Advanced Review

Sex differences in spatial cognition: advancing the conversation

Susan C. Levine,* Alana Foley, Stella Lourenco, Stacy Ehrlich and Kristin Ratliff

The existence of a sex difference in spatial thinking, notably on tasks involving mental rotation, has been a topic of considerable research and debate. We review this literature, with a particular focus on the development of this sex difference, and consider four key questions: (1) When does the sex difference emerge developmentally and does the magnitude of this difference change across development? (2) What are the biological and environmental factors that contribute to sex differences in spatial skill and how might they interact? (3) How malleable are spatial skills, and is the sex difference reduced as a result of training? and (4) Does 'spatializing' the curriculum raise the level of spatial thinking in all students and hold promise for increasing and diversifying the STEM pipeline? Throughout the review, we consider promising avenues for future research.

How to cite this article:
WIREs Cogn Sci 2016. doi: 10.1002/wcs.1380

INTRODUCTION

The question of whether males and females differ in their cognitive abilities, and if so, what factors contribute to these differences, has received substantial attention in the research literature as well as the popular press, especially since the publication of Maccoby and Jacklin's seminal book, *The Psychology of Sex Differences*.[1] For example, Larry Summers, then the President of Harvard University, delivered a speech to the U.S. National Bureau of Economic Research in which he offered some reasons for the underrepresentation of women in science and engineering professions. In his speech, Summers said: '… in the special case of science and engineering, there are issues of intrinsic aptitude, and particularly of the variability of aptitude, and that those considerations are reinforced by what are in fact lesser factors involving socialization and continuing discrimination.'[2] The debates and publications that followed Summers' speech clearly show that the question of whether there are male–female cognitive differences is a hot-button issue.[3,4] Not only does this issue raise questions about gender equality and equity, but it also highlights the importance of systematic scientific research in order to fully understand sex differences in what Summers referred to as 'intrinsic aptitude.'

In this review, we focus on a sex difference on a particular cognitive task—mental rotation. Although sex differences favoring males have been reported on many spatial tasks (mental rotation;[5–7] navigation;[7] discrimination of line orientation;[8] and Piaget's water level task[9]), the largest sex difference has been found for mental rotation.[5–7] We address mental rotation in this review because of the robustness of this sex difference and the large body of research that has accumulated about it. Mental rotation is also of interest because of evidence that it is related to STEM (science, technology, engineering, and math) achievement,[10–16] and evidence that it is malleable.[17,18] However, many of the issues we discuss with respect to mental rotation are relevant to spatial skills more generally.

---

*Correspondence to: s-levine@uchicago.edu

Department of Psychology, University of Chicago, Chicago, IL, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.

We organize the review into four sections that examine major research questions and in each, consider ways to 'advance the conversation' concerning the sex difference in mental rotation. In each section, we also apply a developmental lens to the issues surrounding sex differences in mental rotation. In the first section, we review evidence that a sex difference in mental rotation exists in adults, ask when during development it first emerges, and whether the magnitude of this differences changes over the course of development. In the second section, we review explanations for the sex difference, which broadly include biological and experiential factors, with attention to the ways that that these factors may interact in influencing the development of spatial cognition. Third, we consider the issue of malleability—namely whether spatial thinking can be improved through training, whether training results in a maintenance, widening, or narrowing of the sex difference, and whether the answer to this question depends on the intensity of the training and the age of the participants. Last, we consider educational implications of existing findings concerning sex-related variations in spatial skill in view of findings that spatial thinking is malleable. In particular, we ask whether 'spatializing' the curriculum might raise the level of spatial thinking of all students, whether it is important to start supporting and educating spatial thinking early in the school years, and whether such efforts might increase the size and diversity of the pipeline into STEM.

SEX DIFFERENCES IN MENTAL ROTATION

Evidence for a Sex Difference in Mental Rotation in Adults

Mental rotation skill is typically measured by asking for judgments of whether rotational variants of objects, typically shown pictorially, represent the same object or are mirror images of each other.[19,20] The original way of measuring mental rotation is reported in the seminal Shepard and Metzler[19] study, where the mental rotation task involved judging whether perspective pictures of 3D block structures were the same figure or mirror images of each other. The larger the angle of rotation between the figures, the longer the response time, leading to the conclusion that people are engaging in an analog process of mentally rotating the figures.[19,21] Later, a paper and pencil task was devised to measure mental rotation skill—the widely used Vandenberg and Kuse, Mental Rotation Test (MRT).[22] The MRT involves selecting two perspective pictures of 3D block structures from among four that are the same as a target picture. Distractor pictures are either mirror images of the target picture or structurally different block structures. The MRT is a widely used measure of sex differences in mental rotation, and studies typically report a male advantage on this task, with meta-analyses reporting Cohen's *d* for this difference ranging from .50 to 1.28.[6,23]

Meta-analytic reviews, which have largely focused on adults and older children, provide convincing evidence that there is a substantial male advantage on mental rotation tasks, and that this sex difference is larger than that for other aspects of spatial cognition[6,23] as well as for other aspects of cognition more broadly.[24] Linn and Petersen's meta-analysis found evidence for three spatial factors, two of which showed sex differences. One of these was *Mental Rotation*, which involved mentally rotating figures in the picture plane or in depth, and this factor showed a large sex difference in favor of males, Cohen's $d = 0.73$ ($P < .05$). A second factor, *Spatial Perception*, involved the ability to determine spatial relations despite distracting information, exemplified by the Water Level Test and the Rod-and-Frame Test, and showed a medium size gender difference in favor of males, Cohen's $d = 0.44$, $P < .05$). Finally, another factor, *Spatial Visualization*, involved manipulating spatial information when several stages are needed to produce the correct answer and is exemplified by tasks such as Paper Folding, Embedded Figures, and Block Design. In contrast to *Mental Rotation* and *Spatial Visualization*, this factor did not show a significant sex difference, Cohen's $d = 0.13$ ($P > .05$). These findings have been corroborated by a subsequent meta-analytic review.[6]

Linn and Petersen's meta-analysis also uncovered an interesting pattern within the category of mental rotation tasks. In particular, there was a much larger sex difference on the MRT,[19,22] which involves rotating complex block structures in three dimensions (Cohen's $d = 0.94$), compared to the Primary Mental Abilities Space Test,[25] which involves rotating simpler shapes in the picture plane (Cohen's $d = 0.26$). This effect size difference may indicate that females have particular difficulty mentally rotating objects in depth compared to males.

Other studies show a larger sex difference in favor of males when the objects being rotated are more complex, a finding that may reflect the different strategies that males and females apply to mental rotation tasks. For example, in one study, the mental rotation speed of males was unaffected by complexity (operationalized as the number of vertices the shape had) whereas that of females was slower for more complex shapes, resulting in a larger sex difference for the complex shapes.[26] The authors suggested that the absence of a complexity effect for males reflects their use of a holistic rotation strategy whereas the presence of a complexity effect for women reflects their use of a piecemeal, analytic strategy. The piecemeal strategy might cause difficulty on complex shapes because it involves encoding, rotating, and comparing more individual features.[26–28] Consistent with this idea, on the MRT, females' error patterns are more likely than those of males to reflect the use of a verbal analytic strategy.[27] A caveat, however, comes from a recent study suggesting that objects are not rotated holistically by anyone, but rather by focusing on a particular feature of the object and rotating that feature, suggesting that that the piecemeal versus holistic characterization of female–male differences may not be accurate.[29] Additionally, complexity effects are not always found, and appear to depend on whether participants need to rotate the comparison stimulus in order to determine whether it is the same or different from the target or whether they can make this judgment without engaging in mental rotation.[30,31]

It is important to note that both males and females show the signature increase in reaction time with angular disparity on Shepard–Metzler type mental rotation tasks, suggesting that both groups are engaging in mental rotation (e.g., Refs 32,33; Hooven, Chabris, Ellison, Kievit, & Kosslyn, unpublished data). Moreover, these studies show that the sex difference on the Shepard–Metzler task is not in the slope of reaction times as a function of angular disparity—thought to reflect mental rotation—but rather in the intercept, which is thought to reflect encoding of the shapes, preparation for rotation, decision making, and/or response rather than mental rotation *per se*.[33] In fact, there is evidence to suggest that females have difficulty with forming 3D representations based on 2D images.[34] Thus, it is possible that sex differences on mental rotation tasks may not actually be attributable to a sex difference in mental rotation, but rather to the surrounding processes involved in this task such as encoding the stimulus to be rotated or in decision processes.

We turn next to the question of when during development the sex difference on mental rotation tasks first emerges, and whether the magnitude of the sex difference increases with age. In considering these questions, we keep in mind findings that the type of mental rotation task (picture plane vs in depth; simple vs complex shape) may influence the size of the sex difference. In addition, we consider whether other task demands or procedural variations may influence the presence or magnitude of the sex difference.

Development of the Sex Difference in Mental Rotation

Infant Studies

It was once believed that sex differences on spatial tasks such as mental rotation emerge relatively late, around the time of puberty.[1] Recent studies, however, report a sex difference in favor of males much earlier, as early as infancy.[35–38] Quinn and Liben[37] administered a task in which 3- to 4-month-old infants were familiarized with an identical pair of static stimuli—the numeral '1'—either both in the canonical orientation or both in the mirror image orientation. Each infant was shown seven familiarization trials consisting of identical pairs of the numeral (see Figure 1), presented at seven of eight randomly selected angles (equally spaced from 0° to 315°), with the eighth angle reserved for the test trials. Following

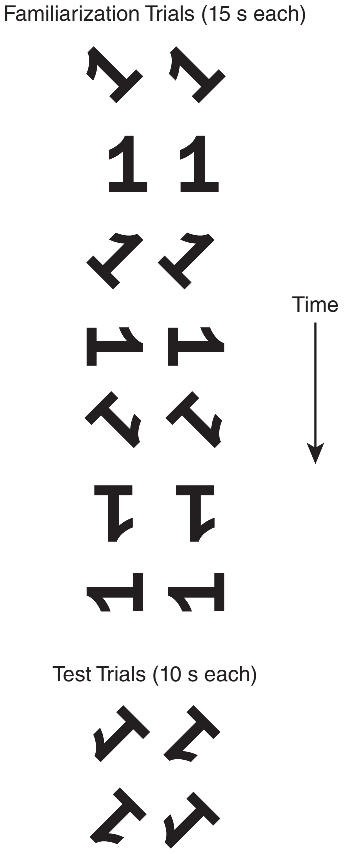

FIGURE 1 | Stimuli shown to infants in the study by Quinn and Liben.[37] The series of figures at the top depict the familiarization stimuli and the figures at the bottom depict the test trials.

familiarization, infants were presented with two test trials in which the two images of the numeral '1' were mirror images of each other rather than identical as was the case during familiarization trials. The researchers measured the amount of time infants looked at the version of the numeral 1 shown during the familiarization trials versus the amount of time they looked at the novel, mirror image of the numeral. Findings showed that boys looked at the novel image 62.6% of the time, which differed significantly from chance, compared to 50.2% of the time for girls, which did not differ from chance. Moreover, boys' novelty preference was significantly greater than that of girls (Cohen's $d = 1.38$).

In another infant study, Moore and Johnson[36] habituated 5-month olds to a simplified Shepard and Metzler cube figure dynamically rotating in depth through 240° of arc around the vertical axis (Figure 2). Once infants habituated (or were shown 12 habituation trials) they were presented, in alternating order, with the same object or its mirror image rotating through the previously unseen 120° of arc. Boys looked significantly longer at the novel (mirror image) stimulus than at the familiar (habituated) stimulus (6.16 s vs 4.16 s, respectively), but girls did not show a significant difference in looking times at the novel and familiar stimulus (6.87 s vs 7.19 s, respectively). As in the Quinn and Liben[37] study, boys' novelty preference was significantly greater than that of the girls (Cohen's $d = 0.67$).

In a follow-up study,[38] Moore and Johnson used the same stimuli with even younger infants—3-month olds. In this study, boys showed the opposite preference, looking significantly longer at the familiar (habituated) stimulus rather than at the novel (mirror image) stimulus (17.50 s vs 13.73 s, respectively, Cohen's $d = 0.46$), whereas girls again showed no difference in looking times (11.66 s vs 10.84 s, respectively, Cohen's $d = 0.13$). Boys' longer looking at the familiar stimulus at 3 months of age and the novel stimulus at 5 months of age may reflect the fact that the younger infants are still gleaning information from the familiar stimulus at the time of the test trials due to their slower processing speed, and hence are still interested in it (e.g., Refs 39–41). Moreover, a comparison of Moore and Johnson's findings to those of Quinn and Liben[37] suggests that the preference seen in 3-month-old boys may vary as a function of stimulus complexity: whereas the 3-month-old male infants in Moore and Johnson's study showed a familiarity preference for depicted 3D block structures rotating in depth, male infants of about the same age in Quinn and Liben's study showed a novelty preference for simpler 2D numeral stimuli rotating in the picture plane. The hypothesis that looking time patterns vary with stimulus complexity could be easily tested in a study that compares 3-month-old infants' looking time patterns for 3D block structures rotating in depth to that of 2D stimuli rotating in the picture plane.

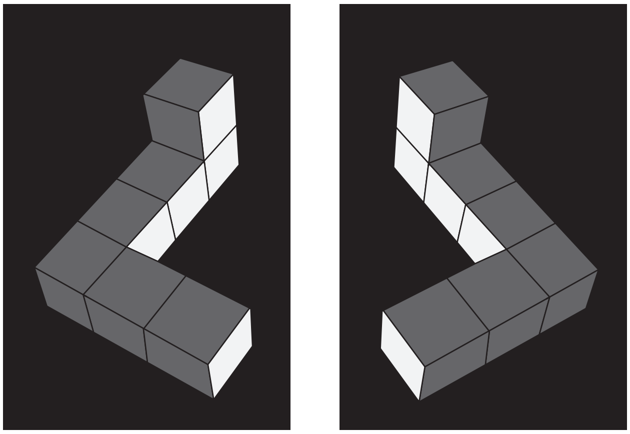

FIGURE 2 | Shepard–Metzler objects shown to infants in study by Moore and Johnson.[36]

A third infant study used a different paradigm to assess the sensitivity of 6- to 13-month-old infants to mirror images.[35] Infants' were simultaneously shown two streams of a two-dimensional *Tetris*-shaped figure. One stream consisted of the identical shape shown in various orientations. In the other stream, every third shape was a mirror image of the previous two. Longer looking at the second stream was interpreted as an index of sensitivity to mirror images. In this study, both male and female infants looked longer at the stream containing mirror images, suggesting that they both detected the mirror images. However, in line with the previous results, the preference for the mirror image was significantly stronger for the male than the female infants (Cohen's $d = 0.57$).

In summary, three different laboratories using different methods report that boys discriminate mirror-imaged stimuli some time during the first year of life,[35–38] whereas girls either do not show evidence of making this discrimination[36–38] or show less robust discrimination than boys.[35] Although it is tempting to conclude from these studies that infant boys show a mental rotation advantage compared to infant girls, the lack of a looking time difference for infant girls in some of the studies[36–38] does not provide definitive evidence that infant girls are not able to discriminate mirror images from nonmirror images. There are many reasons why infants may not look longer at the novel mirror image stimulus (or at the familiar stimulus at younger ages). For example, they may find both test stimuli interesting—after all, both are presented at an angle that was not seen during the habituation trials: in the Quinn and Liben study,[37] during test trials, both the mirror image and familiarized version of the numeral were presented at an angle that had not been seen previously, and in both of the Moore and Johnson studies,[36,38] infants were shown familiar (habituated) and unfamiliar (mirror image) test stimuli rotating back and forth through a previously unseen 120° of arc. The possibility that female infants regard the previously presented and the mirror imaged stimuli as novel because both appeared in new locations would be supported if girls' looking times during test trials were longer than their looking times at the end of habituation or familiarization trials for both novel and familiar stimuli. This pattern would be consistent with evidence that female adults show an advantage over male adults in location coding.[42,43] A comparison of looking times between familiarization trials and test trials could address this possibility, but these data were not reported, i.e., infant girls, more than infant boys, might show increased looking time to both the mirror image and nonmirror imaged stimulus because of the location change.

Quinn and Liben[44] recently tried to rule out this possibility in a study where they showed that both male and female infants are sensitive to the novel location of the originally shown numeral. In particular, they showed that both sexes looked longer when the familiarized stimulus was presented in a novel location (previously unused 120° of arc) versus in a previously seen location (previously used 240° of arc). However, this finding does not rule out the possibility that females find the change in location more novel than the change in rotation. That is, for female infants, location novelty may trump the novelty response to the mirror image, even though they can mentally rotate and discriminate the familiar stimulus from its mirror image. In contrast, for male infants, the novelty of the mirror image may trump location novelty, even though they can discriminate new from old locations. This possibility would be consistent with a sex difference, but not one that reflects an ability of male but not female infants to mentally rotate figures.

Relatedly, more evidence is needed that male infants are actually engaging in mental rotation to discriminate these stimuli rather than noticing more local relations between the parts of the stimulus. For example, it has been suggested that presenting stimuli in various orientations may lead infants to develop 'orientation independent' representations of the familiarized stimulus that would enable them to view the mirror image as novel but that would not entail the ability to mentally rotate the stimulus.[35] Paradigms that vary the angle of rotation would be helpful in addressing the question of whether infants are engaging in mental rotation by providing information as to whether they show the classic signature pattern of mental rotation—more difficulty on stimuli with greater angular disparities. Such studies also could reveal whether there is a sex difference in the slope of this function, which would provide convincing evidence for a sex difference in mental rotation.

It is worth noting that not all infant studies report a sex difference in mental rotation, even though null findings are difficult to interpret.[38,45–50] Möhring and Frick[51] suggest that the mixed findings could be due to methodological differences. One such difference is that infant studies that do not find a sex difference generally employ mental rotation tasks with rich information about 3D forms (e.g., videos of objects moving on a puppet stage or live presentations of 3D physical objects) which, in some cases, infants could touch prior to testing, whereas those that do report a sex difference generally use stimuli that provide less information about 3D forms (e.g., 2D computer graphic representations of 3D stimuli[36]). Under this hypothesis, the sex difference reported in some infant studies may be attributable to differential difficulty females have in forming 3D representations from 2D representations, a pattern that is consistent with the intercept difference seen in studies of male and female adults' performance on mental rotation tasks (e.g., Refs 33,34).

In sum, although the idea that a sex difference on mental rotation tasks is present during infancy is intriguing and is supported by some studies, further investigation is needed in order to gain a full understanding of the nature of this difference. The existing literature provides some hints that a sex difference in ability to form 3D representations from 2D representations may play a role in the infant sex difference in mental rotation, and this possibility is in need of systematic investigation.

Preschool Children

Studies of preschool children also provide somewhat mixed evidence with respect to a sex difference in mental rotation skill. On the Children's Mental Transformation Task (CMTT),[52] which involves mental transformation of 2D shapes in the picture plane—both translation and rotation—a male advantage was found as early as 4.5 years of age.[52,53] Levine et al.[52] showed children pictures of two halves of a shape and asked them which of four shapes the pieces would make if they were moved together (see Figure 3, e.g., of stimuli). Children ranging in age from 4.5 to 7 years showed a male advantage (Cohen's d = 0.25). Importantly, the level of performance of boys and girls overlapped considerably, and the highest performing girls scored at the level of the highest performing boys (see Figure 4). Although the CMTT was designed to engage mental translation and rotation skills, it is possible to solve the problems using feature matching. In fact, in a related study that used the same mental transformation task, Ehrlich and colleagues found that boys gestured about moving the pieces more than girls when asked to explain how they carried out the task (a behavior that was associated with better performance), suggesting that there may also be a sex difference in how children approach this task.

More recently, Frick and colleagues assessed whether there is a sex difference in a 'classic' mental rotation skills in 3- to 5-year-old children, building on older studies of this type that did not examine sex differences (e.g., Refs 54 and 55).[56,57] In one such study, Frick et al.[56] used a novel puzzle task in which children were asked to select which of two mirror image ghost pieces fit into a puzzle by mentally rotating the pieces. Results showed no sex difference in 3- to 5-year olds, but considerable development across this period, with only 10% of 3-year olds responding at above chance levels compared to 95% of 5-year olds (mean percent correct was 54% for 3-year olds versus 83% for 5-year olds, where 50% correct was chance level). In a related study,[57] 4- and 5-year olds had to indicate which hole a piece would fit into to complete a road on a touch screen. Five-year olds, but not 4-year olds, showed increasing reaction times and error rates with increasing angular disparity between the puzzle piece and its placement hole, a hallmark of mental rotation skill. Furthermore, on this mental rotation task, 5-year-old boys outperformed 5-year-old girls but there was no sex difference for 4-year olds, likely because the younger group's performance was near chance.

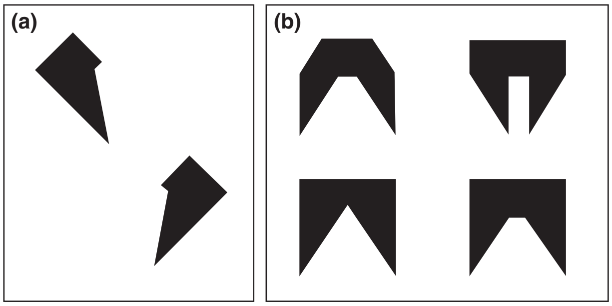

FIGURE 3 | Depiction of the Children's Mental Transformation Task, a 2D mental rotation task used by Levine et al.[52] The child is asked to select the figure from among four choices (b) that can be made from the two pieces shown on the target card (a).

School-Aged Children

Other studies have examined sex differences in mental rotation skill in school-aged children using tasks that involve mentally rotating stimuli in the picture plane or in depth, and discriminating mirror-imaged stimuli from each other. For example, Neuburger et al.[58] assessed second and fourth graders on an MRT that involved deciding whether various stimulus types (animals, letters, and Shepard–Metzler block figures), rotated in the picture plane, were the same or mirror images (Figure 5). They found a sex by grade interaction, such that boys outperformed girls in the fourth grade (Cohen's d = .21) but not in the second grade (Figure 6). There were no interactions involving stimulus type but children performed best for the rotating animals' task, followed by letters, and then Shepard–Metzler block figures. Titze et al.[59] examined sex differences in fourth graders' (ages 9–10) mental rotation skill using a task similar to the MRT that involved rotating Shepard-Metzler block figures in depth. They found a significant male advantage in 10-year-old 4th graders (Cohen's d = 1.47; M_{age} = 10.3 years), but a smaller, nonsignificant sex difference in 9-year-old 4th graders (Cohen's d = 0.45, M_{age} = 9.3 years). They suggest that the lack of a significant sex difference in the younger children might be related to hormonal changes associated with puberty, although they did not measure pubertal stages.

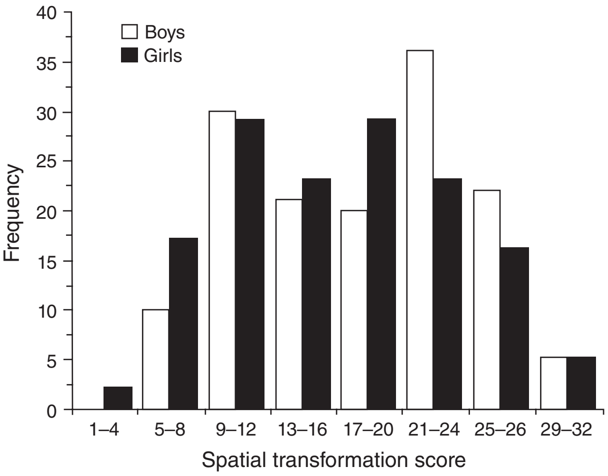

FIGURE 4 | Distributions of girls' and boys' spatial transformation scores in the study by Levine et al.[52]

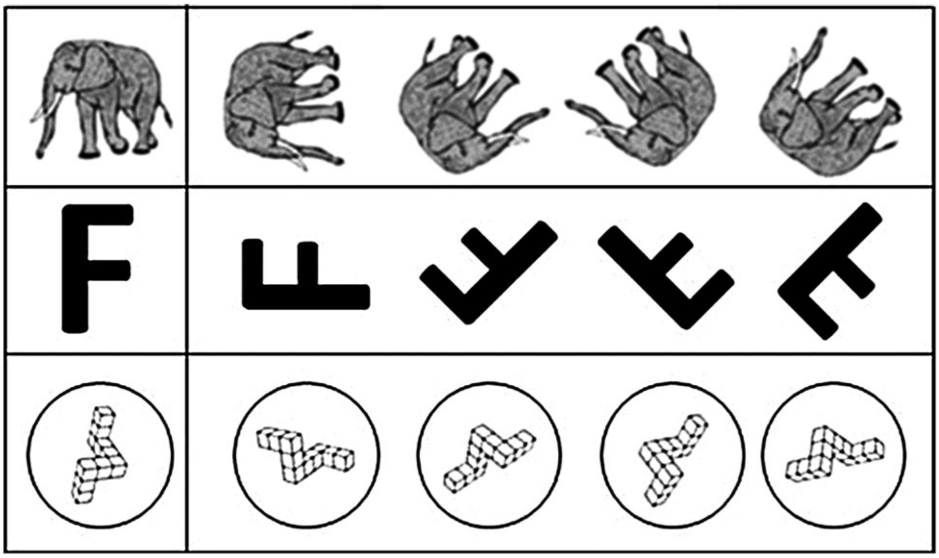

FIGURE 5 | Animal pictures, letters, and block figures used as mental rotation stimuli in the study by Neuburger et al.[58] For each type of stimulus, two of the four choices can be rotated to match the target image shown on the left.

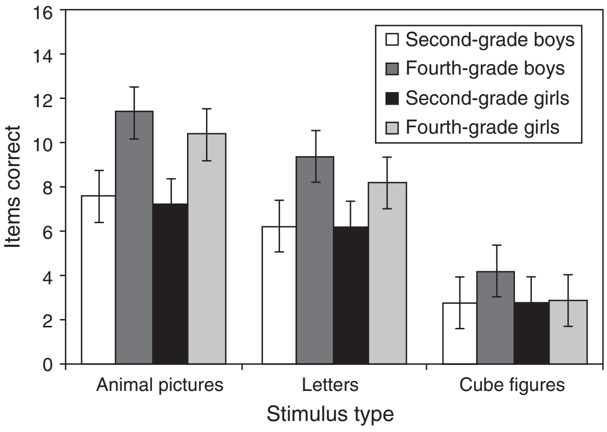

FIGURE 6 | Accuracy on mental rotation tasks as a function of stimulus type (animal pictures, letters, and cube figures), sex, and grade level in the study by Neuburger et al.[58] Error bars indicate 95% confidence intervals.

Age-Related Changes in the Size of Sex Differences

A meta-analysis carried out by Voyer and colleagues[6] showed a male advantage on mental rotation tasks that increased in size over development. In this meta-analysis, the effect size (Cohen's d) of the male advantage in mental rotation for children younger than 13 was 0.33, whereas for 13- to 18-year-olds, it was 0.45, and for adults, it was 0.66. However, as suggested earlier, apparent age-related increases in the magnitude of the sex difference in mental rotation may be attributable to differences in the nature of the mental rotation tasks given to different age groups. Recall the findings, reviewed above, that adults show smaller sex differences for less complex mental rotation problems, just the kind of problems that researchers tend to present to young children. That is, even in adults, sex differences in mental rotation are larger when the rotation is in depth rather than in the picture plane and when the rotation involves more complex and/or less familiar stimuli.[6,23] Linn and Petersen[23] note that the MRT is rarely given to children under 13 years of age, and that when task differences are taken into account, there is no evidence that the magnitude of the sex difference in mental rotation increases with age. In children, an interesting pattern emerges from comparing the effect sizes of the sex differences on the mental rotation tasks used by Neuburger et al.[58] and Titze et al.[59]. The tasks used were similar except that Neuburger et al.'s task involved rotating Shepard–Metzler block figures in the picture plane and Titze et al.'s task involved rotating these figures in depth. The effect size of the male advantage for fourth graders (mean age = 9.9 years) reported in the Neuburger et al.'s study was d = 0.21 whereas the effect size reported in the Titze et al.'s study was d = 1.47 for the older fourth graders (mean age = 10.3 years). Although further study is needed, these findings raise the possibility that the larger sex difference for the older fourth graders in Titze et al.'s study is due to the differential difficulty females have with rotating objects in depth.

Answering the question of when the sex difference in mental rotation emerges and whether the magnitude of the sex difference in mental rotation changes over the course of development is challenging for a number of reasons. First, as discussed, the mental rotation tasks given to young children tend to be easier than those given to older children, often consisting of familiar forms rotated in the 2D plane compared to complex forms rotated in 3D (e.g., Refs 52 and 58), and this factor has been shown to relate to the effect size of sex differences in mental rotation in adults (e.g., Ref 23). Second, even when younger and older groups are given the same kinds of mental rotation problems, the processes they apply may differ, and this may account for changes in the magnitude of the sex difference that are observed in different age groups. For example, a study that paired a mental rotation task with a simultaneous motor rotation task that was congruent or incongruent in its movement to the required mental rotation, indicated that motor processes may play a larger role in mental rotation in younger than older children.[60] Consistent with the importance of manual experience and motor processes in mental rotation at early ages, Möhring and Frick[45] found that 6-month olds given manual experience with an asymmetrical object (e.g., shaped like the letter P) were able to discriminate this object from its mirror image whereas those infants who were not given the opportunity to manually explore the object showed no evidence of discriminating it from its mirror image. Thus, we do not know whether the size of the sex difference in mental rotation would vary with age if the tasks used were the same, or if different age groups recruited the same processes to solve the task, e.g., all age groups recruiting motor processes to the same extent.

Although there is some evidence that the sex difference in mental rotation is larger in older children and adults than in younger children,[6] this finding is not consistent with the large sex difference reported on mental rotation tasks in some infant studies.[36,44] Moreover, the finding of infant competence on mental rotation tasks when older children perform at chance on mental rotation tasks that require explicit selection of correct answers raises fundamental questions. Frick et al.[61] suggest that measures of mental rotation in infants capture spontaneous and immediate responses that may reflect implicit knowledge based on mental simulations, a kind of knowledge that may be too fragile or short-lived to inform the overt verbal or action-based responses required on mental rotation tasks given to children and adults. However, a new longitudinal study[62] reports that infants' performance on the Lauer et al. task, described above, predicts performance on the CMTT[52] at age 4 (Figure 7). Thus, it is possible that infant mental rotation processes lay the foundation for later mental transformation. Longitudinal studies that include a variety of different kinds of mental rotation tasks (e.g., implicit vs explicit; 2D vs 3D, etc.) can provide important information as to whether performance on different kinds of mental rotation tasks are related, either contemporaneously or predictively, and hold promise for advancing our understanding of the developmental trajectory of sex differences in mental rotation.

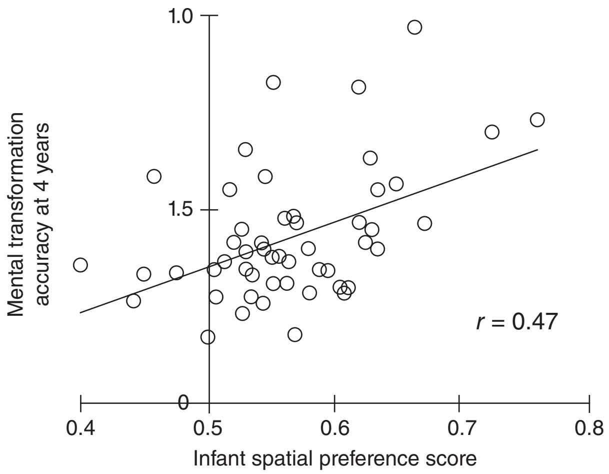

FIGURE 7 | Scatterplot showing association between visuospatial processing in infancy and mental transformation ability at 4 years of age.[62] Infant performance on a spatial change detection task as measured via preference to the stream with the mirror image (chance performance = 0.50), significantly predicted 4-year olds' accuracy (chance performance = 0.25, $r(51) = .47$, $P < .001$) on a mental transformation task.[52]

FACTORS CONTRIBUTING TO SEX DIFFERENCES IN SPATIAL SKILL

We next consider various factors that have been proposed as explanations for the sex difference in mental rotation and spatial skill more generally. These factors fall into two broad categories: biological (e.g., differences in prenatal hormones) and environmental (e.g., differences in play experiences). However, these categories do not capture the complex interactions between biology and experience. Therefore, we consider how these factors, which are frequently studied separately, might interact to produce different levels of spatial skill in girls and boys, as well as among individuals within these groups. Based on the evidence, we argue that not only do biological factors influence experiences but also that experiences influence biology, and that there are multiple routes to developing strong spatial thinking—no one factor is determinative. Individuals who have multiple factors favoring strong spatial skills—a biological advantage, experiences that include frequent and rich play that involves spatial problem solving (e.g., play with blocks and puzzles, certain videogames), membership in a group that is not negatively stereotyped in terms of spatial ability—are likely to develop strong spatial abilities. However, individuals with a biological advantage may not develop strong spatial abilities if they lack experiences that promote spatial thinking. And importantly, individuals may develop strong spatial abilities if they have rich spatial experiences, despite being in a group that is not biologically advantaged. As we will see, spatial thinking is malleable and can be improved in both males and females, a finding that has important educational implications.

Biological Factors

Evidence that sex differences in mental rotation can be explained, at least in part, by biological factors comes from a variety of sources. One is the presence of sex differences in toy preferences that are reported in human neonates as well as in nonhuman primates. A second is the influence that prenatal hormone differences—in particular, androgen levels—have on activity preferences and spatial thinking. Finally, a third kind of biological evidence involves findings of functional and structural sex differences in brain areas thought to be involved in mental rotation. However, it is possible that these neural differences arise from differences in spatially relevant experience, a question that has not been addressed in existing studies.

Are Early Emerging Toy Preferences Evidence for a Biologically Based Sex Difference?

A number of studies report that infant boys prefer inanimate over animate objects and that the reverse is true of infant girls. In an eye-tracking study, Alexander and colleagues[63] found a sex difference in the toy preferences of 3- to 8-month old infants. Whereas infant boys looked longer at a truck than a doll, infant girls showed the reverse preference. The authors argue that the young age of these infants limits the possibility that the observed preferences are products of gender socialization.

Consistent with this position, Baron-Cohen posits that early gender-related toy preferences reflect an innate difference in males' and females' psychological processing, such that males show a preference for 'systematizing' and females for 'empathizing,' which offers an explanation for the sex difference in spatial skills, including mental rotation.[64] In opposition to this position, Spelke[4] argues that the empirical basis for this claim is limited, citing the absence of sex differences in much of the infant literature. Furthermore, she critiques the Connellan et al.[65] study which, because it involved newborns, arguably offers the strongest support for a biological basis of the systematizer–empathizer dichotomy. Although Connellan et al. find that male newborns spent more time looking at an inanimate object than an expressive person whereas the reverse was true for female newborns, Spelke points out that this study did not control for experimenter bias and did not show that the results generalize to other stimuli from the two categories (i.e., other inanimate objects and expressive people). Another issue Spelke raised was the lack of replication of Connellan et al.'s findings. However, since the time of her critique, the Connellan findings have been replicated, albeit with older infants who have greater potential to be influenced by differential experiences.[35,63]

An additional problem with the systematizer–empathizer view is that it is not clear whether this difference translates to a sex difference in mental rotation. Very recently, however, a research group found that toy preference is correlated with mental rotation skill in 6- to 12-month old boys, but not girls.[35] Although environmental factors could play a role in toy preferences by 6 months of age, the authors argue for a biological explanation based on related findings in nonhuman primates[66] and in newborn infants.[65] Furthermore, Lauer et al.[35] propose a possible mechanism for the toy preference that is linked to differential prenatal exposure to androgens in males—specifically, that higher levels of prenatal androgen exposure in males may modulate the development of magnocellular and parvocellular layers of the thalamus.[67] Evidence from animal models suggests that prenatal androgens might accelerate the development of the magnocellular layer (which provides input to the dorsal, object location and motion stream) and delay the development of the parvocellular layer (which provides input to the ventral, object recognition stream).[68,69] If true, this would favor the earlier development of the dorsal stream in male infants and perhaps lead to the male preference for objects with moving parts such as trucks. This interest, in turn, is hypothesized to lead to enhanced mental rotation ability, although details are lacking about how this would unfold. In typically developing females, in contrast, prenatal androgen exposure may be too small to affect the relative timing of development of these neural systems, and thus females may not have the preference for moving parts and thus may have less developed mental rotation ability. Although providing a possible mechanism for the gender difference in toy preference and mental rotation, this hypothesis is in need of further empirical support.

Below, we discuss literature that has more directly implicated gonadal hormones in the sex difference in mental rotation. Of note, although most of the research on biological factors affecting spatial cognition has focused on hormonal differences, more recent approaches to studying sex differences has shown that it is important to consider differences in phenotypes caused by sex chromosome complements (XX vs XY) as well as hormonal differences.[70,71]

Gonadal Hormones

Levels of gonadal hormones such as testosterone have been implicated in sex differences on a variety of spatial tasks in both humans[72,73] and nonhuman animals.[74] Sex hormones may exert their influence prenatally, through shaping the organization of brain regions associated with spatial cognition, and/or postnatally, through activation of these brain regions. Here, we focus on investigations of prenatal influences on postnatal spatial cognition in humans as this pathway has the most evidence. We include studies examining whether androgen exposure levels relate to later spatial skills by reviewing studies of individuals with atypical hormonal profiles and studies of typically developing individuals whose hormonal levels vary.

Atypical Hormone Profiles

One group that exhibits an atypical hormone profile is males with androgen deficiency early in development. This group performs lower on a variety of visuospatial tasks than their typically developing peers.[72] In contrast, males who acquire androgen deficiency after puberty do not appear to exhibit spatial impairments, perhaps reflecting differences between organizational and activational effects of hormones.[75]

Another group that has been widely studied is females with congenital adrenal hyperplasia (CAH), a condition that results in above normal levels of androgens during prenatal and early postnatal development. Although some studies report that females with CAH have enhanced spatial skills compared to unaffected females[73], others report no such advantage.[76] Studies examining play preferences more consistently show differences between females with CAH and unaffected females. Young girls with CAH play more with stereotypical male-type toys (e.g., cars and trucks) and prefer male playmates compared to unaffected girls, even though parents strongly encourage CAH girls to play with stereotypical female toys, such as dolls.[77] Moreover, such preferences persist into adolescence and adulthood.[78] It could be that CAH leads affected girls to engage more in male stereotyped play activities, and that these play activities, rather than the CAH *per se*, predict the sex difference in spatial skill that is sometimes reported. Thus, it is important to examine whether play activities might mediate the relation of CAH to enhanced spatial thinking, which could potentially resolve inconsistencies in the literature. Such a finding would be consistent with widely held view that biological and experiential factors interact to explain gender differences in play and spatial abilities.[79]

Hormonal Variations in Typically Developing Children

Other studies have examined whether variations in prenatal androgen levels in typically developing children are related to variations in mental rotation. For example, Grimshaw et al.[80] found a positive relation between prenatal testosterone level (measured during amniocentesis in the second trimester of pregnancy) and mental rotation at age 7 in girls but not in boys. A recent study[81] assessed the mental rotation skills of young adult twins whose co-twin was either the same or different sex. The findings showed that females with an opposite sex twin performed better on a 3D mental rotation task than females with a same sex twin among dizygotic twins. These findings are consistent with the role of *in utero* testosterone levels on variations in spatial skill. However, they could also be explained by females with an opposite sex twin playing with more male stereotyped toys than females with a same sex twin, because of the greater opportunity to do so. Thus, findings that are sometimes interpreted as supporting a biological difference for the gender difference in mental rotation could also be affected by increased spatial experiences.

Differences in Brain Structure and Function

Brain imaging studies have implicated various brain areas in mental rotation, illustrating the complex network of processes underlying this skill. For example, both men and women show activation in brain areas associated with the motor system, such as the globus pallidus and the premotor cortex, as well as areas involved in mental imagery and self-awareness, such as the precuneus.[82–85] As detailed below, studies of brain structure and function have also revealed sex differences in relevant brain areas—which may be rooted in differences in gonadal hormones,[86–88] as well as in differential experiences, both of which have been implicated in the sex difference in mental rotation.

In terms of sex differences in brain structure, differences in the parietal lobe—a brain region implicated in spatial processing and mental rotation[89–91]—have been identified. Goldstein and colleagues[92] found that volume of the inferior parietal lobe is 20% greater in men than women, even after accounting for overall brain volume. Furthermore, women have smaller parietal lobe surface area, and, importantly, this structural brain difference is related to mental rotation task performance.[93] In terms of brain functioning during mental rotation, differences in activation patterns have been reported for males versus females. Although the specific brain regions implicated differ across studies, likely due to differences in the stimuli used and the nature of the control task, there is some agreement that activation patterns in women reflect more effortful, top–down processing—reflective of verbal-analytic strategies—whereas activation patterns for men reflect more automatic, bottom–up processing—reflecting more visuomotor strategies.[26,28,82] For example, one study examining activation during a mental rotation task of Shepard–Metzler figures found that women showed greater activation than men in the dorsal medial prefrontal cortex and the left temporal–occipital association cortex, thought to be involved in effortful control.[82] In contrast, men showed greater activation than women in areas implicated in automatic, effortless task performance, including the basal ganglia as well as in Broadmann area 7, implicated in memory-related visual imagery.[94]

Only a few studies have examined whether there are sex differences in the brain areas involved in mental rotation in children, and whether these areas differ from those found to be important in adults. One developmental fMRI study compared brain activation patterns in third graders, sixth graders, and adults.[95] Although similar regions were involved across ages, there was a shift from stronger right parietal activation in children to more bilateral parietal activation in adults, perhaps related to greater left hemisphere involvement in more practiced cognitive tasks and greater right hemisphere involvement in more novel tasks, with the mental rotation task being more novel for children than adults.[96] Of note, the investigators used a 2D mental rotation task involving mental rotation of animals rather than the kinds of 3D mental rotation tasks that are typically used in the adult studies, and these task differences need to be taken into account in the interpretation of developmental differences. There was no significant difference in activation patterns between boys and girls, although a difference was detected between adult men and women such that women showed greater activation of the right middle temporal gyrus, the right inferior frontal gyrus, and the left primary motor cortex compared to men.

Other developmental studies have used electrophysiological measures to assess sex differences in brain activation during mental rotation.[97,98] One study showed bilateral parietal activity in preschool boys and left parietal activity in preschool girls on a mental rotation task that involved judging whether two letters were the same or mirror images of each other.[97] Despite the sex difference in patterns of event related potential (ERP) activity, boys and girls showed no difference in their speed or accuracy on the task, which is consistent with previous data showing that there is not typically a sex difference when mentally rotating letters, at least in adults.[23] The investigators speculated that the right hemispheric activation pattern in males may reflect a holistic rotation strategy, whereas the bilateral activation pattern in females might reflect a more analytic, piecemeal approach (but see Ref 29).[26,97,99]

Findings are somewhat inconsistent from study to study, even within the same lab. Some studies report similar patterns across age and gender whereas others find gender and/or age differences. These differing results may be associated with the nature of the mental rotation task, and perhaps to the nature of the baseline task used as a control for the mental rotation task. To sort out the complex set of findings that have been reported, we need more systematic, theory driven research to establish a clear picture of the nature of sex differences in brain activation on mental rotation tasks and whether these vary over the course of development. As is the case for behavioral studies, longitudinal studies examining a variety of different kinds of mental rotation tasks, and that consider not only the child's gender but also their spatial experiences, would help create a more coherent picture of developmental consistency and change. As it stands, a variety of different sex-related patterns of activation are interpreted as supporting a more holistic pattern of mental rotation in males and a more piecemeal pattern of mental rotation in females, which may not adequately capture processing differences in view of recent findings.[29]

Environmental Factors

Play Experiences

Engaging in masculine-stereotyped activities is associated with higher performance on spatial tasks, including mental rotation.[100,101] These include activities such as playing with blocks and other construction toys as well as playing certain videogames.[17,102] Studies of children in their home and school environments show that boys are more likely than girls to play with construction toys (e.g., block building and Legos), a type of play that involves spatial manipulations and transformations as well as part-whole thinking.[103–105] Furthermore, this kind of play is related to better performance on spatial visualization tasks, such as the recognition of geometric figures embedded within more complex pictures.[106] What is less clear is the extent to which this differential spatial play by boys and girls is driven by intrinsic interests, by parent/teacher encouragement and/or by societal stereotypes. Furthermore, it is unclear whether encouraging girls to play with blocks, for example, would result in improvement in their mental rotation skills.

Many reports of sex differences in toy preference are based on parental questionnaires rather than direct observations. However, a recent longitudinal study videotaped the naturalistic interactions of children and their primary caregivers, at 4-month intervals, beginning when children were 14 months of age. Pruden et al.[107] showed that children who heard more spatial language from their parents during nine observation sessions between 14 and 46 months of age performed better on spatial tasks, including a mental transformation task at 54 months, controlling for the overall talk children heard. A mediation analysis showed that parent spatial language predicted children's spatial language, and that this in turn predicted performance on the mental rotation task.[107] Although these findings are correlational, they raise the possibility that that spatial language helps children on mental transformation tasks by giving them ways to describe the stimuli, thereby decreasing the cognitive load of mentally transforming the stimuli. Rich spatial language knowledge also may enhance spatial skill by habitually increasing attention to spatial information in the world. In another investigation based on this longitudinal study, Levine et al.[108] coded children's puzzle play during six observation sessions between 26 and 46 months of age, and found that children who played with puzzles at least once during these sessions scored higher on a mental transformation test that was given at 54 months of age[52] than children who never played with puzzles. Moreover, the frequency of observed puzzle play was correlated with scores on the mental transformation test, supporting a dose–response relationship.[108] Although boys and girls were as likely to engage in puzzle play during the observation sessions, boys played with more difficult puzzles than girls. Furthermore, the parents of boys were more engaged in the puzzle play than the parents of girls and provided them with more spatial language, perhaps because of the greater difficulty of the puzzles the boys were playing with and the greater need to scaffold this play. However, a follow-up study in the laboratory presented parent–child dyads with the exact same puzzles, one a 24 piece puzzle and the other a 48 piece puzzle. Analyses showed that parents provided children with more spatial language on the more difficult puzzle, but that parents of boys provided more spatial language than parents of girls for both the easy and more difficult puzzle.[109] Although the explanation for the greater spatial language input to boys remains an open question, it is possible that this input is helpful to the development of boys' spatial thinking, for the reasons noted above—it may reduce cognitive load while carrying out spatial tasks and may guide children's attention to spatial information on a regular basis.[107]

Studies of college students also reveal a relation between engaging in masculine-stereotyped spatial activities such as carpentry, model-building (based on retrospective reports), and spatial skills (based on performance on tasks such as the MRT).[101] Sex differences in videogame experience have specifically been associated with a sex difference in mental rotation. In a large sample of undergraduates (N = 1278), males reported significantly more experience with videogames than females. Among a subsample of participants (n = 180), males performed significantly higher on the MRT than females, and this sex difference was partially mediated by videogame experience. Moreover, females with greater videogame experience showed higher MRT scores than females with less videogame experience.[102]

The correlations between spatial experiences and spatial abilities do not show that these experiences *cause* higher levels of spatial skill. Instead, it could be that individuals with higher spatial skill, perhaps related to biological factors such as prenatal hormone exposure, engage more in spatially related activities. However, it is also possible that factors such as parental encouragement and societal stereotypes affect the amount and quality of spatial play, spatial interests, and spatial thinking. If this is the case, we might expect sex differences in spatial thinking to vary depending on cultural context.

Gender Stereotypes

There is also evidence that sex differences in children's toy preferences are influenced by societal gender stereotypes. For example, caregivers in Western cultures have been shown to encourage boys more than girls to participate in spatial activities.[100] Additionally, the societal stereotype that females are worse than males at spatial thinking has been found to affect females' performance on spatial tasks, at least in adults. This may be because of stereotype threat—the threat that females, as members of the stereotyped group, may feel when confronted with a spatial task, which in turn can lead to anxiety about their ability to perform the task, deplete their working memory resources, and compromise their task performance.[110] For example, McGlone and Aronson[111] found that female undergraduates performed worse on a MRT when their gender was primed than when their identity as a student at a selective private college was primed prior to taking the test. The reverse was true of male college students, where priming gender led to stereotype lift. Other studies have explicitly manipulated beliefs about sex differences in spatial performance. For example, Campbell and Collaer[112] found a male advantage on a line orientation discrimination task in the explicit stereotype (i.e., 'men do better') and control (no information) conditions, but no sex difference in the stereotype-nullified condition (i.e., 'women and men perform equally'). Importantly, females performed significantly better in the stereotype-nullified condition than in either of the other conditions, particularly on challenging items, whereas male performance did not differ across conditions (see Figure 8). In addition, a study of college students showed that positive same-sex role models had a positive impact on STEM outcomes for female students but not for male students (but see Ref 113). Female students with female math and science professors were more likely to take STEM courses and to graduate with a STEM major than female students with male professors.[114] It is possible that female college professors in the STEM disciplines serve to nullify the stereotype that males are better than females at mathematical and spatial tasks, which results in more females pursuing and succeeding in the STEM disciplines.

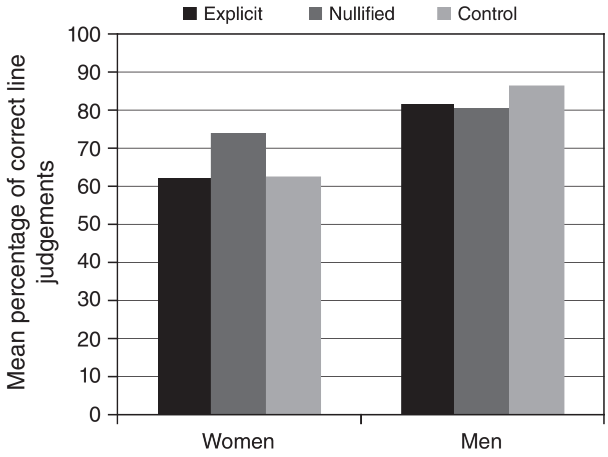

FIGURE 8 | Performance of women and men in explicit stereotype, stereotype nullified and control conditions on a line orientation perception test.[112]

Because research on the effects of spatial stereotypes has focused on adults, we do not have information about when children first become aware of this stereotype and when the stereotype starts to affect their spatial skills and interests. This is an important question, as it can inform efforts to stem deleterious effects of these stereotypes.

Spatial Anxiety

There is little research on the relation of spatial anxiety to spatial learning and performance. In a study examining whether there is a link between spatial anxiety and the sex difference in mental rotation skill, some first and second grade students reported anxiety about engaging in spatial activities, as assessed by a newly developed questionnaire, the Child Spatial Anxiety Scale (CSAQ).[115] Items from this questionnaire include questions such as: (1) 'How do you feel when your teacher asks you whether these shapes are rectangles and why?' (2) 'How would you feel if your teacher asked you to build this house out of these blocks in 5 min?' Results showed that young children with higher levels of spatial anxiety performed more poorly on the Thurstone Spatial Relations Test,[25] a test involving mental rotation, than those with lower levels of spatial anxiety. Moreover, consistent with findings on math anxiety, the negative relation of spatial anxiety and spatial skill was only found for students who were higher in working memory.[115,116,] However, unlike findings for math anxiety, where the anxiety by working memory interaction holds for both males and females,[117] the spatial anxiety by working memory interaction was only significant for girls (see Figure 9). This may be because spatial anxiety leads to verbal ruminations that may differentially interfere with the more verbal analytic strategies that girls may employ on mental rotation tasks.[118]

Spatial anxiety may have negative implications beyond impeding performance on mental rotation tasks. In a series of studies conducted on adults, Ferguson and colleagues[119] found that individuals reporting high levels of spatial anxiety also tended to report high levels of mathematics anxiety, which has been widely reported to be associated with lower levels of math performance.[117,120,121] Critically, the relation between spatial anxiety and math anxiety

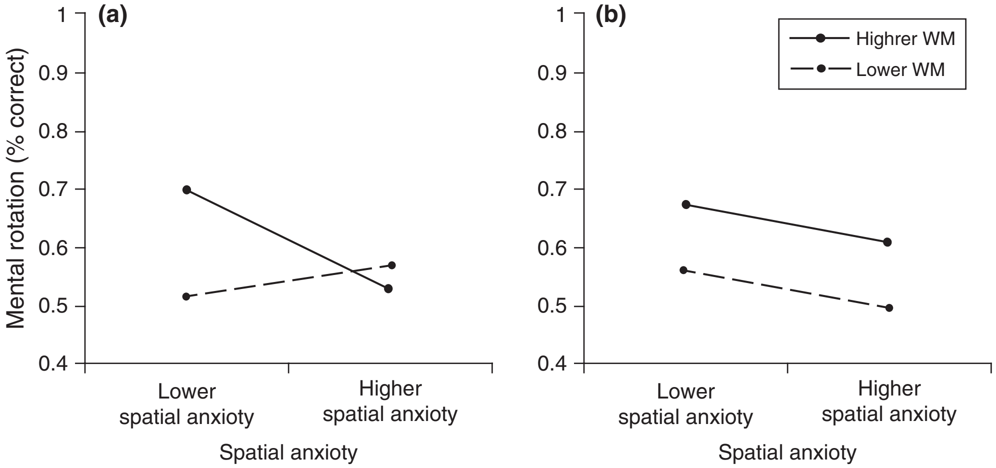

FIGURE 9 | Relation between spatial anxiety and mental rotation task performance as a function of working memory capacity for girls (a) and boys (b).[115]

held even after accounting for sex, general anxiety, and both small- and large-scale spatial skill (sense of direction and mental rotation, respectively). Results also suggested that mental rotation partially accounts for the relation between math and spatial anxiety whereas navigation does not. It could be that spatial anxiety impairs one's mental rotation ability, which might be important to mathematical thinking (e.g., see Refs 15, 122, and 123), thus resulting in greater mathematics anxiety. Although Ferguson et al. did not investigate whether spatial anxiety might account for females' higher levels of math anxiety, they did find that females reported significantly higher levels of spatial anxiety than males in three different samples of participants.

Another important approach to examining the detrimental effects of spatial anxiety on spatial ability involves examining inter-generational effects that can result when spatially anxious adults interact with children. These adults may not provide children with the kinds of input that support the development of spatial thinking. Supporting this possibility, in a study of first and second graders, teachers' level of anxiety about spatial tasks negatively predicted children's growth on the Thurstone Spatial Relations Test[25] across the school year. It is possible that teachers who are high in spatial anxiety are less likely to engage children in their class in spatial activities or that they engage them in lower quality spatial activities, which impedes children's spatial learning.[122]

Role of Cultural Context

Consistent with the role of experience in sex differences in spatial skill, it appears that the sex difference is modulated by a variety of contextual factors that vary both within and across cultures.

Within-Culture Variations

Do sex differences in spatial skill vary as a function of socioeconomic status (SES)? Levine and colleagues[124] reported that sex differences on two different spatial tasks, the Thurstone Spatial Relations Test[125] and a task that involves mapping locations on aerial photographs to locations on maps, varied as a function of SES. In high- and middle-SES groups, boys performed better than girls on both of these spatial tasks. In contrast, there was not a sex difference in a lower SES group (see Figure 10). Moreover, it does not appear that the absence of a sex difference in the lower SES group is attributable to a floor effect because unlike higher SES second graders who show a sex difference on the spatial tasks administered, lower SES third-grade children do not, even though they performed as well as the higher SES second graders on the spatial tasks.

How might such a sex by SES interaction be explained? One possibility is that the male advantage on spatial tasks emerges earlier in high- and middle-SES groups, because in these groups, boys engage in more spatially relevant activities than girls (e.g., playing with Legos, exploring their neighborhood). In contrast, children from lower-SES groups may have less access than their higher-SES counterparts to the kinds of activities that promote spatial thinking, and this may be true for both boys and girls. It is also possible that the experiences of children from lower SES backgrounds with spatial activities such as block play may be qualitatively different than those of their higher SES peers. Thus, we may not see an early sex

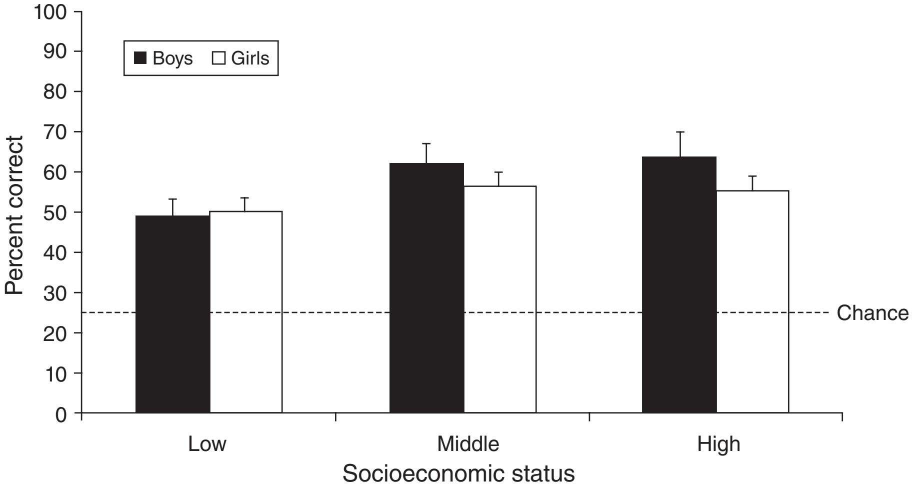

FIGURE 10 | Performance levels of low, middle, and high SES second graders on the Thurstone Spatial Relations Test.[124]

difference on spatial tasks in this group because neither low SES boys nor girls are likely to engage in the kinds of early activities that promote spatial thinking. Over time, lower SES boys may accrue more spatially relevant experience than lower SES girls, and children in this demographic group may also show a male advantage on spatial tasks. Clearly, a longitudinal study is needed to determine the time course over which male advantage on tasks such as mental rotation unfolds in children from different SES groups. Preliminary data from such a study indicate that when low SES girls and boys who do not differ on pretest mental rotation skill are given training on a computer game similar to *Tetris*, a sex difference in favor of males emerges over time (Lourenco and Levine, in preparation).

In apparent contradiction to the Levine et al.[124] findings, Wai et al.[11] analyzed data from Project TALENT—a study conducted with a large representative sample of 9th- to 12th-grade U.S. students (approximately 400,000 in all)—and did not find any evidence of a sex by SES interaction. However, the lack of an interaction with SES in this study could be attributed to a number of study differences. With respect to the spatial skills assessed, children in the Levine et al. study were administered a spatial relations task and an aerial maps task. In contrast, spatial skills in Project TALENT were assessed via a spatial composite score that included mechanical and abstract reasoning. In addition, the Levine et al. study was conducted with second and third graders, whereas the Project TALENT data are from high-school students. Thus, the different findings yielded by these two studies may reflect age-related changes in the interaction between sex and SES–as discussed, males from lower-SES backgrounds may eventually have more spatially relevant experiences than females but these differences may be less marked or even negligible at earlier ages.

Across-Culture Variations

Cross-cultural comparisons have revealed that the gender difference in spatial thinking is robust, and also that certain factors may lead to variations in its size. In adults, a male advantage on spatial tasks has been found across different races, ethnicities, and cultures, such as Japan,[126] Great Britain,[127] and West Africa,[128] as well as India, South Africa, and Australia.[129] An analysis of a large internet survey (including 250,000 respondents) by the British Broadcasting Corporation found that men outperformed women on a 3D mental rotation task and a line orientation judgment task that involved selecting the line from among 15 choices that matched the orientation of a target line. This was true in more than 40 different countries and 7 ethnic groups.[127,130] Lippa and colleagues[127] found that the sex difference in spatial skill increased as gender equity and economic development of countries increased—a finding that is reminiscent of the Sex × SES interaction reported by Levine et al., which found a larger sex difference in more privileged socioeconomic groups.

However, the Lippa et al. finding conflicts with the Gender Stratification Hypothesis,[131–133] which posits that sex differences will be smaller when opportunities for men and women are more equitable. In an attempt to explain their surprising finding, Lippa et al. suggest that women in more egalitarian societies may have more exposure to gender stereotypes about spatial skills, which may negatively impact their performance, as described above. Of note, the larger sex difference in spatial skill in countries with more gender equity contrasts with a study reporting that the sex difference in math performance on the 2003 PISA is smaller in countries with greater gender equity,[132,134] (but see Stoet and Geary[135] for analyses showing inconsistent findings and outlier effects on the math findings, leading them to argue that there is not a relation between math ability and gender equity of countries). One possible reason for the opposite relation of societal gender equity to sex differences in spatial versus math skill is that mathematics is formally taught in school whereas spatial thinking is not. Thus, in societies with greater gender equity, both males and females have more similar access to mathematics courses, and as a result, sex differences in math may diminish. In contrast, because schools do not typically focus on spatial learning, stereotypes and sex-related differences in interests may continue to prevail in countries with more gender equity. Furthermore, as argued by Lippa et al.,[127] the toll that stereotypes take on females' spatial skill may be greater in more gender equitable societies because awareness of gender stereotypes about spatial thinking may be more evident in these countries.

Other studies find that cultural practices that are highly spatially demanding are related to the sex differences in spatial skill and level of spatial skill. For example, in societies where both sexes rely on a high level of visuo-spatial skills for survival, such as the Canadian Eskimo Inuit, men and women do not differ in their performance on spatial tasks.[128] Furthermore, experience with character-based languages (e.g., Mandarin), which incorporate a Euclidean frame for students learning to write, is related to higher mental rotation skill. Men and women who can write in Mandarin scored significantly higher on the MRT compared to men and women who could not write in Chinese, controlling for being multilingual.[136–138]

An innovative study compared sex differences in spatial skill in two distinct, but genetically related tribes in Northeast India that differ in that one is patrilineal and the other matrilineal.[139] One of these tribes, the Karbi, is patrilineal (e.g., it is socially unacceptable for women to own land, and land is inherited by the eldest son), showed a male advantage in time to complete a puzzle assembly task. The other tribe, the Khasi, which is matrilineal (e.g., it is socially unacceptable for men to own land, and land is inherited by the youngest daughter), showed no sex difference on the puzzle assembly task. A remaining question is the mechanism that explains these different male–female differences in puzzle assembly, and whether this pattern of male and female advantage is specific to spatial tasks or is true of performance in other domains as well.

An Integrative Approach

Contemporary researchers are no longer arguing about whether biological or experiential factors impact differences in spatial skill. Rather, they acknowledge that multiple factors affect levels of spatial skill, as is the case for other cognitive functions. A question that needs to be answered is how these factors together contribute to the sex difference in spatial thinking as well as to the wide range of differences in spatial thinking that exist within gender. One such study examined genetic and environmental contributions to spatial ability, measured by the Jigsaws and Hidden Shapes tests,[140] in a sample of over 4000 pairs of identical and fraternal 12-year-old twins.[141] Findings showed that spatial ability was only moderately heritable, with 27% of the variance in spatial scores being explained by genetics and 73% being explained by environmental factors. Although sex differences, favoring boys, were observed in spatial scores, results suggested that the underlying genetic and environmental contributions were the same for both girls and boys.

There also have been attempts to study how genes and environment interact to influence spatial cognition. An early attempt to do this led to the 'bent-twig' hypothesis.[118] According to this hypothesis, an early biologically based spatial advantage for males would lead to a 'bending of the twig' such that males would be more interested in and engage more in spatial activities, which in turn could lead to a greater spatial advantage. An alternative version of this hypothesis is that societal stereotypes rather than, or in addition to, biological factors could bend the twig toward greater male engagement in spatial skills, or more likely, both kinds of influence are important in bending the twig. Working within the 'bent twig' framework, Pezaris and Casey[118] used family handedness—in particular, the presence of a left-handed first-degree family member—as a marker of biological advantage among females,[142] and examined the strategies that right-handed adolescent males and right-handed adolescent females with or without left-handed family members used on a mental rotation task. They found that only females who were considered biologically advantaged (i.e., right-handed with left-handed relatives) *and* who had high amounts of relevant experience (as indexed by high achievement in math and science) performed the mental rotation task using the visuo-spatial strategies typically favored by the males.[83]

Currently, neuroconstructivist theories[143,144,] hold promise for providing a framework that integrates the multiple factors that affect brain and behavioral development. In addition to the many factors that can affect phenotypes, neuroconstructivist theories highlight the fact that the timing of various events, both biological and experiential, can have important influences on developmental outcomes. These theories are beginning to be applied to research on cognitive development,[145,146] although there are still relatively few empirical investigations of spatial development that have taken such an approach.

Some recent work investigating mathematical development through a neuroconstructivist lens (e.g., Ref 147) serves as a useful example of how a more integrative approach could inform our understanding of spatial development and sex differences in spatial thinking. Hart and colleagues used a latent class analysis in a large twin study to explore familial transmission—representing shared genes and environments—of different profiles of cognitive and affective correlates of math achievement in 12-year olds. Students who had the highest math achievement tended to have high numerical skills and low math anxiety, and showed strong evidence of familial transmission. Conversely, students who showed lower math achievement and higher math anxiety, despite having high numerical skills showed little evidence of familial transmission. Children in this group were more likely to have acquired math anxiety through experiences that were not shared with their co-twins (perhaps from teachers[148])—hence the evidence of low familial transmission—and the deleterious effects of math anxiety may have outweighed any benefit of having strong numerical skills. Although this study does not provide evidence of causal relations, it illustrates that multiple, interacting factors contribute to mathematical development. In particular, cognitive and affective factors relate to math achievement differentially depending on child-specific (i.e., nonfamilial) experiences, which may alter the impact of familial experiences and biological factors. Similar studies exploring levels of familial transmission of spatial skills could help us better understand the contribution of various biological, experiential and attitudinal factors to spatial thinking and sex differences in spatial thinking. For example, is spatial anxiety, like math anxiety, acquired at least in part through nonshared experiences? Additionally, longitudinal twin studies would have the potential to capture the intricacies of biological and environmental influences on sex differences by examining changes in familial transmission over developmental time.

Evolutionary Theories

Several investigators have argued that the male advantage in spatial cognition is an evolutionary consequence of sexual selection, a theory that integrates experiential and biological factors over a long-time frame. The common core of these arguments is that over evolutionary time, males traveled farther from their home bases and engaged in more spatially demanding survival practices than females, which in turn created greater selective pressure for the evolution of dynamic spatial skills (e.g., navigation) in males than females (e.g., Ref 149). Newcombe[150] summarizes various evolutionary accounts of sex differences in spatial skill, placing them into one of two categories: the 'Man as Hunter, Woman as Gatherer' theories, and the 'Man Who Gets Around' theories. In the former category, the central argument is based on the division of labor in traditional hunter–gatherer societies: women gathered edible plants, whereas men hunted animals, which was contingent on spatial skills including making tools and tracking and aiming at moving animals. The latter category of theories centers on the idea that men who mated with multiple females had a reproductive advantage and had to rely on spatial–navigational ability in order to find more females than their competitors. As Newcombe points out, both sets of theories have only limited empirical support (but see Ref 151 for related findings on voles). Moreover, existing evolutionary theories are not directly relevant to the sex difference in mental rotation, which is only modestly related to navigation ability,[152–154] and may even be negatively related to a Euclidean navigation strategy[155] (see Ref 156 for a contrasting view).

Studies of nonhuman animals also provide a lens onto epigenetic mechanisms that may contribute to sex differences in spatial skills. A comprehensive review of the relevant animal literature is beyond the scope of this paper, but we highlight one illustrative study (also see Ref 157 for review). Deer moles expand their home range in order to seek out potential mates when they reach sexual maturity, which also co-occurs with improved spatial navigation ability. However, if the moles were exposed to endocrine-disrupting compounds, such as bisphenol A (BPA), earlier in their development—often transmitted through feeding due to contamination of the mother's diet—they do not show these enhanced spatial navigational abilities at the time of sexual maturity. Conversely, female deer moles exposed to endocrine-disrupting compounds showed enhanced spatial navigational abilities.[158] This study illustrates how biological and environmental factors can interact to affect spatial abilities.

The research that has taken an integrative approach to studying development underscores the fact that there are multiple interacting factors at work, and biological precursors for developing strong or weak spatial skills are not always realized in terms of phenotypic spatial ability levels, as intervening environmental factors can alter the course of this development, and even modulate preexisting biological factors. Thus, it is clear that we need more research at multiple levels of analysis, ranging from the molecular to behavioral to societal, to gain a deeper understanding of how the development of spatial skills unfold in males and females.

MALLEABILITY

As mentioned above in discussing biological and experiential factors, real-world experiences such as playing with blocks, puzzles, and videogames are related to individual differences in spatial skill.[17,102] However, these studies leave the question of causation open—i.e., does engaging in these activities build spatial thinking, or do males, because they are on average more interested and better at spatial thinking, differentially engage in these activities, or is there a bidirectional relationship between engaging in spatial activities and spatial ability? Training studies that experimentally manipulate spatial experiences can assess whether particular experiences are causally related to spatial skill development, and whether enriched spatial experiences lead to a maintenance, narrowing, or widening of the male advantage on spatial tasks such as mental rotation. Training studies—as well as meta-analyses of these studies—provide the strongest causal evidence that spatial skill is malleable.

Training Studies

There is ample evidence that a wide range of experiences lead to better performance on spatial tasks such as mental rotation. These experiences often include training that involves the motor system, e.g., through gesturing the movement of objects or through acting on objects, in line with imaging and behavioral studies that implicate the motor system in mental rotation.[82–85] One study found that 5-year-old boys and girls who were instructed to gesture about the movement of target pieces showed greater improvement on a two-dimensional mental transformation task[52] than children who were instructed to produce a pointing gesture or to observe an experimenter produce a movement or pointing gesture (see Figure 11).[159] A follow-up study showed that training that involved actually moving the objects also resulted in improvement, and this improvement was significantly greater than in the gesture training at immediate posttest. However, because children in the gesture condition continued to show improvement between the immediate and delayed posttests whereas those in the object movement condition did not, there was no significant difference between the two conditions 1 week after training (Levine, Goldin-Meadow, Carlson, & Hemani-Lopez, unpublished data). In yet another study, 5-year-old boys and girls improved similarly when given training that involved practice with mental transformation problems or observing the experimenter move the pieces together. However, girls improved more than boys when instructed to imagine the movement of objects on a mental transformation task, possibly because boys but not girls are already doing this.[53]

There is also evidence that computer-game training improves performance on mental rotation tasks, reducing the sex difference in adults.[160] For example, experience with rotating objects (e.g., using a joystick to move objects) has been shown to play an important role in increasing performance on mental rotation tasks, even eliminating the male

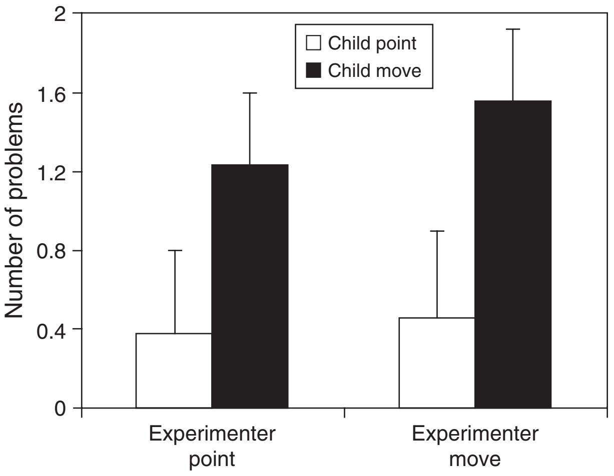

FIGURE 11 | Amount of improvement when children were trained to make either point or movement gestures while the experimenter made either point or move gestures.[159]

advantage.[161] Such training appears to be particularly effective for children with lower levels of spatial skill.[162] Some researchers have argued that improvements in mental rotation elicited by video game training may be due, in part, to improvements in visual selective attention.[163] In one study, college students played an action video game 1 h every day for 10 days and showed improved visual attention following training, compared to control participants who played the video game *Tetris*.[164] Although *Tetris* contains a visuo-motor component, the authors argue that the game only requires focus on one object at a time, which would not be expected to alter visual attention to the same degree as an action video game that involves more dynamic visualization components, such as tracking enemies, detecting new enemies, and avoiding harm. Computer programming that involves designing linear graphics (e.g., *Logo*) has also been shown to improve mental rotation skill. Fifth and sixth grade students who participated in this type of programming in 30-min sessions every other day for one school year showed significant gains in performance on the Thurstone, regardless of sex and grade level.[165]

Practice alone also benefits mental rotation skill.[18,160,166] Furthermore, as with other kinds of training, practice effects have been shown to transfer to different spatial tasks but less to verbal tasks, suggesting that improvements are domain specific.[166] Wright et al. found that long-term practice (daily practice sessions over 21 days) on a computerized version of either the classic Shepard and Metzler mental rotation task (MRT) or a computerized adaptation of the mental paper-folding task (MPFT)[19,167] resulted in improved performance not only on the task used in training but also on the untrained spatial task (either the MRT or the MPFT). Critically, this transfer effect was domain specific, as spatial practice did not improve performance on a computerized verbal analogies test.[168] Of note, the improvement on the unpracticed mental rotation task occurred at the level of encoding, suggesting that generalized spatial training effect may impact the quality of the encoding of spatial stimuli, rather than mental rotation *per se*.[169] In yet another training study, participants were given a pretest, posttest, and delayed posttest on a computerized version of the MRT.[22] Half of the participants were given the pretest, training and posttest in the morning and half in the evening, with the delayed posttest given 12 h later to both groups. While males performed better than females on the pretest, posttest, and delayed posttest, the difference was only significant at pretest. In addition, although both the evening and the morning training groups showed comparable improvement on the immediate posttest, the two groups differed on the delayed posttest—the evening-trained group, showed further improvement following a night of sleep whereas the morning-trained group showed some decrement following a day of normal activities. These findings implicate the role of sleep in consolidating spatial learning on a mental rotation task.[170]

There is also evidence that training spatial thinking may have consequences for STEM success. For example, Sorby developed a spatial workbook for engineering students that emphasized 3D visualization skills.[16] Students who used the workbook not only improved their spatial visualization skills but also were more likely to successfully complete an engineering course. In a related study that employed a regression discontinuity design, Sorby and colleagues[171] assigned engineering students who scored below a cutoff score on a pretest measure of spatial visualization skills to a spatial skills remediation course—involving the same exercises used in the spatial workbook. Students in the intervention course (45% female) showed improved spatial visualization skills following the course, as well as better calculus grades a semester later, compared to their peers who had scored above the cutoff score (16% female) on the pretest spatial visualization measure.

A common thread that emerges from training studies is that the effectiveness of the training depends on the duration of the specific training or practice. In general, longer training results in more improvement,[17] but one study reports that males typically increase mental rotation performance at a faster pace before reaching asymptote, whereas females increase performance more gradually over time (see Figure 12).[166] Thus, studies that occur over a short duration may favor males and may even widen the sex difference in task performance, whereas those that extend over several weeks are more likely to maintain or reduce the sex difference that was present at pretest. This difference in timing may explain some of the inconsistencies in the literature about whether training leads to parallel gains in males and females or whether it closes the gap.[18,160,163,172] Because most training studies are short in duration, sex differences may be maintained or even enlarged after training. The question of why the rate of improvement differs for males and females is an important one. This could be attributable to a male biological advantage and/or to their higher levels of prior experience with spatial tasks. What is clear is that the spatial thinking of both males and females improves with training, with most studies reporting parallel gains.[18]

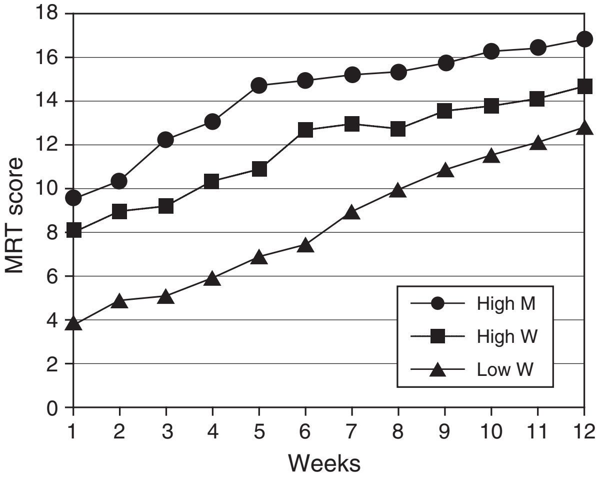

FIGURE 12 | Growth pattern in mental rotation performance over time for high and low spatially experienced women and for high spatially experienced men.[160]

Meta-Analyses of Training Studies

The training studies reviewed above provide evidence of malleability. However, the most compelling evidence of training effects on spatial skill is provided by meta-analyses. Two meta-analyses have focused specifically on the question of the malleability of spatial thinking and whether training impacts males and females to the same extent or differentially. The first of these, Baenninger and Newcombe,[17] found that spatial abilities are malleable, and that training had similar benefits for males and females. Thus, the significant male advantage that was present at pretest, remained at posttest, although both males and females improved significantly from pre- to posttest. This meta-analysis also identified the attributes of training that were most beneficial. In particular, short-term training (e.g., single training sessions, or multiple brief training sessions that occurred over a period of less than three weeks) did not produce effects that differed significantly from practice. In contrast, training interventions that were medium in duration (e.g., more than one training session over more than three weeks or training that is part of curriculum that lasts less than a semester) were more beneficial than either practice or short-term training. Furthermore, training that was specific, focusing on the spatial task assessed at pre- and posttest, benefits performance on that task the most, followed by more general spatial training (using a variety of spatial tasks), followed by indirect training that involves coursework related to spatial ability (e.g., engineering coursework). However, this meta-analysis does not tell us the degree to which these different kinds of spatial training generalize to other spatial tasks or to tasks that are not typically classified as spatial tasks but that may be impacted by spatial skills (e.g., spatial representations of number such as representing numbers as units on a number line).[122,173,] That is, it is possible that general and indirect training generalizes more to untrained spatial tasks than specific training. Furthermore, because there was only one training study involving children at the time Baenninger and Newcombe conducted their meta-analysis,[104] it could not shed light on whether spatial training is more impactful at certain developmental time points, i.e., the question of whether there is any evidence that there is a sensitive period for spatial learning.

In broad agreement with Baenninger and Newcombe's[17] findings, a recent meta-analysis[18] of spatial training studies found that spatial skills are malleable with a weighted effect size relative to available controls of $g = .47$ (Hedges g rather than Cohen's d was used). This study found a sex difference in favor of males at pretest and that training did not eliminate this difference because males and females responded very similarly to training ($g = .54$ for males and $g = .53$ for females). Again, it is important to investigate whether the sex difference is more likely to narrow or close with long-term training given Terlecki and Newcombe's finding that males respond more quickly to training than females. This meta-analysis[18] also addressed the question of whether there is any evidence of a sensitive period for the training of spatial skills. Unfortunately, only a few training studies examined multiple age groups, increasing between-study heterogeneity, and decreasing the opportunity to detect significant differences if they exist. Findings showed a larger effect size of training for children under age 13 than for 13- to 18-year olds or adults, but this difference was not statistically significant. Thus, although this meta-analysis does not provide evidence for a sensitive period for spatial learning, it leaves this possibility open.

EDUCATIONAL IMPLICATIONS

Finding ways to improve females' mental rotation skill is important given accumulating evidence that mental rotation skill matters for educational outcomes and career paths. For example, mental rotation skill has been found to mediate the sex difference in middle school students' science performance[174] and high-school students' math performance.[14] However, it is also important to consider what role mental rotation actually plays in the classroom and in STEM occupations. Is mental rotation directly required for performing various problem-solving tasks involved in STEM disciplines? In some cases, the answer is yes. A salient example is that chemistry students use mental rotation to decide whether two symmetrical chemical structures are the same molecule or stereoisomers, although reliance on mental rotation decreases with expertise.[175] However, as discussed below, it could also be the case that mental rotation ability might be a proxy for spatial thinking more generally, and it is this general capacity for spatial thinking—rather than mental rotation *per se*—that facilitates learning and performance in STEM.[176–178]

Spatial thinking can play a role in even some of the most fundamental aspects of mathematical learning, such as the ability to mentally represent numerical symbols and number words in terms of their relative magnitudes. There is accumulating evidence these representations are built on a highly spatial framework, analogous to a mental number line.[179–181] Moreover, the quality of the spatial organization of children's numerical representations is predictive of their mathematics achievement.[122,123,182] Therefore, it is particularly worrisome that 7- to 9-year-old boys demonstrate more accurate number line representations than their female peers.[183] The sex difference in spatial representations of number is also seen in adults: the association of space and number (SNARC effect) appears to be stronger and the acuity of spatial number representations appears to be finer in males than in females.[173] However, it is possible that the accuracy of females' mental number line can be improved by bolstering their spatial thinking. Indeed, there is evidence that performance on the CMTT[52] at age 5 predicts children's accuracy on a mental number line at age 6, which in turn, predicts their approximate calculation skills at age 8.[122] Additionally, board games that emphasize spatial representations of number magnitudes have been found to elicit more accurate number line representations in young children.[184,185]

Spatial thinking may also serve as a learning or problem-solving tool. For example, one study reports that training children's spatial skills—in particular mental transformation skills—led to improvement on math tasks such as missing term problems such as 2 + ___ = 7,[123] (see Ref 186 for contrasting results). The authors hypothesize that the spatial learning may have primed children's ability to transform the missing term problems into the more familiar, conventional format (e.g., 7 − 2 = ___). Additionally, children who were trained to memorize solutions to complex addition problems with a spatial tool—specifically, a graphic display of addend magnitudes on a number line—demonstrated greater learning than those who were asked to learn the solutions by rote memorization.[182]

Strong spatial thinking is also related to STEM interests, course taking, and to STEM careers, even when controlling for math and verbal skills.[130,187] In their analysis of Project TALENT, which consisted of a large stratified random sample drawn from U.S. high schools in 1960, an important result was that spatial skills of both males and females assessed during high school predicted earning advanced degrees in the STEM disciplines 11 years later, even controlling for verbal and mathematics skills (see Figure 13).[11–13] Apart from the issue of the sex difference in spatial skill, these findings are enough to

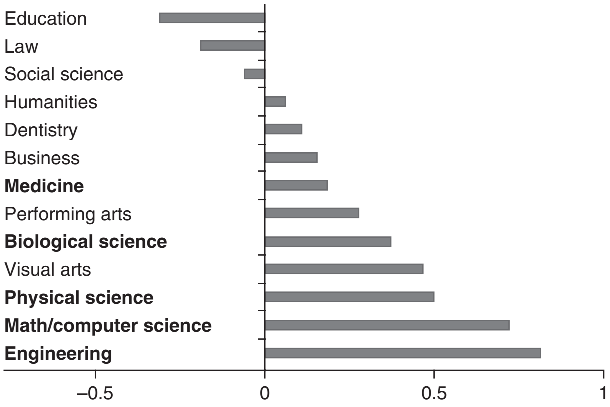

FIGURE 13 | The distribution of mean spatial scores (in standard deviations) by occupations entered 11 years after spatial performance was assessed (adapted from Wai et al.[11]). STEM occupations (bolded) all have positive spatial performance scores.

motivate greater educational efforts to improve spatial thinking.

Despite this evidence, educating spatial thinking has received little attention in education policy and curricula. One possible reason for this blind spot is that, although spatial thinking is malleable, promising research literature on how to actually promote this kind of thinking in the classroom and in home environments is just beginning to emerge.[107,108,188–190] It is also possible that curriculum developers, teachers, and students tacitly endorse essentialist views about spatial ability. For example, anecdotally, people often say that they cannot read maps or follow directions to assemble furniture. In STEM courses, students complain that they have difficulty with 3D visualization and mental rotation, a problem that negatively impacts their success in courses such as organic chemistry and engineering.

Although the existence of sex differences on certain spatial tasks can be viewed as an unfortunate situation, the glass is actually half full. By developing lessons that help children and adults learn to think spatially and incorporating these lessons into the existing curriculum, we are likely to both improve spatial thinking for all learners and to increase STEM achievement and the pipeline of students entering the STEM disciplines.[11,18,177,191] Such instructional efforts need to cover a broad range of spatial skills that are relevant to success in STEM, including mental rotation, but extending to other STEM-related spatial skills such as encoding spatial relationships, cross-sectioning, and scaling. Fluency with using spatial language and symbolic spatial representations such as maps, models, graphs, and diagrams also needs greater attention in the curriculum. In addition, the use of spatial learning tools that highlight spatial information and support spatial thinking (e.g., gesture, spatial language, and spatial alignment) can be used to support STEM learning.[191] In addition to supporting spatial learning, instructional efforts need to focus on reducing spatial anxiety and susceptibility to societal stereotypes about sex differences in spatial abilities. With respect to sex differences, we hypothesize that once schools focus on spatial learning, the greater gender equity of societies is likely to translate into a reduction of the sex difference in spatial skill, as has been the case for mathematics achievement, at least according to some studies.[132,134]

Developing and investigating supports for spatial learning is an increasingly productive direction for research and is an important way to advance the conversation about sex differences in spatial cognition. However, this is not to say that improving spatial thinking through educational efforts is the only route to increasing females' path into the STEM pipeline. Notably, there is a large sex difference (Cohen's *d* ranges from 0.36 to 1.11) in interests and attitudes toward STEM,[192] and there is also evidence that female interest in STEM can be increased by emphasizing communal goals in STEM[193] and by making the physical environment in which STEM learning occurs more appealing.[194] Clearly, there are many complex and interacting factors that influence STEM achievement, interest, and career choice. By taking a multipronged approach that includes 'spatializing' the curriculum as well as nurturing interests and attitudes about STEM learning and careers, we have the best chance of preparing students for the challenges of STEM, and of increasing and diversifying the STEM pipeline, a highly important 21st century goal.

SUMMARY

This review reveals that significant progress has been made in understanding the sex difference in mental rotation skill and the development of this difference. The sex difference in mental rotation skill, in favor of males, is firmly established. Furthermore, this sex difference appears to emerge early in life, at least when measured with certain, implicit measures. We also know that spatial thinking predicts STEM success, both in normative and highly selected talented samples.[11–13] However, there are many remaining questions and new research directions that are needed, which we have flagged as ways to 'advance the conversation.' Among these is a lack of clear understanding of the developmental trajectory of sex differences in spatial thinking partly due to the paucity of longitudinal studies. Additionally, although a variety of biological and environmental explanatory factors of the sex difference have been identified, what is less firmly established is how biological propensities interact with experiences to create different levels of spatial thinking and different attitudes about spatial thinking, as few research studies have attempted an integrative approach. In terms of malleability, we now have strong evidence that spatial thinking improves with training for both males and females, but know little about the effects of long-term spatial training and instruction, and the extent to which the effectiveness of long term training may differ over the course of development, both in terms of durability and generalizability. Relatedly, we lack information about the kinds of curricular changes that would most effectively support high levels of spatial thinking and whether these educational efforts to increase spatial thinking would lead to increased STEM achievement, STEM majors, and STEM career paths, particularly for underrepresented groups such as women. These issues are not only ripe for scientific advances but are of great importance in view of our society's need for a larger STEM workforce.[195]

ACKNOWLEDGMENTS

We thank the National Science Foundation Spatial Intelligence Learning Center Grants SBE-1041707 and SBE-0541957 for support of this work.

REFERENCES

1. Maccoby EE, Jacklin CN. *The Psychology of Sex Differences*, vol. 1. Palo Alto, CA: Stanford University Press; 1974.
2. Summers LH. Remarks at NBER Conference on Diversifying the Science and Engineering Workforce, January, 2005. Available at: http://www.president.harvard.edu/speeches/2005/nber.html. (Accessed February 25, 2010).
3. Ceci SJ, Williams WM. *Why Aren't More Women in Science: Top Researchers Debate the Evidence*. Washington, DC: American Psychological Association; 2007.
4. Spelke ES. Sex differences in intrinsic aptitude for mathematics and science? A critical review. *Am Psychol* 2005, 60:950.
5. Hyde JS. New directions in the study of gender similarities and differences. *Curr Dir Psychol Sci* 2007, 16:259–263.
6. Voyer D, Voyer S, Bryden MP. Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol Bull* 1995, 117:250.
7. Scott DM, Elizabeth H, Maria H. Navigation in a "virtual" maze: sex differences and correlation with psychometric measures of spatial ability in humans. *Evol Hum Behav* 1998, 19:73–87.
8. Liben LS, Golbeck SL. Adults' demonstration of underlying Euclidean concepts in relation to task context. *Dev Psychol* 1986, 22:487.
9. Casey BM, Nuttall RL, Pezaris E. Spatial–mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *J Res Math Educ* 2001, 32:28–57.
10. Geary DC. Sexual selection and sex differences in mathematical abilities. *Behav Brain Sci* 1996, 19:229–247.
11. Wai J, Lubinski D, Benbow CP. Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J Educ Psychol* 2009, 101:817.
12. Shea DL, Lubinski D, Benbow CP. Importance of assessing spatial ability in intellectually talented young adolescents: a 20-year longitudinal study. *J Educ Psychol* 2001, 93:604.
13. Lubinski D. Spatial ability and STEM: a sleeping giant for talent identification and development. *Pers Individ Differ* 2010, 49:344–351.
14. Casey BM, Nuttall R, Pezaris E, Benbow CP. The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Dev Psychol* 1995, 31:697.
15. Casey MB, Nuttall RL, Pezaris E. Mediators of gender differences in mathematics college entrance test scores: a comparison of spatial skills with internalized beliefs and anxieties. *Dev Psychol* 1997, 33:669.
16. Sorby SA. Educational research in developing 3-D spatial skills for engineering students. *Int J Sci Educ* 2009, 31:459–480.
17. Baenninger M, Newcombe N. The role of experience in spatial test performance: a meta-analysis. *Sex Roles* 1989, 20:327–344.
18. Uttal DH, Meadow NG, Tipton E, Hand LL, Alden AR, Warren C, Newcombe NS. The malleability of spatial skills: a meta-analysis of training studies. *Psychol Bull* 2013, 139:352.
19. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. *Science* 1971, 171:701–703.
20. Cooper LA, Shepard RN. Mental transformation in the identification of left and right hands. *J Exp Psychol Hum Percept Perform* 1975, 1:48.
21. Cooperau L, Shepard R. The time required to prepare for a rotated stimulus. *Mem Cognit* 1973, 1:246–250.
22. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. *Percept Mot Skills* 1978, 47:599–604.
23. Linn MC, Petersen AC. Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Dev* 1985, 56:1479–1498.
24. Hyde JS, Fennema E, Lamon SJ. Gender differences in mathematics performance: a meta-analysis. *Psychol Bull* 1990, 107:139–155.

25. Thurstone TG. *PMA Readiness Level*. Chicago, IL: Science Research Associates; 1974.
26. Heil M, Jansen-Osmann P. Sex differences in mental rotation with polygons of different complexity: do men utilize holistic processes whereas women prefer piecemeal ones? *Q J Exp Psychol* 2008, 61:683–689.
27. Geiser C, Lehmann W, Eid M. Separating" rotators" from" nonrotators" in the mental rotations test: a multigroup latent class analysis. *Multivar Behav Res* 2006, 41:261–293.
28. Hugdahl K, Thomsen T, Ersland L. Sex differences in visuo-spatial processing: an fMRI study of mental rotation. *Neuropsychologia* 2006, 44:1575–1583.
29. Xu Y, Franconeri SL. Capacity for visual features in mental rotation. *Psychol Sci* 2015, 26:1241–1251.
30. Farah MJ, Kosslyn SM. Structure and strategy in image generation*. *Cognit Sci* 1981, 5:371–383.
31. Folk MD, Luce RD. Effects of stimulus complexity on mental rotation rate of polygons. *J Exp Psychol Hum Percept Perform* 1987, 13:395.
32. Hooven CK, Chabris CF, Ellison PT, Kosslyn SM. The relationship of male testosterone to components of mental rotation. *Neuropsychologia* 2004, 42:782–790.
33. Boone AP, Hegarty M. Strategy differences do not account for gender differences in mental rotation. In: *Poster Presented at 37th Annual Cognitive Science Society Meeting*, Pasadena, CA, July, 2015.
34. Voyer D, Doyle RA. Item type and gender differences on the Mental Rotations Test. *Learn Individ Differ* 2010, 20:469–472.
35. Lauer JE, Udelson HB, Jeon SO, Lourenco SF. An early sex difference in the relation between mental rotation and object preference. *Front Psychol* 2015, 6:558.
36. Moore DS, Johnson SP. Mental rotation in human infants a sex difference. *Psychol Sci* 2008, 19:1063–1066.
37. Quinn PC, Liben LS. A sex difference in mental rotation in young infants. *Psychol Sci* 2008, 19:1067–1070.
38. Moore DS, Johnson SP. Mental rotation of dynamic, three-dimensional stimuli by 3-month-old infants. *Infancy* 2011, 16:435–445.
39. Hunter MA, Ames EW. A multifactor model of infant preferences for novel and familiar stimuli. *Adv Infancy Res* 1988, 5:69–95.
40. Colombo J. On the neural mechanisms underlying developmental and individual differences in visual fixation in infancy: two hypotheses. *Dev Rev* 1995, 15:97–135.
41. Roder BJ, Bushnell EW, Sasseville AM. Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy* 2000, 1:491–507.
42. Galea LAM, Kimura D. Sex differences in route-learning. *Pers Individ Differ* 1993, 14:53–65.
43. Silverman I, Eals M. *Sex Differences in Spatial Abilities: Evolutionary Theory and Data*. New York, NY: Oxford University Press; 1992.
44. Quinn PC, Liben LS. A sex difference in mental rotation in infants: convergent evidence. *Infancy* 2014, 19:103–116.
45. Möhring W, Frick A. Touching up mental rotation: effects of manual experience on 6-month-old infants' mental object rotation. *Child Dev* 2013, 84:1554–1565.
46. Frick A, Wang S-h. Mental spatial transformations in 14-and 16-month-old infants: effects of action and observational experience. *Child Dev* 2014, 85:278–293.
47. Hespos SJ, Rochat P. Dynamic mental representation in infancy. *Cognition* 1997, 64:153–188.
48. Rochat P, Hespos SJ. Tracking and anticipation of invisible spatial transformations by 4-to 8-month-old infants. *Cogn Dev* 1996, 11:3–17.
49. Schwarzer G, Freitag C, Buckel R, Lofruthe A. Crawling is associated with mental rotation ability by 9-month-old infants. *Infancy* 2013, 18:432–441.
50. Schwarzer G, Freitag C, Schum N. How crawling and manual object exploration are related to the mental rotation abilities of 9-month-old infants. *Front Psychol* 2013, 4:97.
51. Frick A, Möhring W. Mental object rotation and motor development in 8-and 10-month-old infants. *J Exp Child Psychol* 2013, 115:708–720.
52. Levine SC, Huttenlocher J, Taylor A, Langrock A. Early sex differences in spatial skill. *Dev Psychol* 1999, 35:940.
53. Ehrlich SB, Levine SC, Goldin-Meadow S. The importance of gesture in children's spatial reasoning. *Dev Psychol* 2006, 42:1259.
54. Marmor GS. Development of kinetic images: when does the child first represent movement in mental images? *Cogn Psychol* 1975, 7:548–559.
55. Marmor GS. Mental rotation and number conservation: are they related? *Dev Psychol* 1977, 13:320.
56. Frick A, Hansen MA, Newcombe NS. Development of mental rotation in 3-to 5-year-old children. *Cogn Dev* 2013, 28:386–399.
57. Frick A, Ferrara K, Newcombe NS. Using a touch screen paradigm to assess the development of mental rotation between 3$1/2$ and 5$1/2$ years of age. *Cogn Process* 2013, 14:117–127.
58. Neuburger S, Jansen P, Heil M, Quaiser-Pohl C. Gender differences in pre-adolescents' mental-rotation performance: do they depend on grade and stimulus type? *Pers Individ Differ* 2011, 50:1238–1242.

59. Titze C, Jansen P, Heil M. Mental rotation performance and the effect of gender in fourth graders and adults. *Eur J Dev Psychol* 2010, 7:432–444.

60. Frick A, Daum MM, Walser S, Mast FW. Motor processes in children's mental rotation. *J Cogn Dev* 2009, 10:18–40.

61. Frick A, Möhring W, Newcombe NS. Development of mental transformation abilities. *Trends Cogn Sci* 2014, 18:536–542.

62. Lauer JE, Lourenco SF. *Stability in spatial ability from infancy to preschool age. In:* Poster presented at the Biennial Meeting of the Society for Research in Child Development (SRCD), Philadelphia, PA, March 2015.

63. Alexander GM, Wilcox T, Woods R. Sex differences in infants' visual interest in toys. *Arch Sex Behav* 2009, 38:427–433.

64. Baron-Cohen S. Why so few women in math and science? In: Sommers CH, ed. *The Science of Women in Science*. Washington, DC: AEI Press; 2009.

65. Connellan J, Baron-Cohen S, Wheelwright S, Batki A, Ahluwalia J. Sex differences in human neonatal social perception. *Infant Behav Dev* 2000, 23:113–118.

66. Hassett JM, Siebert ER, Wallen K. Sex differences in rhesus monkey toy preferences parallel those of children. *Horm Behav* 2008, 54:359–364.

67. Alexander GM. An evolutionary perspective of sex-typed toy preferences: pink, blue, and the brain. *Arch Sex Behav* 2003, 32:7–14.

68. Bachevalier J, Hagger C. Sex differences in the development of learning abilities in primates. *Psychoneuroendocrinology* 1991, 16:177–188.

69. Salyer DL, Lund TD, Fleming DE, Lephart ED, Horvath TL. Sexual dimorphism and aromatase in the rat retina. *Dev Brain Res* 2001, 126:131–136.

70. Arnold AP, Chen X. What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? *Front Neuroendocrinol* 2009, 30:1–9.

71. McCarthy MM, Arnold AP, Ball GF, Blaustein JD, De Vries GJ. Sex differences in the brain: the not so inconvenient truth. *J Neurosci* 2012, 32:2241–2247.

72. Buchsbaum MS, Henkin RI. Perceptual abnormalities in patients with chromatin negative gonadal dysgenesis and hypogonadotropic hypogonadism. *Int J Neurosci* 1980, 11:201–209.

73. Resnick SM, Berenbaum SA, Gottesman II, Bouchard TJ. Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Dev Psychol* 1986, 22:191.

74. Williams CL, Barnett AM, Meck WH. Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. *Behav Neurosci* 1990, 104:84.

75. Slabbekoorn D, Van Goozen SHM, Megens J, Gooren LJG, Cohen-Kettenis PT. Activating effects of cross-sex hormones on cognitive functioning: a study of short-term and long-term hormone effects in transsexuals. *Psychoneuroendocrinology* 1999, 24:423–447.

76. Hines M, Fane BA, Pasterski VL, Mathews GA, Conway GS, Brook C. Spatial abilities following prenatal androgen abnormality: targeting and mental rotations performance in individuals with congenital adrenal hyperplasia. *Psychoneuroendocrinology* 2003, 28:1010–1026.

77. Pasterski VL, Geffner ME, Brain C, Hindmarsh P, Brook C, Hines M. Prenatal hormones and postnatal socialization by parents as determinants of male-typical toy play in girls with congenital adrenal hyperplasia. *Child Dev* 2005, 76:264–278.

78. Cohen-Bendahan CCC, van de Beek C, Berenbaum SA. Prenatal sex hormone effects on child and adult sex-typed behavior: methods and findings. *Neurosci Biobehav Rev* 2005, 29:353–384.

79. Sameroff A. A unified theory of development: a dialectic integration of nature and nurture. *Child Dev* 2010, 81:6–22.

80. Grimshaw GM, Sitarenios G, Finegan J-AK. Mental rotation at 7 years-relations with prenatal testosterone levels and spatial play experiences. *Brain Cogn* 1995, 29:85–100.

81. Vuoksimaa E, Kaprio J, Kremen WS, Hokkanen L, Viken RJ, Tuulio-Henriksson A, Rose RJ. Having a male co-twin masculinizes mental rotation performance in females. *Psychol Sci* 2010, 21:1069–1071.

82. Butler T, Imperato-McGinley J, Pan H, Voyer D, Cordero J, Zhu Y-S, Stern E, Silbersweig D. Sex differences in mental rotation: top–down versus bottom–up processing. *Neuroimage* 2006, 32:445–456.

83. Jordan K, Wüstenberg T, Heinze H-J, Peters M, Jäncke L. Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia* 2002, 40:2397–2408.

84. Vingerhoets G, De Lange FP, Vandemaele P, Deblaere K, Achten E. Motor imagery in mental rotation: an fMRI study. *Neuroimage* 2002, 17:1623–1633.

85. Richter W, Somorjai R, Summers R, Jarmasz M, Menon RS, Gati JS, Georgopoulos AP, Tegeler C, Ugurbil K, Kim S-G. Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *J Cogn Neurosci* 2000, 12:310–320.

86. Dietrich T, Krings T, Neulen J, Willmes K, Erberich S, Thron A, Sturm W. Effects of blood estrogen level on cortical activation patterns during cognitive activation as measured by functional MRI. *Neuroimage* 2001, 13:425–432.

87. Hausmann M, Slabbekoorn D, Van Goozen SHM, Cohen-Kettenis PT, Güntürkün O. Sex hormones affect spatial abilities during the menstrual cycle. *Behav Neurosci* 2000, 114:1245.
88. Maki PM, Resnick SM. Effects of estrogen on patterns of brain activity at rest and during cognitive activity: a review of neuroimaging studies. *Neuroimage* 2001, 14:789–801.
89. Save E, Poucet B. Hippocampal–parietal cortical interactions in spatial cognition. *Hippocampus* 2000, 10:491–499.
90. Culham JC, Kanwisher NG. Neuroimaging of cognitive functions in human parietal cortex. *Curr Opin Neurobiol* 2001, 11:157–163.
91. Jagaroo V. Mental rotation and the parietal question in functional neuroimaging: a discussion of two views. *Eur J Cogn Psychol* 2004, 16:717–728.
92. Goldstein JM, Seidman LJ, Horton NJ, Makris N, Kennedy DN, Caviness VS Jr, Faraone SV, Tsuang MT. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb Cortex* 2001, 11:490–497.
93. Koscik T, O'Leary D, Moser DJ, Andreasen NC, Nopoulos P. Sex differences in parietal lobe morphology: relationship to mental rotation performance. *Brain Cogn* 2009, 69:451–459.
94. Fletcher PC, Frith CD, Baker SC, Shallice T, Frackowiak RS, Dolan RJ. The mind's eye—precuneus activation in memory-related imagery. *Neuroimage* 1995, 2:195–200.
95. Kucian K, von Aster M, Loenneker T, Dietrich T, Mast FW, Martin E. Brain activation during mental rotation in school children and adults. *J Neural Transm* 2007, 114:675–686.
96. Goldberg E, Costa LD. Hemisphere differences in the acquisition and use of descriptive systems. *Brain Lang* 1981, 14:144–173.
97. Hahn N, Jansen P, Heil M. Preschoolers' mental rotation: sex differences in hemispheric asymmetry. *J Cogn Neurosci* 2010, 22:1244–1250.
98. Roberts JE, Bell MA. The effects of age and sex on mental rotation performance, verbal performance, and brain electrical activity. *Dev Psychobiol* 2002, 40:391–407.
99. Corballis MC. Mental rotation and the right hemisphere. *Brain Lang* 1997, 57:100–121.
100. Caldera YM, Huston AC, O'Brien M. Social interactions and play patterns of parents and toddlers with feminine, masculine, and neutral toys. *Child Dev* 1989, 60:70–76.
101. Nazareth A, Herrera A, Pruden SM. Explaining sex differences in mental rotation: role of spatial activity experience. *Cogn Process* 2013, 14:201–204.
102. Terlecki MS, Newcombe NS. How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles* 2005, 53:433–441.
103. Tracy DM. Toys, spatial ability, and science and mathematics achievement: are they related? *Sex Roles* 1987, 17:115–138.
104. Connor JM, Serbin LA. Behaviorally based masculine-and feminine-activity-preference scales for preschoolers: correlates with other classroom behaviors and cognitive tests. *Child Dev* 1977, 48:1411–1416.
105. Saracho ON. Preschool children's cognitive style and their selection of academic areas in their play. *Early Child Dev Care* 1995, 112:27–42.
106. Caldera YM, Mc Culp A, O'Brien M, Truglio RT, Alvarez M, Huston AC. Children's play preferences, construction play with blocks, and visual-spatial skills: are they related? *Int J Behav Dev* 1999, 23:855–872.
107. Pruden SM, Levine SC, Huttenlocher J. Children's spatial thinking: does talk about the spatial world matter? *Dev Sci* 2011, 14:1417–1430.
108. Levine SC, Ratliff KR, Huttenlocher J, Cannon J. Early puzzle play: a predictor of preschoolers' spatial transformation skill. *Dev Psychol* 2012, 48:530.
109. Levine SC, Petersen L, Ping R. Parent spatial language inequity during spatial play. In: *Poster Presented at the NSF Spatial Language Conference*, San Diego, CA, January, 2015.
110. Beilock SL. Math performance in stressful situations. *Curr Dir Psychol Sci* 2008, 17:339–343.
111. McGlone MS, Aronson J. Stereotype threat, identity salience, and spatial reasoning. *J Appl Dev Psychol* 2006, 27:486–493.
112. Campbell SM, Collaer ML. Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychol Women Q* 2009, 33:437–444.
113. Betz DE, Sekaquaptewa D. My fair physicist? Feminine math and science role models demotivate young girls. *Soc Psychol Person Sci* 2012, 3:738–746.
114. Carrell SE, Page ME, West JE. Sex and science: how professor gender perpetuates the gender gap, 2009. Available at: http://www.nber.org/papers/w14959. (Accessed February 25, 2010).
115. Ramirez G, Gunderson EA, Levine SC, Beilock SL. Spatial anxiety relates to spatial abilities as a function of working memory in children. *Q J Exp Psychol* 2012, 65:474–487.
116. Beilock SL, Carr TH. When high-powered people fail working memory and "choking under pressure" in math. *Psychol Sci* 2005, 16:101–105.
117. Ramirez G, Gunderson EA, Levine SC, Beilock SL. Math anxiety, working memory, and math achievement in early elementary school. *J Cogn Dev* 2013, 14:187–202.

118. Pezaris E, Casey BM. Girls who use "masculine" problem-solving strategies on a spatial task: proposed genetic and environmental factors. *Brain Cogn* 1991, 17:1–22.

119. Ferguson AM, Maloney EA, Fugelsang J, Risko EF. On the relation between math and spatial ability: the case of math anxiety. *Learn Individ Differ* 2015, 39:1–12.

120. Hembree R. The nature, effects, and relief of mathematics anxiety. *J Res Math Educ* 1990, 21:33–46.

121. Ashcraft MH, Kirk EP. The relationships among working memory, math anxiety, and performance. *J Exp Psychol Gen* 2001, 130:224.

122. Gunderson EA, Ramirez G, Beilock SL, Levine SC. The relation between spatial skill and early number knowledge: the role of the linear number line. *Dev Psychol* 2012, 48:1229.

123. Cheng Y-L, Mix KS. Spatial training improves children's mathematics ability. *J Cogn Dev* 2014, 15:2–11.

124. Levine SC, Vasilyeva M, Lourenco SF, Newcombe NS, Huttenlocher J. Socioeconomic status modifies the sex difference in spatial skill. *Psychol Sci* 2005, 16:841–845.

125. Thurstone LL, Thurstone TG. *Factorial Studies of Intelligence*. Chicago, IL: University of Chicago; 1941.

126. Mann VA, Sasanuma S, Sakuma N, Masaki S. Sex differences in cognitive abilities: a cross-cultural perspective. *Neuropsychologia* 1990, 28:1063–1077.

127. Lippa RA, Collaer ML, Peters M. Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Arch Sex Behav* 2010, 39:990–997.

128. Berry JW. Temne and Eskimo perceptual skills. *Int J Psychol* 1966, 1:207–229.

129. Porteus SD. *Porteus Maze Tests: Fifty Years' Application*. Palo Alto, CA: Pacific Books Publishing; 1965.

130. Silverman I, Choi J, Peters M. The hunter-gatherer theory of sex differences in spatial abilities: data from 40 countries. *Arch Sex Behav* 2007, 36:261–268.

131. Baker DP, Jones DP. Creating gender equality: cross-national gender stratification and mathematical performance. *Sociol Educ* 1993, 66:91–103.

132. Guiso L, Monte F, Sapienza P, Zingales L. Culture, gender, and math. *Science* 2008, 320:1164.

133. Riegle-Crumb C. The cross-national context of the gender gap in math and science. In: Hedges LV, Schneider B. *The Social Organization of Schooling*. New York, NY: The Russell Sage Foundation; 2005, 227–243.

134. Else-Quest NM, Hyde JS, Linn MC. Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol Bull* 2010, 136:103.

135. Stoet G, Geary DC. Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence* 2015, 48:137–151.

136. Li Y, O'Boyle MW. Differences in mental rotation strategies for native speakers of Chinese and English and how they vary as a function of sex and college major. *Psychol Rec* 2011, 61:1.

137. Li C, Nuttall RL, Zhu W. Writing Chinese characters and success on mental rotation test. *Percept Mot Skills* 1999, 88:1261–1270.

138. Li C, Nuttall RL, Zhao S. The effect of writing Chinese characters on success on the water-level task. *J Cross-Cult Psychol* 1999, 30:91–105.

139. Hoffman M, Gneezy U, List JA. Nurture affects gender differences in spatial abilities. *Proc Natl Acad Sci USA* 2011, 108:14786–14788.

140. Smith P, Lord TR. *Spatial Reasoning 6–14 Series*. London: NFER Nelson; 2002.

141. Tosto MG, Hanscombe KB, Haworth CMA, Davis OSP, Petrill SA, Dale PS, Malykh S, Plomin R, Kovas Y. Why do spatial abilities predict mathematical performance? *Dev Sci* 2014, 17:462–470.

142. Annett M. *Left, Right, Hand and Brain: The Right Shift Theory*. London, UK: Lawrence Erlbaum Associates; 1985.

143. Fine C. His brain, her brain? *Science* 2014, 346:915–916.

144. Joel D. Genetic-gonadal-genitals sex (3G-sex) and the misconception of brain and gender, or, why 3G-males and 3G-females have intersex brain and intersex gender. *Biol Sex Differ* 2012, 3:6.

145. Karmiloff-Smith A. The tortuous route from genes to behavior: a neuroconstructivist approach. *Cogn Affect Behav Neurosci* 2006, 6:9–17.

146. Ansari D, Karmiloff-Smith A. Atypical trajectories of number development: a neuroconstructivist perspective. *Trends Cogn Sci* 2002, 6:511–516.

147. Hart SA, Logan JAR, Thompson LA, Kovas Y, McLoughlin G, Petrill SA. A latent profile analysis of math achievement, numerosity, and math anxiety in twins. *J Educ Psychol*. In press.

148. Beilock SL, Gunderson EA, Ramirez G, Levine SC. Female teachers' math anxiety affects girls' math achievement. *Proc Natl Acad Sci USA* 2010, 107:1860–1863.

149. Geary DC. Reflections of evolution and culture in children's cognition: implications for mathematical development and instruction. *Am Psychol* 1995, 50:24.

150. Newcombe NS. On tending to our scientific knitting: thinking about gender in the context of evolution. In: *Handbook of Gender Research in Psychology*. New York, NY: Springer; 2010, 259–274.

151. Gaulin SJC. Evolution of sex difference in spatial ability. *Am J Phys Anthropol* 1992, 35:125–151.

152. Allen GL, Kirasic KC, Dobson SH, Long RG, Beck S. Predicting environmental learning from spatial abilities: an indirect route. *Intelligence* 1996, 22:327–355.
153. Montello DR, Lovelace KL, Golledge RG, Self CM. Sex-related differences and similarities in geographic and environmental spatial abilities. *Ann Assoc Am Geogr* 1999, 89:515–534.
154. Wolbers T, Hegarty M. What determines our navigational abilities? *Trends Cogn Sci* 2010, 14:138–146.
155. Dabbs JM, Chang EL, Strong RA, Milun R. Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evol Hum Behav* 1998, 19:89–98.
156. Vashro L, Cashdan E. Spatial cognition, mobility, and reproductive success in northwestern Namibia. *Evol Hum Behav* 2014, 36:123–129.
157. Jašarević E, Geary DC, Rosenfeld CS. Sexually selected traits: a fundamental framework for studies on behavioral epigenetics. *ILAR J* 2012, 53:253–269.
158. Jašarević E, Sieli PT, Twellman EE, Welsh TH, Schachtman TR, Roberts RM, Geary DC, Rosenfeld CS. Disruption of adult expression of sexually selected traits by developmental exposure to bisphenol A. *Proc Natl Acad Sci USA* 2011, 108:11715–11720.
159. Goldin-Meadow S, Levine SC, Zinchenko E, Yip TK, Hemani N, Factor L. Doing gesture promotes learning a mental transformation task better than seeing gesture. *Dev Sci* 2012, 15:876–884.
160. Terlecki MS, Newcombe NS, Little M. Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Appl Cogn Psychol* 2008, 22:996–1013.
161. Cherney ID. Mom, let me play more computer games: they improve my mental rotation skills. *Sex Roles* 2008, 59:776–786.
162. Subrahmanyam K, Greenfield PM. Effect of video game practice on spatial skills in girls and boys. *J Appl Dev Psychol* 1994, 15:13–32.
163. Feng J, Spence I, Pratt J. Playing an action video game reduces gender differences in spatial cognition. *Psychol Sci* 2007, 18:850–855.
164. Green CS, Bavelier D. Action video game modifies visual selective attention. *Nature* 2003, 423:534–537.
165. Miller RB, Kelly GN, Kelly JT. Effects of Logo computer programming experience on problem solving and spatial relations ability. *Contemp Educ Psychol* 1988, 13:348–357.
166. Wright R, Thompson WL, Ganis G, Newcombe NS, Kosslyn SM. Training generalized spatial skills. *Psychon Bull Rev* 2008, 15:763–771.
167. Shepard RN, Feng C. A chronometric study of mental paper folding. *Cogn Psychol* 1972, 3:228–243.
168. Morrison RG, Krawczyk DC, Holyoak KJ, Hummel JE, Chow TW, Miller BL, Knowlton BJ. A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *J Cogn Neurosci* 2004, 16:260–271.
169. Amorim M-A, Isableu B, Jarraya M. Embodied spatial transformations: "body analogy" for the mental rotation of objects. *J Exp Psychol Gen* 2006, 135:327.
170. Debarnot U, Piolino P, Baron J-C, Guillot A. Mental rotation: effects of gender, training and sleep consolidation. *PLoS ONE* 2013, 8:e60296.
171. Sorby S, Casey BM, Veurink N, Dulaney A. The role of spatial training in improving spatial and calculus performance in engineering students. *Learn Individ Differ* 2013, 26:20–29.
172. Just MA, Carpenter PA. Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychol Rev* 1985, 92:137.
173. Bull R, Cleland AA, Mitchell T. Sex differences in the spatial representation of number. *J Exp Psychol Gen* 2013, 142:181.
174. Ganley CM, Vasilyeva M, Dulaney A. Spatial ability mediates the gender difference in middle school students' science performance. *Child Dev* 2014, 85:1419–1432.
175. Stieff M. Sex differences in the mental rotation of chemistry representations. *J Chem Educ* 2013, 90:165–170.
176. O'Boyle MW, Cunnington R, Silk TJ, Vaughan D, Jackson G, Syngeniotis A, Egan GF. Mathematically gifted male adolescents activate a unique brain network during mental rotation. *Cogn Brain Res* 2005, 25:583–587.
177. Halpern DF, Benbow CP, Geary DC, Gur RC, Hyde JS, Gernsbacher MA. The science of sex differences in science and mathematics. *Psychol Sci Public Interest* 2007, 8:1–51.
178. Casey BM, Andrews N, Schindler H, Kersh JE, Samper A, Copley J. The development of spatial skills through interventions involving block building activities. *Cogn Instr* 2008, 26:269–309.
179. Dehaene S. *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press; 1997.
180. Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S. Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* 1999, 284:970–974.
181. Dehaene S, Bossini S, Giraux P. The mental representation of parity and number magnitude. *J Exp Psychol Gen* 1993, 122:371–396.
182. Booth JL, Siegler RS. Numerical magnitude representations influence arithmetic learning. *Child Dev* 2008, 79:1016–1031.

183. Thompson CA, Opfer JE. Costs and benefits of representational change: effects of context on age and sex differences in symbolic magnitude estimation. *J Exp Child Psychol* 2008, 101:20–51.

184. Laski EV, Siegler RS. Learning from number board games: you learn what you encode. *Dev Psychol* 2014, 50:853.

185. Siegler RS, Ramani GB. Playing linear numerical board games promotes low-income children's numerical development. *Dev Sci* 2008, 11:655–661.

186. Hawes Z, Moss J, Caswell B, Poliszczuk D. Effects of mental rotation training on children's spatial and mathematics performance: a randomized controlled study. *Trends Neurosci Educ* 2015, 4:60–68.

187. Humphreys LG, Lubinski D, Yao G. Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *J Appl Psychol* 1993, 78:250.

188. Casey BM, Dearing E, Dulaney A, Heyman M, Springer R. Young girls' spatial and arithmetic performance: the mediating role of maternal supportive interactions during joint spatial problem solving. *Early Child Res Q* 2014, 29:636–648.

189. Anthamatten P. Spatial thinking concepts in early grade-level geography standards. *J Geogr* 2010, 109:169–180.

190. Young CJ, Cartmill EA, Levine SC, Goldin-Meadow S. Gesture and speech are interlocking pieces: The development of children's jigsaw puzzle assembly ability. In: *36th Annual Meeting of the Cognitive Science Society*, Quebec City, Canada, July 2014.

191. Newcombe NS. Picture this: increasing math and science learning by improving spatial thinking. *Am Educ* 2010, 34:29.

192. Su R, Rounds J, Armstrong PI. Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol Bull* 2009, 135:859.

193. Diekman AB, Brown ER, Johnston AM, Clark EK. Seeking congruity between goals and roles a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol Sci* 2010, 21:1051–1057.

194. Cheryan S, Plaut VC, Davies PG, Steele CM. Ambient belonging: how stereotypical cues impact gender participation in computer science. *J Pers Soc Psychol* 2009, 97:1045.

195. United States Department of Education. Science, technology, engineering and math: education for global leadership. Available at: http://www.ed.gov/stem. (Accessed August 18, 2015).