Acceptance and Effects of Role Models in the Spatial Domain

A Storytelling Experiment with Fourth Graders

Sarah Neuburger[*1], Vera Ruthsatz[1], Petra Jansen[2], Martin Heil[3], Claudia Quaiser-Pohl[1]

[1] University of Koblenz-Landau, Institute of Psychology, Universitaetsstrasse 1, 56070 Koblenz, Germany

[2]University of Regensburg, Institute of Sports Science, Universitaetsstrasse 31, 93053 Regensburg, Germany

[3]University of Duesseldorf, Institute of Experimental Psychology, Universitaetsstrasse 1, 40225 Duesseldorf, Germany

[*1]neuburger@uni-koblenz.de

***Abstract*- Previous research has demonstrated that both real-world and fictional role models can influence gender stereotypes, performance, and self-concept in adults and children. The current study investigated such modeling processes in the domain of spatial skills by examining how fourth-grade boys and girls respond to a spatially skilled fictional role model, who was either male or female. The study had an experimental pre-post design, in which 263 children were examined in three intervention groups (female role model, male role model, no role model). Results showed that model acceptance and model gender ratings were more strongly influenced by the model's sex than by her or his male-stereotyped spatial skills. Model acceptance differed between boys and girls only for the female model, who was accepted to a greater degree by girls, while the male model was accepted to the same degree by boys and girls. Furthermore, there was a gender-specific effect of the stories on girls' general self-esteem, which slightly increased in the female-model group and slightly decreased in the male-model group and the control group. However, the hypothesized gender-specific effects of the story interventions were not found with regard to gender stereotypes, spatial performance, and spatial self-concept; instead, there was an intervention-independent decrease in girls' gender stereotypes and spatial self-concept. These results are discussed with regard to the influence of spatial experience on the gender effect in spatial performance, the effectiveness of fictional versus real-world models, and the problem of upward comparison with highly skilled role models.**

Keywords- Role Models; Spatial Cognition; Narratives; Gender Stereotypes; Elementary-School Children

I. INTRODUCTION

Spatial skills are important determinants for educational and professional success in technical, mathematical, and scientific domains, in which a gender gap still exists [see e.g. 1]. In contrast to the mathematical domain, where performance differences between males and females seem to disappear [2, 3], the male advantage in spatial performance still prevails [4, 5, 6]. The influence of gender stereotypes on this performance difference is increasingly recognized and investigated in adults [e.g. 7]. However, the existence, the impact, and the malleability of *children's* spatial stereotypes have rarely been examined. The present study aimed at contributing to a better understanding of this topic by investigating male and female fourth graders' gender stereotypes, performance, and self-concept in the spatial domain, and by assessing their responses to a (presumed) stereotype-inconsistent female and a (presumed) stereotype-consistent male role model.

A. *The Gender Effect in Spatial Performance*

Spatial performance is an umbrella term comprising various types of non-verbal stimuli processing, and males have been found to outperform females in some of these spatial subdomains [e.g. 8, 9, 10], with the "Mental Rotation Test" [11, 12] producing the largest and most reliable gender effect. With about one standard deviation, the gender effect in mental rotation is considered to be the largest cognitive gender effect [13]. Mental rotation refers to the ability to rotate two- or three-dimensional objects in mind [14]. Several studies suggest the male performance advantage to emerge before adolescence [e.g. 15]. There are equivocal findings with regard to the age in which the gender effect first occurs [see e.g. 16 vs. 17], and habituation studies suggest that there is already a male advantage in infants [18, 19]. In psychometric paper-pencil tests, recent studies indicate that a gender difference in favour of males can be reliably found from ten years onwards [20, 21]; however, Levine and colleagues [22] report a significant male advantage in mental rotation already in second graders who have a middle or high socio-economic background. With regard to the underlying causes of the gender effect, current models adopt an integrative view considering both biological and sociocultural factors. Such psychobiosocial models [23, 24] seem to be the most appropriate approach, because there is some evidence in favor of genetic and hormonal influences on the gender difference in mental rotation [e.g. 25, 26], but pure biological explanations are not unequivocally supported by the data [e.g. 27, 28, 29], and the influence of social-psychological and experiential variables on spatial performance has been demonstrated in several ways: Task instructions inducing effort attributions or emphasizing the role of stereotypes in explaining the gender difference have been found to promote mental-rotation performance [30, 31]. Similarly, self-confidence affects mental-rotation performance and mediates the gender difference [32]. Furthermore, substantial effects of spatial trainings on mental rotation have been reported [33]. In recent studies, training effects of playing computer games [e.g. 34, 35], practicing manual rotation [36], and directly enhancing the "spatial sense" [37] were shown. Gender-related differences in spatial experience [e.g. 38, 39] are probably enlarged by gender stereotypes, which can be defined as “behaviors and attitudes expected of or held as a standard” [40, p. 36] for males and females, or as “culturally shared assumptions and expectations about sex differences in abilities, personality traits, activities, and roles” [41, p. 1493]. According to Nash [42], children in the second grade begin to perceive spatial skills as masculine. Besides influencing the degree to which male and females gather experience in a performance domain, gender stereotypes are also more directly related to performance by inducing “stereotype threat” effects [43], which are induced by evaluative, gender-biased task framing, identity priming or blatant stereotype activation. Stereotype threat triggers feelings of being “at risk of confirming, as self-characteristic, a negative stereotype about one’s group” [44, p.797], which impairs cognitive performance. Females’ performance is influenced by stereotype threat in both the mathematical [e.g. 45] and the spatial domain [e.g. 7, 46, 47, 48]. Such effects have also been demonstrated in children [e.g. 49, 50, 51], but, dependent on age, stereotype activation method and performance measures, children’s susceptibility to stereotype threat could not always be confirmed [52, 53]. In face of these in parts contradictory findings, the impact and malleability of children’s gender stereotypes in the spatial domain should be more closely examined.

B. The Influence of Role Models on Gender Stereotypes, Performance, and Self-Concept

According to the social-cognitive theory of gender development [54], gender stereotypes are established by three mechanisms: modeling, enactive experience followed by evaluative reactions, and direct tuition. Modeling processes can occur in face-to-face contacts with significant social agents in the individual’s real-world environment or by presentation of gender roles in the media. Role models in the positive sense are “individuals who provide an example of the kind of success that one may achieve, and often also provide a template of the behaviors that are needed to achieve such success” [55, p. 36]. The impact of the exposition to a counter-stereotypic individual on subsequent attitude changes depends on perceived group typicalness [56]. If the divergence from the respective group schema is too large, exposition more likely results in subtyping than in schema change [57]. Therefore, only female role models embodying both counter-stereotypic and stereotypic features, thus being perceived as sufficiently typical for females’ own gender, are likely to change gender-related attitudes. With regard to developmental differences in the processing of gender-stereotype consistent and inconsistent information, two cognitive trends have been found. On the one hand, gender schemas become more elaborated with age, and “their tendency to influence and bias information organization, memory, and behavior may increase” [58, p. 988]. On the other hand, however, gender schemas become more consciously controlled and flexible during middle childhood [58, 59, 60], and an increase in the ability to process and recall gender-inconsistent material during elementary-school years has been found [61, 62]. In face of these opposite trends, it cannot clearly be predicted if counter-stereotype interventions are more effective with younger or older elementary-school children.

As indicated above, Bussey and Bandura [54] consider media representations of gender roles as influential sources of gender schemata, stating that “children are continually exposed to models of gender-linked behaviors in readers, storybooks, video games, and […] television” [54, p. 701]. Therefore, the present study investigated modeling processes in response to the presentation of male and (stereotype-inconsistent) female story protagonists, whose features can be deliberately designed and controlled. The impact of gender-role narratives on children’s gender stereotypes has been reported e.g. by Ashby and Wittmayer [63], who demonstrated an increase in fourth graders’ appropriateness ratings of traditionally male jobs and characteristics for females after the exposition to stories depicting women in nontraditional roles. Similarly, Scott and Feldman-Summers [64] found that the exposure to female story protagonists demonstrating nontraditional, socially desirable roles (e. g. explorer, achiever, decision maker, leader, intellectual problem solver) influenced children’s sex role perceptions: Male and female third and fourth graders who had read more stories with a counter-stereotypic female protagonist than stories with a stereotypic male protagonist considered it more likely that girls could do what the protagonist did. Recent studies on fictional role models have focused more strongly on role model presentation on television than on written texts. Female gender role expectations have been found to be influenced by films depicting aggressive, attractive female protagonists [40]. This effect was mediated by the perception that the protagonist was a good role model for women, which is in line with general stereotype research [56].

With regard to the impact of role models on performance, several studies have shown that role models who are members of the individual’s stereotyped group can prevent stereotype threat: Females’ math performance has been found to be protected against stereotype threat by a female experimenter with high perceived math competence, even without the physical presence of this experimenter [65]. Similarly, 11-13 year-old schoolgirls were found to perform better in a math test when positive same-gender role models were accessible [66]. There are several mechanisms by which role models can positively influence performance, e.g. reduction of extra-task thinking [67], enhancement of self-evaluations and motivation [e. g. 65, 68], inspiration and guidance of academic aspirations [69]. Recent studies suggest that role models are more likely to promote girls’ performance in stereotyped domains when the role model’s success is attributed to stable internal causes, i.e. ability [67, 70]. Dependent on their behaviors and features, role models can have either positive or debilitating effects on performance, as the results of David, Spencer, Quinn, and Gerhardstein [71] indicate: after viewing gender-stereotypic television commercials, women underperformed in a math test, avoided math tasks, and showed less interest in educational/vocational options related to quantitative domains. Although most of the studies in this area have been conducted with adults or adolescents, even young children’ s performance is supposedly affected by the sex of a story protagonist: McArthur and Eisen [72] showed that task persistence of male and female preschoolers was affected by the content of a previously heard story: Boys persisted significantly longer in a difficult task when the story depicted a boys' achievement behavior compared to persistence after listening to the same story with a female protagonist, and girls tended to persist longer when the story depicted a girls' achievement behavior.

In addition and related to gender stereotypes and performance, children's self-concept can be influenced by fictional role models, as demonstrated by Ochman [73]: female and male third graders were exposed to twelve stories over four weeks. The stories, in which either a boy or a girl goes through magic adventures, were presented by videotapes of an actor who read the story to the participants. In the nongender-role stereotyped stories with the female protagonist, the girl displayed many traits, behaviors, and activities that are not part of the female stereotype, e.g. the girl rescues a kidnapped boy. The non-stereotypic stories with the female protagonist increased girls' self-esteem and reduced boys' self-esteem, while the stories with the male protagonist resulted in the reverse pattern. Why does the sex of the protagonist influence self-concept changes after story exposition? Lockwood [55] suggests two reasons: first, the protagonist's sex influences the attention that is given to him or her, with gender-mismatched role models being viewed as irrelevant for the self and therefore might be ignored; second, gender-mismatched role models might increase the salience of existing difficulties of one's own gender and thus they might have detrimental effects on self-concept. The latter assumption implies that the sex of a role model is more influential when the participant's own sex is struggling with difficulties or is disadvantaged in a domain, which was empirically confirmed [55]: females were found to be more inspired by outstanding females than by male role models, while males equally benefitted from female and male models. Furthermore, females tended to identify more with the female model. Similar results were reported by Connor and Serbin [74], who examined the responses of male and female fourth, sixth, and eighth graders to stories with either a male or a female protagonist, found that both boys and girls more strongly liked to be the same-sex protagonist, to do the things he/she did, and to be friends with him/her. Thus, for adults as well as for children, the role model's sex seems to influence model acceptance and identification tendencies and therefore affects self-concept changes due to model exposition. Another reason for expecting the model's sex to influence the outcome of role model interventions is based on social comparison theory: Similar others are more likely used as sources of self-relevant information than dissimilar others [75]. Concerning the impact of such comparisons with positive, high-achieving same-sex role models on self-concept, both self-enhancing and self-devaluing effects can occur, depending on whether assimilation (based on gender identity) or contrast effects (emphasizing the differences between self and model) prevail.

In contrast to the mathematic domain, the effects of spatially skilled role models on girls' and women's performance and self-concept have not yet been examined. In face of the predictive value of spatial skills on the success in STEM (science, technology, engineering, and mathematics) [1] and the large size of the gender effect in some aspects of spatial performance, the lack of research in this domain is astonishing. Therefore, the current study assessed the effects of a fictional role model on girls' and boys' gender stereotypes, spatial performance, spatial self-concept and self-esteem. Furthermore, the study explored boys' and girls' responses to male and female spatially skilled models with regard to model acceptance and perceived gender typicality. The study examined these variables in fourth graders, because sex differences in psychometric spatial tests probably emerge around the age of ten [20, 21] and basic steps in the development of children's self-concept of ability occur during elementary-school years [76, 77]. Moreover, as the accessibility of gender stereotype domains increases during early and middle childhood [78], older elementary-school children are more likely influenced by stereotypes in the spatial domain than younger age groups.

C. Design and Hypotheses

The study had a quasi-experimental pre-post design, in which subjects' gender stereotypes, spatial performance, spatial self-concept, and self-esteem were assessed before and after the storytelling intervention. The design of the study included three groups: (1) the female role model group, in which a spatially skilled girl was presented in four stories, (2) the male role model group, in which a spatially skilled boy was presented in four stories, and (3) the control group, in which no role model stories but four gender-neutral texts were presented. In the female and the male role model group, model acceptance and perceived gender typicality were assessed during the intervention after each of the four stories.

Hypothesis 1: The female model was expected to be accepted to a higher degree by girls, while the male model was expected be accepted to a higher degree by boys.

Hypothesis 2: The role model's sex was expected to influence children's gender ratings of the model: The female role model should be rated as more typical for girls than for boys, and the male role model should be rated as more typical for boys than for girls.

Concerning the effects of the intervention, the following results were expected:

Hypothesis 3: The female role model should reduce children's gender stereotypes; therefore, both boys' and girls' posttest stereotype scores (controlled for pretest scores) should be lower in the female group than in the male role model group and in the no role model group.

Hypothesis 4: The female role model should promote girls' spatial performance; therefore, the gender effect in the posttest spatial performance scores (controlled for pretest scores) should be lower in the female role model group than in the male role model group and in the no role model group.

Hypothesis 5: The female role model should promote girls' spatial self-concept; therefore, the gender effect in the posttest spatial self-concept scores (controlled for the pretest scores) should be lower in the female group than in the male role model group and in the no role model group.

Hypothesis 6: Same-sex role models should promote children's self-esteem; therefore, there should be a gender effect in favour of girls' posttest self-esteem scores (controlled for pretest scores) in the female role model group and a gender effect in favour of boys in the male role model group.

II. METHOD

A. Participants

Participants were recruited from 13 public German schools. Local headmasters and teachers were invited via E-Mail or telephone call to participate in the study. 329 fourth graders, whose parents gave their informed consent, participated in the study. As 66 of the participating children missed, mainly because of illness, at least one of the six experimental sessions (pretest, four intervention sessions, posttest), only 263 children (*mean age* = 10.17 years, *SD* = 0.53, *range*: 8.92-12.17) were included in data analysis. The sample included children from families with low (17.1%), middle (24.3%) and high (41.1%) socio-economic status (SES), 17.5% of the subjects' parents did not fill in the SES questionnaire. No reward was given to the participants, because supervisory school authority did not permit rewarding in order to ensure voluntary participation. Each class was assigned to one of three experimental conditions: the female role model group, in which stories with a female protagonist were read (tested: 61 boys/ 51 girls, analyzed: 47 boys/ 40 girls), the male role model group, in which stories with a male protagonist were read (tested: 62 boys/ 59 girls, analyzed: 45 boys/ 52 girls), and the control group, in which non-fictional texts about animals and environment protection were read (tested: 46 boys/ 50 girls, analyzed: 38 boys/ 41 girls).

B. Materials

1) Intervention Material:

In all three groups, intervention material consisted of four texts, which were read aloud and handed out to the children, and four worksheets referring to the texts. For the female and the male role model group, four age-appropriate stories were written (length: 1844 to 1956 words), in which either a girl or a boy displayed good spatial skills. All four stories were constructed around one protagonist, a fourth-grade child named Lena (female role model) or Lukas (male role model). The protagonist displayed the following spatial skills (the activities and performances used to illustrate the respective skill are listed in parentheses): spatial orientation (knowledge of the geographic directions and application of this knowledge in spatial orientation, finding one's way in the city, succeeding in a treasure quest, understanding maps, helping a new class member to orient in and around school), mental rotation (rotating photographs in mind fast and correctly, inventing a cryptograph based on rotating letters) and spatial visualization (understanding and playing with an architect's drafts and computer models, performing well in geometry tasks in school, constructing LEGO® buildings). The stories were nearly identical in the female and the male role model group , except for (1) few details that were adapted to the gender of the protagonist in order to create realistic, "round" characters (e.g. Lena draws a plan of her dollhouse, whereas Lukas draws a plan of his Playmobil town), and (2) the gender of other characters in the stories, which aimed at presenting further positive role models for girls or boys, e.g. a female vs. male architect, female vs. male intelligent classmates, and female vs. male ancient heroes, and at balancing the number of male and female characters in the stories. In the control group, the texts (length: 681 to 1035 words) described four animals (kakapo, orang-utan, wolf, jellyfish), and included information about environment and wildlife protection. In all three groups, worksheets contained six multiple-choice questions about the content of the respective text, in which the right answer had to be chosen out of four alternatives.

2) Model Acceptance:

In the female and the male role model group, the text-understanding items were followed by three questions assessing model acceptance, which were: How much did you like the story? (not at all – a little – much – very much), Would you like to be like Lena/ Lukas? (no – rather no – rather yes – yes), Would you like to be Lena's/ Lukas' friend? (no – rather no – rather yes – yes). A total model-acceptance score was computed by summarizing the answers given to the three questions in the four intervention sessions. Cronbach's α of the resulting twelve-item scale was .92.

3) Model Gender Ratings:

Children's gender typicality ratings were assessed in the final question on the intervention worksheets, which was: Do you think that in this story, Lena/ Lukas is like most girls or like most boys? (exactly like most girls – rather like most girls – neither nor – rather like most boys – exactly like most boys). Answers on the gender typicality question were coded from -2 (exactly like most girls) to +2 (exactly like most boys). Thus, scores above zero reflect a tendency towards the male pole of the scale, and scores below zero reflect a tendency towards the female pole. An overall gender typicality score was computed by averaging the answers to the gender-typicality question in the four intervention sessions. Cronbach's α of this four-item scale was .89.

4) Gender Stereotypes:

A 15-item questionnaire, which assessed stereotypes about play preferences and activities, personality traits, and spatial activities and skills, was constructed. The five spatial activities and skills were supposed to be male stereotyped (performing well on mental-rotation tasks, willingly building something according to construction plans, well understanding maps and road schemes, playing with LEGO®, and performing well on line-reflection tasks). Of the non-spatial items, five items were female stereotyped (playing with dolls, playing with necklaces and bracelets, dancing and going to ballet lessons, crying frequently, wanting to look beautiful and to please other people), and five items were male stereotyped (playing with trucks, playing Cowboys and Indians, playing soccer, being brave, always want to tell what to do). These non-spatial items were based on the findings of Trautner, Helbing, Sahm, and Lohaus [79] and Miller et al. [78]. Item formulation followed Trautner et al. [79]: "There are children who [activity or trait]. Who [activity or trait]?" Answers were given on a picture-supported, five-point scale ("only girls", "more girls than boys", "as many girls as boys", "more boys than girls", "only boys") and coded with numbers from -2 (only girls) to +2 (only boys), so that scores above zero reflected a male stereotype and scores below zero reflected a female stereotype. The total gender-stereotype score was computed by summarizing the absolute values of all answers. Cronbach's α of the complete scale was .77. Item statistics suggested the elimination of one supposedly male item ("There are children who always want to rule. Who does always want to rule?"), because this elimination increased Cronbach's α to .81. Thus, only 14 of the original 15 items were included in further analyses.

5) Spatial Performance:

As measure of spatial test performance, we used the "Mental Rotations Test" (MRT) of Peters et al. [12] in two shortened versions [version A and version B, see 80]. Half of the participants solved Version A in the pretest and Version B in the posttest, the other half of the participants solved Version B in the pretest and Version A in the posttest. In order to ensure task understanding, a training task was administered before the MRT in the pretest. This training task was constructed in the same way as the MRT, except that letters instead of three-dimensional cube figures had to be rotated. Both rotation tasks consisted of twelve test items with one target on the left side and four comparison stimuli on the right. Two of the four comparisons were rotated versions of the target and had to be crossed out by the participants. For the letters rotation task, test duration was two minutes, for the more difficult cube-figures rotation task, the performance measure, test duration was three minutes. Figure 1 shows one sample item of each rotation task. According to Peters et al. [12], spatial performance was defined as number of correctly solved items, i.e. one point was given for each item in which both correct comparisons and none of the mirror images had been crossed out. Thus, a maximum score of 12 could be achieved.

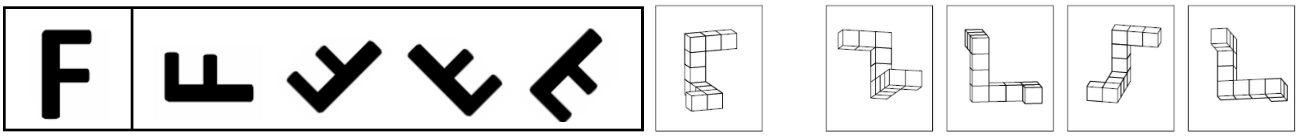

Fig. 1 Example items from the mental-rotation tasks (left: training task, right: pre-/post-performance measure)

6) Spatial Self-Concept:

The questionnaire consisted of nine items referring to activities that require spatial skills: solving mental rotation tasks, finding one's way in an unknown environment, understanding maps and road schemes, solving line-reflection tasks, continuing cube-figure drawings, building something according to construction plans, playing with LEGO®, playing Gameboy or Playstation, and playing action or adventure computer games. For each of the spatial activities, one of the following three types of questions was used: (1) How good are you at [*spatial activity*]? (2) How much fun is it for you to [*spatial activity*]? (3) How easy do you find [*spatial activity*]? Answers were given at a three-point scale and coded with 2, 1, and 0 (2 = better/more fun/easier than other children; 1 = as good as/ as much fun as/ as easy as other children; 0 = worse than/ less fun than/ more difficult than other children). A total spatial self-concept score was computed by summarizing the answers to the nine questions. Cronbach's α of the scale was only .65, which might reflect the heterogeneity of the construct "spatial abilities".

7) Self-Esteem:

The self-esteem measure was a shortened version of the German "Statements List of self-esteem for children and adolescents" [ALS; 81]. The questionnaire comprised nine positive self statements (e.g. I am very satisfied with myself.", "I am a valuable person.", "Others take me seriously.") and nine negative self statements (e.g. "I am afraid of making mistakes.", "I am a loser.", "Others laugh at me."). For each statement, participants were asked to indicate on a four-point answer scale how often it applies to themselves (0 = never; 1 = sometimes; 2 = most of the time; 3 = always). The self-esteem score was computed by reversing the values of the answers to the negative self statements and adding them to the values of the answers to the positive self statements. Cronbach's α of the scale was .85

8) Intelligence:

We controlled for general cognitive abilities by administering the subtest "Reasoning" of the "Cognitive Ability Test" [KFT 1-3; 82]. It consists of 15 items, which are made up of five pictures, respectively. Four of the five pictures belong to the same category (e.g. vegetable). The picture that does not fit into this category has to be detected and marked by the child. As the KFT is a power-test, there is no time constraint. The test was chosen because it measures general non-verbal cognitive abilities, and because performance in the test does not too strongly depend on visual-spatial abilities. Reasoning performance was defined as the number of correctly solved items.

9) Socio-Economic Status (SES):

Participants' SES was assessed with a questionnaire filled in by their parents, who were asked to indicate their school leaving certificate and professional degree. Following Jöckel et al. [83], rank-ordered values were assigned to the different combinations of school and professional graduation levels.

C. Procedure

Pre- and posttest as well as the intervention took place in classrooms of the participating schools. All parts of the experiment were administered in groups of 10-25 children. There were six experimental sessions (Session 1: pretest, Sessions 2-5: intervention, Session 6: posttest), which were administered within four weeks (Week 1: pretest, Weeks 2-3: intervention sessions twice a week, Week 4: posttest). In each class, one female experimenter administered the pre- and posttest sessions, and another female experimenter administered the intervention sessions. Before and during the study, neither teachers nor parents nor participants were told about the expected effect of the stories on the tests and questionnaires; instead, they were told that the study included two different parts (measurement of various cognitive competences and attitudes on the one hand, and training and evaluation of text understanding on the other hand). In the pre- and posttest, instructions did not highlight the diagnostic, evaluative purpose of the spatial test, instead, it was intended (a) to make children feel comfortable in the test situation so that they would enjoy participating in the study, and (b) to clearly communicate that performance would not affect the teachers' recommendation for secondary school.

1) Pretest:

In Session 1, which required about 90 minutes including a break of about 10-20 minutes, the pretest measures were administered and intelligence was assessed. First, the experimenter gave the following short introduction: "We are currently working on a research project, in which we administer various tasks with children in several schools. These tasks are interesting and you won't get grades for your performance. Our aim is to examine the thinking and learning of children. As your teacher has already told you, this research project comprises several lessons over a period of one month." After this introduction, the Reasoning Test was explained by the experimenter. When all participants had finished the Reasoning Test, the self-esteem scale was administered. The experimenter read aloud each of the questions to compensate for possible reading problems. Afterwards, the concept of "mentally rotating objects" was introduced by rotating a real, familiar object (a pair of scissors), and the mental-rotation training task was administered. Then, children's spatial performance in the cube-figures test was assessed. Either between the two mental-rotation tasks or after the two tasks, there was a short break. Afterwards, spatial self-concept was assessed. As with the self-esteem scale, all items were read aloud by the experimenter. Finally, gender stereotypes were assessed. The scale was introduced in the following way: "This questionnaire is about girls and boys. Again, there are no right or wrong answers. We are interested in your opinion." Again, all items were read aloud by the experimenter. Afterwards, children received a "scientist passport", in which all experimental sessions were documented.

2) Intervention:

In each of the four intervention sessions, the text was first read aloud by the experimenter. Then, the text and the worksheet were handed out to the children. The experimenter said: "This worksheet contains different types of questions: the first six (control group: eight) questions assess how well you understood the text. For answering the questions you are allowed to check up in the text. The right answer has to be chosen out of four alternatives. Question seven-ten (control group: nine-ten) are about your opinion on the text. It is important to work on your own, do not talk to your neighbor." Children were given as much time as needed for solving the worksheet. Using the text for answering the question was explicitly allowed because the worksheets were not intended to assess recall performance, but to induce deeper elaboration of the text. After finishing the worksheet, they delivered it to the experimenter and were instructed to paint a picture referring to the text. When all children had solved the tasks, answers to the text understanding questions were discussed in the group.

3) Posttest:

The posttest session followed the same procedure as the pretest session, but without the mental-rotation training task (letters task) and the Reasoning Test. Finally, children were thanked for participating in the study. Debriefing was done by delivering written project information to the schools.

III. RESULTS

A. Reasoning Ability

A 2 (sex) x 3 (group)-ANOVA with standardized reasoning scores, i.e. age corrected T-values, as dependent variable revealed a significant sex difference with regard to general cognitive abilities: girls' scores (M = 54.19; SD = 8.10) were significantly higher than boys' scores (M = 52.09; SD = 8.16), F (1,248) = 4.82; p = .029; η^2 = .019. There was neither a main effect of group nor an interaction between sex and group, $p > .10$. Because of the sex difference, reasoning score was included as covariate in all of the subsequent analyses in which boys were compared to girls.

B. Socio-economic Status (SES)

In order to examine possible group differences with regard to SES, a Chi ²-Test was computed, in which the percentages of children with high, middle, and low SES were compared between the female role model group, the male role model group, and the no role model group. Since no significant differences between groups were found, $chi^2(4) = 5.39$; $p = .25$, SES was not included in the further analyses. More detailed analyses of the sample with regard to interactive effects of SES and gender on spatial performance and self-concept are reported by Ruthsatz, Neuburger and Quaiser-Pohl [84].

C. Analysis of Dropouts

Excluded cases were equally distributed over the three groups and two sexes, $chi^2(2) = 1.04$; $p > .10$, they did not differ from the analyzed sample with regard to the SES distribution, $chi^2(2) = 0.18$; $p > .10$, age, $p > .05$, reasoning ability, $p > .10$, or pretest spatial performance, $p > .10$.

D. Intercorrelations of Dependent Measures

The descriptive results of pre- and posttest measures are shown in Table I. Before computing the correlations between the pretest dependent measures, scores were z-standardized on the basis of schools in order to control for effects of cluster sampling. Then, Pearson’s correlation coefficients were computed separately for boys and girls. For boys, no significant correlations were found between the standardized pretest dependent measures, $p > .10$; for girls, there was a significant correlation between self-esteem and spatial performance, $r = .19$; $p < .05$, and a marginally significant correlation between self-esteem and spatial self-concept, $r = .17$; $p = .06$, none of the remaining correlations reached significance, $p > .10$.

TABLE I PRE- AND POSTTEST GENDER STEREOTYPES (GS), SPATIAL PERFORMANCE (SP), SPATIAL SELF-CONCEPT (SSC), AND SELF-ESTEEM (SE)

M (SD)	**Female Role Model**		**Male Role Model**		**No Role Model**	
Boys	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
GS	10.19 (4.10)	9.53 (5.11)	12.13 (3.85)	11.49 (4.54)	10.79 (4.35)	11.12 (7.97)
SP	3.64 (2.89)	4.93 (3.13)	3.70 (2.39)	4.98 (2.73)	4.61 (2.37)	5.39 (3.05)
SSC	12.23 (2.24)	12.16 (2.60)	12.47 (2.53)	12.51 (2.68)	12.26 (2.78)	11.40 (3.14)
SE	39.11 (9.05)	38.46 (8.99)	36.97 (5.57)	36.88 (6.56)	37.64 (6.21)	37.79 (7.52)
Girls	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
GS	8.87 (4.36)	6.79 (3.05)	8.63 (3.13)	6.96 (3.73)	7.90 (3.66)	5.95 (3.49)
SP	2.49 (1.92)	3.03 (1.99)	2.06 (1.53)	3.29 (1.97)	2.46 (1.66)	4.00 (2.28)
SSC	9.54 (2.46)	9.03 (3.06)	9.73 (2.82)	8.94 (2.73)	10.08 (3.28)	9.38 (3.01)
SE	37.56 (7.62)	38.63 (7.03)	36.58 (7.54)	34.20 (8.08)	37.31 (6.59)	36.03 (8.07)

In order to examine if model acceptance and model gender typicality ratings were associated with the pretest scores or with each other, correlations between these two variables and the z-standardized pretest scores (gender stereotypes, spatial performance, spatial self-concept, and self-esteem) were computed separately for boys and girls and the two role model groups. For *boys in the female role model group*, gender typicality ratings of the role model marginally correlated with spatial performance scores, $r = -.26$; $p = .08$, and with gender stereotype scores, $r = -.26$; $p = .08$, i.e. for boys, more feminine/less masculine ratings of the spatially skilled female role model were in tendency associated with better spatial performance and lower overall gender stereotypes. For *boys in the male role model group*, model acceptance significantly correlated with gender stereotype scores, $r = -.39$; $p < .05$, and with gender typicality ratings of the model, $r = .40$; $p = .01$, i.e. for boys, a higher acceptance of the spatially skilled male role model was associated with lower gender stereotype scores and more masculine/less feminine ratings of the male model. For *girls in the female role model group*, gender typicality ratings of the role model significantly correlated with spatial performance scores, $r = .40$; $p = .01$, and with self-esteem, $r = .50$; $p < .01$, i.e. with girls, more masculine/less feminine ratings of the spatially skilled female role model were associated with better spatial performance and higher general self-esteem. For *girls in the male role model group*, no significant correlations were found, $p > .10$. Because of these, for the most part, only non-significant to small correlations between the dependent variables, a univariate approach was chosen for the further analyses.

E. Model Acceptance and Gender Typicality Ratings (H1, H2)

For the female and the male role model group, the influence of participants’ sex and group on model acceptance and on gender typicality ratings was examined by 2x2 between-subjects analyses of covariance (ANCOVA). In line with the hypotheses, there was an interaction between group and sex for model acceptance, $F(1,158) = 4.00$; $p = .047$; $\eta^2 = .03$. Unexpectedly, there were also main effects of group, $F(1,158) = 5.30$; $p = .023$; $\eta^2 = .03$, and sex, $F(1,158) = 9.93$; $p = .002$; $\eta^2 = .06$. No significant effect of the covariate reasoning ability was found, $p > .10$, and an ANOVA without the covariate confirmed the above described effects. Figure 2a shows that girls’ model acceptance was higher in the female than in the male role model group. However, boys’ model acceptance was approximately the same in the two role model groups. The female model was accepted to a greater degree than the male model, and girls more strongly accepted both the female and the male model more than boys did. Bonferroni corrected simple main effects showed that there was a significant difference between boys' and girls' acceptance scores in the female model group, F (1,79) = 16.51; p< .01; η^2 = .17, but not in male model group, F (1,78) = 0.54; p> .10.

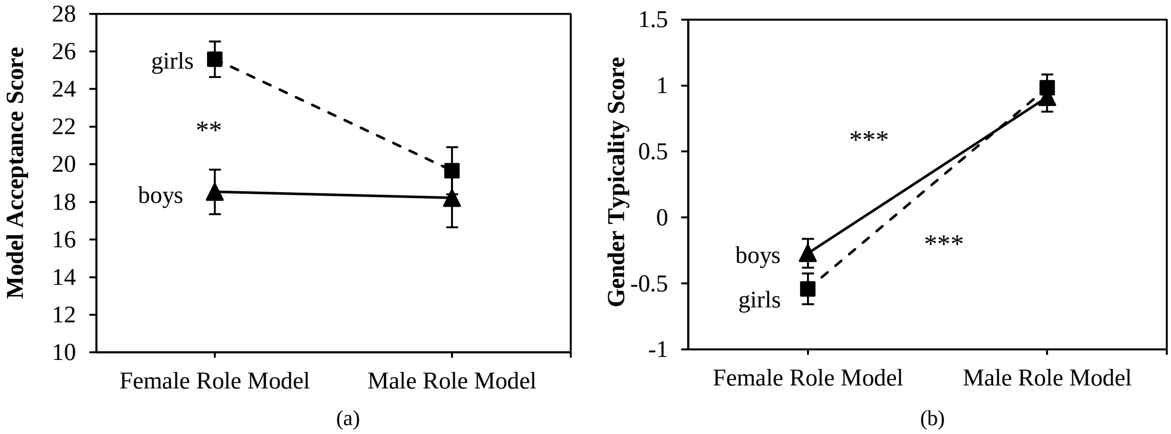

Fig. 2 (a) Means and standard errors of model acceptance; (b) Means and standard errors of gender typicality ratings as a function of gender and group

The 2x2-ANCOVA with gender typicality score as dependent variable revealed, in line with the hypotheses, a main effect of group, F (1,172) = 155.10; p< .001; η^2 = .47. Neither the main effect of sex nor the interaction between sex and group reached significance, all p> .10. No significant effect of the covariate reasoning ability was found, p> .10, and an ANOVA without the covariate confirmed the above described results. Figure 2b shows that gender ratings reflected the role model's sex; One-sample T-Tests showed that the gender typicality ratings of the female role model significantly differed from the neutral value of the scale (=0) towards the *female* pole (= -2), t (86) = 4.34; p< .001, while the gender typicality ratings of the male role model significantly differed from the neutral value of the scale (=0) towards the *male* pole (= 2), t (96) = 15.60; p< .001.

F. Gender Stereotypes (H3)

In order to examine the effect of the intervention on gender stereotypes, a 2x2-ANCOVA with the factors group and sex and the dependent variable posttest stereotype score was conducted; reasoning score and pretest stereotype score were included as covariates in the ANCOVA. There was a main effect of sex, F (1,235) = 16.83; p< .001; η^2 = .07 (see Figure 3), and a significant effect of the covariate pretest stereotype score, F (1,235) = 130.88; p< .001; η^2 = .36, but neither a main effect of group, p > .10, nor an interaction of group and sex, p> .10. Although there was a marginally significant effect of the covariate reasoning ability, F (1,235) = 3.68; p = .056; η^2 = .02, the above described results were confirmed in an ANCOVA without the covariate reasoning ability.

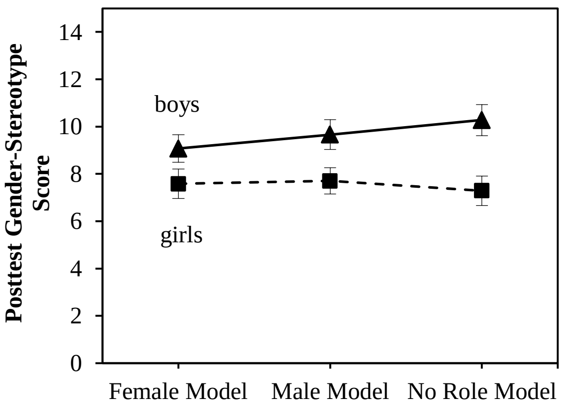

Fig. 3 Girls' and boys' posttest gender-stereotype scores corrected for pretest scores and reasoning scores (estimated marginal means and standard errors)

A 2 (group) x 2 (sex)-ANCOVA on the pretest gender stereotype score revealed a significant main effect of sex, F (1,240) = 26.23; p< .001; η^2 = .10. Boys' stereotype scores (M = 11.08; SD = 4.15) were higher than girls' stereotype scores (M = 8.42; SD = 3.71). None of the other effects reached significance, $p \geq$ .10. There was no significant effect of the covariate reasoning ability, p> .10, and an ANOVA without the covariate confirmed the above described results. Boys' and girls' pretest ratings on the five spatial items are shown in Figure 4. Scores above zero represent a bias towards the male pole of the scale, and scores below zero represent a bias towards the female pole. In order to examine the degree to which the spatial items of the gender-stereotypes scale were actually stereotyped, boys' and girls' pretest ratings on these items were compared to the gender-neutral score zero by one-sample t-tests. Boys' ratings were significantly above zero on all of the spatial items (mental rotation: t (127)

= 3.54; p = .001, building according to construction plans: t (127) = 13.15; p< .001, playing with LEGO®: t (127) = 14.13; p< .001, understanding maps and road schemes: t (127) = 4.89; p < .001, line reflection: t (127) = 3.68; p< .001). Girls' ratings were significantly above zero on only two spatial items, which were building according to construction plans, t (131) = 11.38; p< .001, and playing with LEGO®: t (131) = 12.13; p< .001. Girls' ratings did not significantly diverge from zero for mental rotation and map understanding (p> .10), and they were significantly below zero for line reflection, t (131) = 3.47; p = .001.

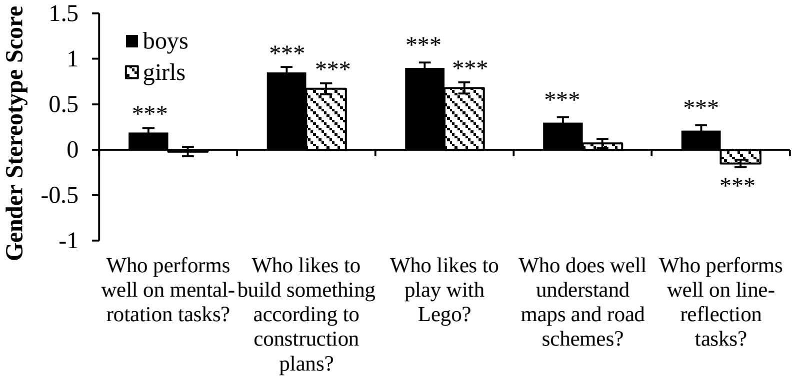

Fig. 4 Means and standard errors of boys' and girls' ratings on the spatial items of the gender-stereotype scale; possible values range from -2 ("only girls") to +2 ("only boys")

G. Spatial Performance (H4)

A 2 (group) x 2 (sex) – ANCOVA on the pretest spatial performance score revealed a main effect of sex, F (1,247) = 43.81; p< .001; η^2 = .15, and a main effect of group, F (2,247) = 3.50; p< .05; η^2 = .03, but no interaction of group and sex, p> .10. Although there was a significant effect of the covariate reasoning ability, F (1,247) = 9.41; p< .01; η^2 = .04, the above described results were confirmed in an ANOVA without the covariate. Boys (M = 4.15; SD = 2.50) outperformed girls (M = 2.52; SD = 1.64), and the female role model group (M = 3.73; SD = 2.29) outperformed the no role model group (M = 3.47; SD = 2.28), which in turn outperformed the male role model group (M = 2.81; SD = 2.12).

In order to examine the effect of the intervention on spatial performance, a 2x2-ANCOVA with the factors group and sex and the dependent variable posttest performance score was conducted; reasoning score and pretest performance score were included as covariates in the ANCOVA. There was a main effect of sex, F (1,246) = 9.24; p< .01; η^2 = .04, a significant effect of the covariate pretest performance, F (1,246) = 76.28; p< .001; η^2 = .24, and an interaction of group and sex, F (2,246) = 3.23; p< .05; η^2 = .03. No main effect of group was found, p> .10. There was no significant effect of the covariate reasoning ability, p> .10, and an ANCOVA without the covariate largely confirmed the above described results, with the exception that the previously significant interaction group x sex was only marginally significant, F (1,255) = 2.89; p = .058; η^2 = .022. Figure 5a shows that the direction of the group x sex interaction was not in line with the hypotheses: There was a significant gender effect in favour of boys in the female role model group, which was neither found in the male role model group nor in the no role model group. Overall, boys' posttest scores (M = 5.37; SD = 2.91) were higher than girls' posttest scores (M = 3.53; SD = 2.04).

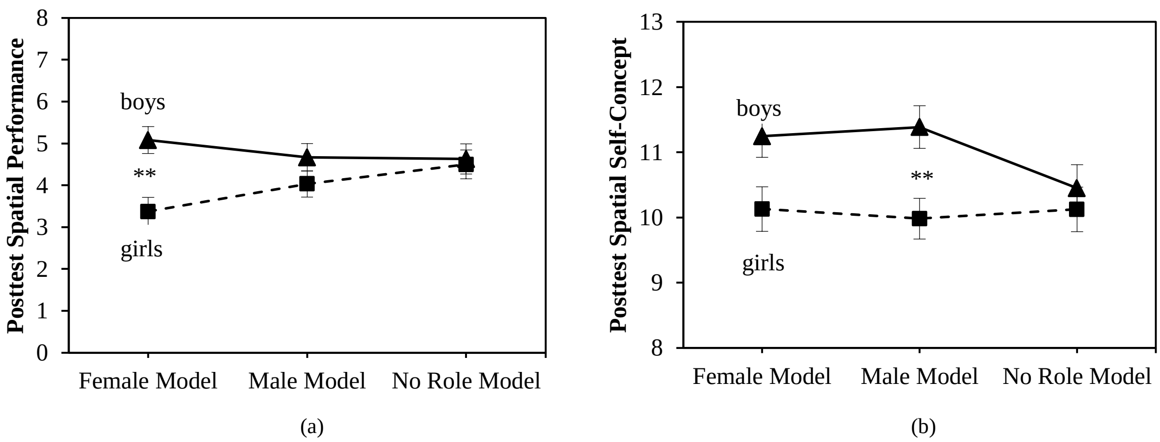

Fig. 5 (a) Girls' and boys' posttest spatial performance scores corrected for pretest scores and reasoning scores (estimated marginal means and standard errors); (b) Girls' and boys' posttest spatial self-concept scores corrected for pretest scores and reasoning scores (estimated marginal means and standard errors)

H. Spatial Self-Concept (H5)

A 2 (group) x 2 (sex) – ANCOVA on the pretest spatial self-concept revealed a main effect of sex, F (1,242) = 54.96; p< .001; η^2 = .19, but neither a main effect of group nor an interaction group x sex, p> .10. Boys' pretest spatial self-concept (M = 12.33; SD = 2.45) was higher than girls' pretest spatial self-concept (M = 9.78; SD = 2.83). There was no significant effect of the covariate reasoning ability, p> .10, and an ANOVA without the covariate confirmed the above described results.

In order to examine the effect of the intervention on spatial self-concept, a 2x2-ANCOVA with the factors group and sex and the dependent variable posttest self-concept score was conducted; reasoning score and pretest self-concept score were included as covariates in the ANCOVA. There was a main effect of sex, F (1,239) = 12.47; p< .001; η^2 = .05, and a significant effect of pretest self-concept, F (1,239) = 312.65; p< .001, but neither the main effect of group nor the interaction of group and sex reached significance, p> .10. There was no significant effect of the covariate reasoning ability, p> .10, and an ANCOVA without the covariate confirmed the above described results. Figure 5b shows that, contrary to the hypotheses, the gender effect in spatial self-concept after the intervention (which, however, reached significance only in the male role model group) was about the same in the two role model groups and lower in the no role model group. Overall, boys' posttest spatial self-concept (M = 12.07; SD = 2.80) was higher than girls' posttest spatial self-concept (M = 9.10; SD = 2.90).

I. Self-Esteem (H6)

A 2 (group) x 2 (sex) – ANCOVA on the pretest self-esteem score revealed no main effects and no interaction, p> .10. There was no significant effect of the covariate reasoning ability, p> .10, and an ANOVA without the covariate confirmed the above described results. In order to examine the effect of the intervention on self-esteem, a 2x2-ANCOVA with the factors group and sex and the dependent variable posttest self-esteem score was conducted; reasoning score and pretest self-esteem score were included as covariates in the ANCOVA. Neither a main effect of group nor a main effect of sex was found, p> .10, but there was a significant effect of the covariate pretest self-esteem, F (1,210) = 352.68; p< .001; η^2= .63, and a marginally significant interaction of group and sex, F (2,210) = 2.84; p = .061; η^2 = .03. No significant effect of the covariate reasoning ability was found, p> .10, and an ANCOVA without the covariate reasoning ability largely confirmed the above described results, with the exception that the previously marginally significant interaction of group and sex reached significance, F (2,218) = 3.51; p = .032; η^2 = .03. Figure 6 shows that the direction of this interaction was in line with the hypotheses: In the male role model group and in the no role model group, boys' posttest self-esteem was higher than that of girls, while in the female role model group, girls' posttest self-esteem was higher.

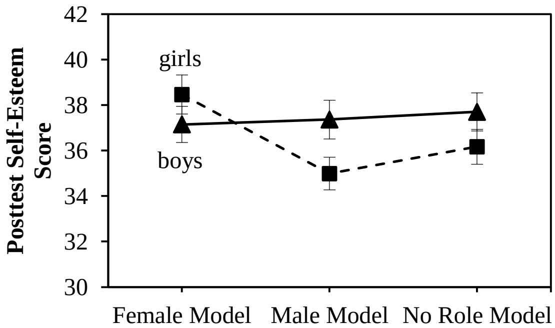

Fig. 6 Girls' and boys' posttest self-esteem corrected for pretest scores and reasoning scores (estimated marginal means and standard errors)

J. Correlations between Change Scores and Model Ratings

As with Section III.D., z-standardized scores (controlled for inter-school differences) were used in the following analyses. Change scores were computed by subtracting pretest scores from the posttest scores. For *boys in the female role model group*, spatial self-concept change positively correlated with gender stereotypes change (r = .43; p< .01), and self-esteem change positively correlated with model acceptance (r = .37; p< .05). For *boys in the male role model group*, self-esteem change positively correlated with spatial self-concept change (r = .48; p< .01), and gender stereotypes change (r = .40; p< .05); furthermore, spatial performance change negatively correlated with model acceptance (r = -.34; p< .05) and model gender ratings, i.e. the model's subjectively perceived masculinity (r = -.31; p< .05). For *girls in the female role model group*, spatial performance change positively correlated with self-esteem change (r = .37; p< .05) and negatively correlated with model gender ratings, i.e. the model's subjectively perceived masculinity (r = -.35; p< .05); furthermore, self-esteem negatively correlated with gender stereotypes change (r = -.37; p< .05). For *girls in the male role model group*, gender stereotype change positively correlated with model gender ratings, i.e. the model's subjectively perceived masculinity (r = .35; p = .01).

IV. DISCUSSION

In this study, a storytelling intervention based on the modeling aspect of social-cognitive gender-role learning theory [54] was examined in the domain of spatial abilities. Hypothesis 1, which referred to children's preference for same-sex models, was partially confirmed: as expected, the female model was more strongly accepted by girls than by boys; however, the male model was equally accepted by boys and girls. Hypothesis 2, which assumed that the sex of the role model would influence children's gender ratings, was confirmed: the female role model was rated as more typical for girls than for boys, and the male role model was rated as more typical for boys than for girls. Hypotheses 3-5 could not be confirmed: neither children's gender stereotypes nor the male advantages in spatial performance and spatial self-concept were reduced by the female role model. With regard to Hypothesis 6, results suggest that the same-sex role model slightly promoted children's self-esteem, although the interaction of gender and group was relatively small.

Concerning model acceptance and model gender ratings, the following aspects of the results should be more closely examined: for girls, results suggest that the sex of the role model is important for their acceptance ratings; a same-sex model is more strongly accepted than a different-sex model. In contrast to the study of Connor and Serbin [74], where *boys'* model acceptance was more sensitive to the sex of the fictional role model, the current results suggest that *girls* more strongly differentiate between male and female models, at least in the case of spatially skilled models. This would be in line with the theoretical assumptions and the findings of Lockwood [55]: in case of existing disadvantages or difficulties of the individual's own social group/sex, the role model's group status/sex becomes more important. For boys, the perception of the role model's gender seems to be more important: while boys' model acceptance did not differ between the male and the female role model, who both displayed the same behaviours, there was a positive correlation between the subjectively perceived masculinity and the acceptance of the male role model. For girls, no association between model acceptance and gender typicality ratings was found; however, interestingly, there was a positive correlation between the subjectively perceived masculinity of the female role model on the one hand and self-esteem and spatial performance on the other hand. This is astonishing, because it might have been expected that girls with a higher self-esteem and better spatial skills would, on the basis of their own skills and behaviours, associate high spatial skills more strongly with femininity than with masculinity and thus would perceive a female role model more feminine/less masculine than girls with lower self-esteem and spatial skills. The opposite direction of this association possibly indicates that girls with high spatial skills are not necessarily free from gender stereotypes, or, rather, they do not ignore the actually existing sex difference in spatial performance, despite their own highly developed spatial skills; instead, because of everyday experiences of the discrepancy between themselves and the majority of the other girls, they might be even more strongly aware of the exceptional status of a girl who shows behaviours and skills as the role model did. For both boys and girls, the sex-congruent bias in the gender ratings of the model was more pronounced for the male role model, which was probably due to the fact that the role model displayed high spatial abilities, which are more strongly associated with masculinity than with femininity. However, the female role model was clearly perceived as more feminine than the male role model, despite her stereotype-inconsistent spatial talent. This is in line with Conkright et al. [62], who demonstrated that elementary-school children's gender assignments to imaginary characters are primarily based on the sex indicated by the pronoun (he vs. she) used in the stories, and not on the activity preferences of the characters. Taken together, model acceptance and gender typicality ratings suggest that the spatially skilled girl in our stories was accepted as a positive female role model by the participating girls, which is an important prerequisite for subsequent attitude changes [56, 91]. The fact that girls who more strongly perceived the female role model as typical for their own gender showed a stronger increase in spatial performance supports the assumption that gender ratings indeed influence the effect of role models.

The expected reduction of gender stereotypes after the female role model intervention was not confirmed. However, besides the effects of the storytelling intervention, some interesting aspects of the present findings concerning gender stereotypes deserve attention. First, the existence of a spatial gender stereotype was confirmed for boys and partially confirmed for girls. Both sexes agreed that building and construction activities are more typical for boys than for girls. As girls did not show a male stereotype for other spatial activities and skills (mental rotation, spatial orientation, line reflection), further stereotype research in this area should take a differentiated perspective on the specific subcomponents of the umbrella construct "spatial ability". In contrast to findings with adult females [24,85], no gender stereotype was found for mental rotation with girls; but as the cube figures of the MRT, the most commonly used mental-rotation task, resemble building blocks, this test nevertheless might activate a male stereotype. Recent data show that the MRT cube figures are indeed perceived as more typical for males than for females by fifth graders [86], although the cube figures were not as strongly stereotyped as other stimuli, e. g. tractor or revolver, in that study. Thus, the role of the stimulus material used in spatial tests should be more closely examined. Rather than using such strongly stereotyped tasks like block building in assessing spatial skills [87], researchers should design more gender-fair measurements, as it has been done in other cognitive domains. Furthermore, since the present findings are based on a descriptive-stereotypes questionnaire assessing what children think of how boys and girls actually are, modified measurements should also assess the degree to which prescriptive spatial gender stereotypes of how boys and girls should be exist (e.g. "girls should not play with LEGO"). With regard to the overall gender-stereotype ratings, girls' stereotype scores were lower than boys' scores. The lower stereotype scores of the female participants are contrary to the higher proportion of girls' stereotypes reported by Miller et al. [78]. However, this might have been an artifact of the stereotype questionnaire used in the current study, which included more positively connoted male than female items and more negatively connoted female than male items. Another interesting finding of the present study is the group-independent decrease in girls' stereotype scores from pre- to posttest. This might reflect the impact of the female experimenter, who served as a real-world, face-to-face model of spatial abilities by administering scientific spatial tests. Notably however, this modeling process influenced girls' gender stereotypes more strongly than boys' stereotypes.

Concerning spatial performance, the female role model did not reduce the male advantage. Instead, there was an overall increase in spatial performance, which was probably, for the most part, an effect of repeated testing [33], and in the female role model group, boys' spatial performance increased even more than girls' performance. One reason for the absence of an intervention-specific increase in girls' spatial performance might have been the non-evaluative instruction given in the pre- and posttest, which emphasized the scientific, non-diagnostic interest of the experimenter. Such instructions have been found to counteract stereotype threat [88]; therefore, the present study might have failed to demonstrate the benefits of same-sex role models on females' performance found in other studies. In face of this methodic limitation, the present results should not be considered as questioning the positive impact of female role models in the spatial domain.

Similar to the results on spatial performance, the female role model did not promote girls' spatial self-concept; both pre- and posttest results showed a clear difference in favor of males between boys' and girls' spatial self-concept. In line with the findings regarding spatial gender stereotypes, the gender effect in spatial self-concept suggests leisure-time activities contributing to spatial-skill development to be more prevalent in boys' than in girls' everyday experiences. The finding that the female role model did not reduce the male advantage in children's spatial self-concept might be due to the fact that same-sex role models can induce two different social comparison processes, which affect the self-concept in opposite directions: as outlined in Section I.B., assimilation effects lead to self-enhancing, while contrast effects lead to self-devaluing. The latter might explain the somewhat paradox finding of negative effects of same-sex role models on self-concept, as reported by Potter [89], who investigated the effect of reading storybooks about a superior and an inferior problem solver on third graders' self-concept. In one condition, the superior problem solver was a girl, and the inferior problem solver was a boy; in the other condition, the sexes of the main characters were reversed. Unexpectedly, the self-concept of the female subjects increased when the superior problem solver was a boy and the inferior problem solver was a girl, and stayed the same in the reversed condition; the self-concept of the male subjects also increased when the inferior, not the superior, problem solver was a boy, and dropped in the reversed condition. Potter [89] interprets her results on the basis of self-devaluing upward-comparisons induced by same-sex model and concludes that "the dynamics involved [in the influence of the model's sex] are far more complex […] and the direction of change is not easily predicted" [89, p. 13]. In the present study, both self-devaluing upward-comparisons and self-enhancing assimilation effects might have occurred, with the latter resulting in the positive impact of the female role model on girls' self-esteem. Interestingly, there were significant positive correlations between girls' pretest self-esteem and pretest spatial performance, as well as between the change in girls' self-esteem and the change in their spatial performance – this association should be further investigated, especially with regard to its underlying causal mechanisms.

Apparently, role-model interventions have to be carefully designed in order to avoid threatening girls' self-concept. One important moderator of role-model effects seems to be the attainability of the model's high performance, which is higher when a model is somewhat older than the subjects [69, 70]. Therefore, in the present study, the age similarity between the model and the participants possibly reduced the attainability and thus the inspiring, self-elevating potential of the stories. Upward comparison to the highly skilled fictional model cannot explain the unexpected decrease in both boys' and girls' spatial self-concept in the control group, in which texts about animals and nature protection were read. Perhaps these texts unintentionally challenged children's spatial skills because they contained some geographical information. Another methodic aspect that should be considered with regard to the control condition is the length of the texts, which were somewhat shorter than the role model stories. Thus, the present results should be verified in designs where the length of control texts is adjusted to that of the intervention texts; furthermore, the inclusion of "no intervention" control groups might be reasonable. Concerning the experimental procedure, test instructions (diagnostic vs. non-diagnostic) as well as the effect of completing gender-stereotypes questionnaires on performance and self-concept should be considered: Hausmann et al. [24], e. g., implicitly activated stereotype threat by assessing gender-stereotypes before administering cognitive tests. In order to avoid such effects in the present study, gender stereotypes were assessed after the spatial test had been administered; nevertheless, the questionnaire might have resulted in a long-term activation of gender stereotypes beyond the pretest session and thus might have influenced children's performance and answers in the questionnaires in the posttest.

Narratives easily attract children's – and adults' – attention. Probably since the existence of mankind, stories have been used as educational medium by conveying knowledge, moral standards, and social conventions. Future studies should examine how content (fiction vs. facts) and presentation mode (e.g. written/read stories vs. television) influence the impact of story interventions in the spatial domain. In other contexts, fictional stories presented on television have been found to influence children's real-world schemata, although some schema aspects were more influenced by the factual television program, at least in the short term and in case of socially unrealistic or obviously non-factual features of the fictional program [90]. Furthermore, the influence of female experimenters and other face-to-face presented, real-life role models on girls' spatial performance provides an important aspect for further research.The present study, which was intended to examine how storytelling might promote females with regard to spatial performance and spatial self-concept, provides first insights into children's responses to spatially skilled female and male fictional role models. In contrast to studies in other domains [e.g. 66, 72], the positive impact of same-sex models in the spatial domain could not be confirmed. However, results indicate that female fourth graders do accept spatially skilled fictional models and, despite their stereotype-inconsistent spatial skills, do not perceive them as "unfeminine". Further research should explore how these modeling prerequisites of spatially skilled female models can be effectively used to reduce the gender gap in the spatial domain. The present findings suggest that such interventions should be enriched by offering opportunities for girls to make new spatial experiences on their own, in an optimal scenario guided by inspiring female mentors, who display high, but attainable, spatial skills, thus enabling each girl to tell her own spatial-adventures story.

ACKNOWLEDGMENT

This study was supported by the German Research Foundation (DFG QU/96 4-1/JA 889/7-1).We thank Madeleine Stein, Marie-Christin Wagner, and Viktoria Goergen for administering the tests and the intervention procedure, and all children and teachers of the participating schools.

REFERENCES

[1] Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*, 817-835.

[2] Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin, 136*, 103-127.

[3] Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 25*, 494-495.

[4] 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.

[5] Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: A meta-analysis. *Psychonomic Bulletin and Review, 18*, 267-277.

[6] Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgements are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior, 39*, 990-997.

[7] Moè, A., & Pazzaglia, F. (2006). Following the instructions! Effects of gender beliefs in mental rotation. *Learning and Individual Differences, 16*, 369-377.

[8] Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

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

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

[11] Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations. A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599-604.

[12] Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition, 28*, 39-58.

[13] Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th edition). New York: Psychology Press.

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

[15] Johnson, E. S., & Meade, A. C. (1987). Developmental patterns of spatial ability: An early sex difference. *Child Development, 58*, 725-740.

[16] Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35*, 940-949.

[17] Quaiser-Pohl, C. (2003). The mental cutting test “Schnitte” and the picture rotation test – two new measures to assess spatial ability. *International Journal of Testing, 3*, 219-231.

[18] Moore, D. S. & Johnson, S. P. (2008). Mental rotation in human infants: a sex difference. *Psychological Science, 19*, 1063-1066.

[19] Quinn, P. C. & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science, 19*, 1067-1070.

[20] Titze, C., Jansen, P., & Heil, M. (2010). Mental rotation performance and the effect of gender in fourth graders and adults. *European Journal of Developmental Psychology, 7*, 432-444.

[21] Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents’ mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences, 50*, 1238-1242.

[22] Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. C. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science, 16*, 841-845.

[23] Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S. & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1-51.

[24] Hausmann, M., Schoofs, D., Rosenthal, H. E. S., & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences. A psychobiosocial approach. *Psychoneuroendocrinology, 34*, 389-401.

[25] Pezaris, E., & Casey, M. B. (1991). Girls who use “masculine” problem-solving strategies on a spatial task: proposed genetic and environmental factors. *Brain and Cognition, 17*, 1-22.

[26] Grimshaw, G.M., Sitarenios, G., & Finegan, J.-A.K. (1995). Mental rotation at 7 years: relations with prenatal testosterone levels and spatial play experiences. *Brain and Cognition, 29*, 85-100.

[27] Hines, M., Fane, B. A., Pasterski, V. L., Mathews, G. A., Conway, G. S., & Brooke, C. (2003). Spatial abilities following prenatal androgen abnormality: targeting and mental rotations performance in individuals with congenital adrenal hyperplasia. *Psychoneuroendocrinology, 28*, 1010-1026.

[28] Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Biosocial factors, sexual orientation and neurocognitive functioning. *Psychoneuro-endocrinology, 29*, 867-881.

[29] Halari, R., Hines, M., Kumari, V., Mehrotra, R., Wheeler, M., Ng, V., & Sharma, T. (2005). Sex differences and individual differences in cognitive performance and their relationship to endogenous gonadal hormones and gonadotropins. *Behavioral Neuroscience, 119*, 104-117.

[30] Moè, A., & Pazzaglia, F. (2010). Beyond genetics in mental rotation test performance. The power of effort attribution. *Learning and Individual Differences, 20*, 464-468.

[31] Moè, A. (2012). Gender difference does not mean genetic difference: Externalizing improves performance in mental rotation. *Learning and Individual Differences, 22*, 20-24.

[32] Estes, Z., & Felker, S. (2011) Confidence mediates the sex difference in mental rotation performance. *Archives of Sexual Behavior*, DOI 10.1007/s10508-011-9875-5.

[33] Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2012, June 4). The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. *Psychological Bulletin*. Advance online publication. doi: 10.1037/a0028446.

[34] Newcombe, N. S. (2007). Taking science seriously: Straight thinking about spatial sex differences. In Ceci, S. J., & Williams, W. M. (Eds.), *Why aren't more women in science: Top researchers debate the evidence* (pp. 69-77). Washington, D.C.: American Psychological Association.

[35] Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850-855.

[36] Wiedenbauer, G., & Jansen-Osmann, P. (2008). Manual training of mental rotation in children. *Learning and Instruction, 18*, 30-41.

[37] Tzuriel, D., & Egozi, G. (2010). Gender differences in spatial ability of young children. The effects of training and processing strategies. *Child Development, 81*, 1417-1430.

[38] Quaiser-Pohl, C. Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences, 40*, 609-619.

[39] Cherney, I. D., & London, K. (2006). Gender-linked differences in the toys, television shows, computer games, and outdoor activities of 5- to 13-year-old children. *Sex Roles, 54*, 717-726.

[40] Taylor, L. D., & Setters, T. (2011). Watching aggressive, attractive, female protagonists shapes gender roles for women among male and female undergraduates viewers. *Sex Roles, 65*, 35-46.

[41] Weinraub, M., Clemens, L. P., Sockloff, A., Ethridge, T., Gracely, E., & Myers, B. (1984). The development of sex role stereotypes in the third year: Relationships to gender labeling, gender identity, sex-typed toy preference, and family characteristics. *Child Development, 55*, 1493-1503.

[42] Nash, S. C. (1979). Sex role as mediator of intellectual functioning. In M. A. Wittig & A. C. Petersen (Eds.), *Sex-related differences in cognitive functioning* (pp. 263−302). New York: Academic Press.

[43] Steele, C.M. (1997). A threat in the air. *American Psychologist, 52*, 613-629.

[44] Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology, 69*, 797-811.

[45] 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.

[46] Sharps, M. J., Welton, A. L., & Price, J. L. (1993). Gender and task in the determination of spatial cognitive performance. *Psychology of Women Quarterly, 17*, 71-83.

[47] McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486-493.

[48] Moè, A. (2009). Are males always better than females in mental rotation? Exploring a gender belief explanation. *Learning and Individual Differences, 19*, 21-27.

[49] Ambady, N., Shih, M., Kim, A. & Pittinsky, T.L. (2001). Stereotype susceptibility in children: Effects of identity acitivation on quantitative performance. *Psychological Science, 12*, 385-390.

[50] Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2012). A threat in the classroom: Gender stereotype activation and mental-rotation performance in elementary-school children. *Zeitschrift für Psychologie, 220*, 61-69.

[51] Muzzatti, B. & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology, 43*, 747-759.

[52] Hartley, B. L. & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development*. Advanve online publication. doi: 10.1111/cdev.12079.

[53] Ganley, C. M., Mingle, L. A., Ryan, A. M., Ryan, K., Vasilyeva, M., & Perry, M. (2013). An examination of stereotype threat effects on girls' mathematic performance. Developmental Psychology. Advance online publication. doi: 10.1037/a0031412.

[54] Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106*, 676-713.

[55] Lockwood, P. (2006). "Someone like me can be successful": Do college students need same-gender role models? *Psychology of Women Quarterly, 30*, 36-46.

[56] Wilder, D. A. (1984). Intergroup contact: The typical member and the exception to the rule. *Journal of Experimental Social Psychology, 20*, 177-194.

[57] Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology, 45*, 961-977.

[58] Ruble, D. N., & Martin, C. L. (1998). Gender development. In Damon, W. & Eisenberg, N. (Eds.), *Handbook of Child Psychology. 5th edition, Volume 3: Social, emotional, and personality development* (pp.943-965). New York: John Wiley & Sons, Inc.

[59] Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: a meta-analytic review. *Developmental Review, 13*, 147-183.

[60] Katz, P. A. & Ksansnak (1994). Developmental aspects of gender role flexibility and traditionality in middle childhood and adolescence. *Developmental Psychology, 30*, 272-282.

[61] Welch-Ross, M., & Schmidt, C. (1996). Gender-schema development and children's constructive story memory: Evidence for a developmental model. *Child Development, 67*, 820-835.

[62] Conkright, L., Flannagan, D., Dykes, J. (2000). Effects of pronoun type and gender role consistency on children's recall and interpretation of stories. *Sex Roles*, 43, 481-497.

[63] Ashby, M. S., & Wittmaier, B. C. (1978). Attitude changes in children after exposure to stories about women in traditional or nontraditional occupations. *Journal of Educational Psychology, 70*, 945-949.

[64] Scott, K. P., & Feldman-Summers, S. (1979). Children's reactions to textbook stories in which females are portrayed in traditionally male roles. *Journal of Educational Psychology, 71*, 396-402.

[65] Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin, 28*, 1183-1193.

[66] Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*, 545-560.

[67] McIntyre, R., Paulsen, R., Taylor, C., Morin, A. L., & Lord, C. G. (2011). Effects of role model deservingness on overcoming performance deficits induced by stereotype threat. *European Journal of Social Psychology*, *41*, 301-311.

[68] Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of in-group versus out-group social comparisons on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, 36, 519-530.

[69] Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*, 91-103.

[70] Bagès, C., & Martinot, D. (2011). What is the best model for girls and boys faced with a standardized mathematics evaluation situation: A hardworking role model or a gifted role model? *British Journal of Social Psychology, 50*, 536-543.

[71] Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*, 1615-1628.

[72] McArthur, L. Z., & Eisen, S. V. (1976). Achievements of male and female storybook characters as determinants of achievement behavior by boys and girls. *Journal of Personality and Social Psychology, 33*, 467-473.

[73] Ochman, J. M. (1996). The effects of nongender-role stereotyped, same-sex role models in storybooks on the self-esteem of children in grade three. *Sex Roles, 35*, 711-735.

[74] Connor, J. M., & Serbin, L. A. (1978). Children's responses to stories with male and female characters. *Sex Roles, 4*, 637-645.

[75] Wood (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, 106, 231-248.

[76] Berk, L. E. (2004). *Development through the lifespan* (3rd ed.). Boston: Allyn and Bacon.

[77] Ruble, D. N. (1987). The acquisition of self-knowledge: A self-socialization perspective. In N.Eisenberg (Ed.), *Contemporary Topics in Developmental Psychology* (pp. 243-270). Wiley-Interscience, New York.

[78] Miller, C. F., Lurye, L. E., Zosuls, K. M., & Ruble, D. N. (2009). Accessibility of gender stereotype domains: Developmental and gender differences in children. *Sex Roles, 60*, 870-881.

[79] Trautner, H. M., Helbing, N., Sahm, W. B., & Lohaus, A. (1988). Unkenntnis - Rigidität - Flexibilität: Ein Entwicklungsmodell der Geschlechtsrollen-Stereotypisierung. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie, 19*, 105-120.

[80] Titze, C., Jansen, P., & Heil, M. (2010). Mental rotation performance in fourth graders: No effect of gender beliefs (yet)? *Learning and Individual Differences, 20*, 459-463.

[81] Schauder, T. (1996). *Aussagen-Liste zum Selbstwertgefühl für Kinder und Jugendliche.* Göttingen: Beltz.

[82] Heller, K., & Geisler, H. - J. (1983). *KFT 1-3. Kognitiver Fähigkeits-Test (Grundschulform)*. Weinheim: Beltz Test GmbH.

[83] Jöckel, K.- H., Babitsch, B., Bellach, B.- M., Bloomfield, K., Hoffmeyer-Zlotnik, Winkler, J., & Wolf, C. (1998). Messung und Quantifizierung soziodemografischer Merkmale in epidemiologischen Studien. In W. Ahrens, B.-M. Bellach & K.-H. Jöckel (Eds.), *Messung soziodemografischer Merkmale in der Epidemiologie* (S. 7-38). München: MMV Münchner Medizin Verlag.

[84] Ruthsatz, V., Neuburger, S., & Quaiser-Pohl, C. (in press). The social relevance and the socio-cultural origins of gender differences in spatial abilities. *Folia Sociologica*.

[85] Halpern, D. F., Straight, C. A., & Stephenson, C. L. (2011). Beliefs about cognitive gender differences: accurate for direction, underestimated for size. *Sex Roles, 64*, 336-347.

[86] Neuburger, S., Heuser, V., Jansen, P., & Quaiser-Pohl, C. (2012). Influence of rotational axis and gender-stereotypical nature of rotation stimuli on the mental-rotation performance of male and female fifth graders. In C. Stachniss, K. Schill, & D. H. Uttal (Eds.), *Spatial Cognition 2012*, LNAI 7463 (pp. 220-229). Berlin: Springer-Verlag.

[87] Casey, B. M., Pezaris, E. E., & Bassi, J. (2012). Adolescent boys' and girls' block constructions differ in structural balance: A block-building characteristic related to math achievement. *Learning and Individual Differences, 22*, 25-36.

[88] Nguyen, H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93,* 1314-1334.

[89] Potter, B. A. (1978). Sex of the protagonist in children's storybooks: effect on self-concept. *Journal of Instructional Psychology, 5*, 6-14.

[90] Huston, A. C., Wright, J. C., Fitch, M., Wroblewski, R., & Piemyat, S. (1997). Effects of Documentary and fictional television formats on children's acquisition of schemata for unfamiliar occupations. *Journal of Applied Developmental Psychology, 18*, 563-585.

[91] Bandura, A. (1974). Analysis of modeling processes. In A. Bandura (Ed.), *Psychological Modeling: Conflicting Theories*. New York, N. Y.: Lieber-Atherton.