cross-cultural differences in the estimated multifaceted intelligence quotient score for self, parents and siblings. *Personality and Individual Differences, 26*, 1025–1034.
- Geary, D. C., & DeSoto, M. C. (2001). Sex differences in spatial abilities among adults from the United States and China: implications for evolutionary theory. *Evolution and Cognition, 7*, 172–177.
- Gibbons, L. J., Stiles, D. A., & Shkodriani, G. M. (1991). Adolescents' attitudes toward family and gender roles: an international comparison. *Sex Roles, 25*, 625–643.
- Gittler, G. (1990). *Dreidimensionaler Würfeltest (3DW) [Three-dimensional cube test]*. Beltz: Weinheim, Germany.
- Guillot, A., Champely, S., Batier, C., Thiriet, P., & Collet, C. (2007). Relationship between spatial abilities, mental rotation and functional anatomy learning. *Advances in Health and Sciences Education, 12*, 1573–1677.
- Guimond, S., Branscombe, N. R., Brunot, S., Buunk, A. P., Chatard, A., Desert, M., et al. (2007). Culture, gender, and the self: variations and impact of social comparison processes. *Journal of Personality and Social Psychology, 92*, 1118–1134.
- Halpern, D. F. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Hamilton, C. J. (1995). Beyond sex differences in visuo-spatial processing. The impact of gender trait possession. *British Journal of Psychology, 86*, 1–20.

- Harris, A. C. (1994). Ethnicity as a determinant of sex role identity: a replication study of item selection for the Bem Sex Role Inventory. *Sex Roles, 31*, 241–273.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: perception or rationalization. *Journal of Personality and Social Psychology, 58*, 197–208.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581–592.
- Kail, R., Carter, P., & Pellegrino, J. (1979). The locus of sex differences in spatial ability. *Perception and Psychophysics, 26*, 182–186.
- Kaufman, J. C. (2006). Self-reported differences in creativity by ethnicity and gender. *Applied Cognitive Psychology, 20*, 1065–1082.
- Kimura, D. (1992). Sex differences in the brain. *Scientific American, 267*, 119–125.
- Lawton, C. A. (1994). Gender differences in way-finding strategies: relationship to spatial ability and spatial anxiety. *Sex Roles, 30*, 765–779.
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterisation of gender differences in spatial abilities: a meta-analysis. *Child Development, 56*, 1479–1498.
- Luehring, J., & Altman, J. D. (2000). Factors contributing to sex differences in the mental rotation task. *Psi Chi, 5*, 39–35.
- Masters, M. S. (1998). The gender difference on the mental rotations test is not due to performance factors. *Memory and Cognition, 26*, 444–448.
- McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486–493.
- Newcombe, N., & Bandura, M. M. (1983). Effect of age at puberty on spatial ability in girls: a question of mechanism. *Developmental Psychology, 19*, 215–224.
- Olson, J. E., Frieze, I. H., Wall, S., Zdaniuk, B., Ferligoj, A., Kogovsek, T., et al. (2007). Beliefs in equality for women and men as related to economic factors in Central and Eastern Europe and the United States. *Sex Roles, 56*, 297–308.
- Oosthuizen, S. (1991). Sex-related differences in spatial ability in a group of South African students. *Perceptual and Motor Skills, 73*, 51–54.
- Peters, M., Manning, J. T., & Reimers, S. (2007). The effects of sex, sexual orientation, and digit ratio (2D:4D) on mental rotation performance. *Archives of Sexual Behavior, 36*, 251–260.
- Petrusic, W. M., Varro, L., & Jamieson, D. G. (1978). Mental rotation validation of two spatial ability tests. *Psychological Research, 40*, 139–148.
- Rammstedt, B., & Rammsayer, T. H. (2000). Sex differences in self-estimates of different aspects of intelligence. *Personality and Individual Differences, 29*, 869–880.
- Rasanen, L. (1991). Girls and the learning of physical concepts. *Finnish Journal of Education, 22*, 185–194.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.
- Resnick, S. M. (1993). Sex differences in mental rotations: an effect of time limits. *Brain and Cognition, 21*, 71–79.
- Sanders, B., & Soares, M. P. (1986). Sexual maturation and spatial ability in college students. *Developmental Psychology, 22*, 199–203.
- Scali, M., Brownlow, S., & Hicks, J. (2000). Gender differences in spatial task performance as a function of speed or accuracy orientation. *Sex Roles, 43*, 359–376.
- Sharps, M. J., Price, J. L., & Williams, J. K. (1994). Spatial cognition and gender. Instructional and stimulus influence on mental image rotation performance. *Psychology of Women Quarterly, 18*, 413–425.
- Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype performance boosts: the impact of self-relevance and the manner of stereotype activation. *Journal of Personality and Social Psychology, 83*, 638–647.
- 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.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4–28.
- Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.
- Vandenberg, S. G. (1971). *Mental rotation test*. University of Colorado: Boulder.
- Voyer, D., Rodgers, M. A., & McCormick, P. A. (2004). Timing conditions and the magnitude of gender differences on the mental rotation task. *Memory and Cognition, 23*, 72–82.
- 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.
- Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences, 35*, 863–875.
- Wheeler, S. C., Jarvis, W. B. G., & Petty, R. E. (2001). Think unto others: the self-destructive impact of negative stereotypes. *Journal of Experimental Social Psychology, 37*, 173–180.
- Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: a review of possible mechanisms. *Psychological Bulletin, 127*, 797–826.
- Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multinational study*. Newbury Park: Sage.
- Windischberger, C., Lamm, C., Bauer, H., & Moser, E. (2003). Human motor cortex activity during mental rotation. *NeuroImage, 20*, 225–232.