The Importance of Gesture in Children's Spatial Reasoning

Stacy B. Ehrlich, Susan C. Levine, and Susan Goldin-Meadow
University of Chicago

On average, men outperform women on mental rotation tasks. Even boys as young as 4½ perform better than girls on simplified spatial transformation tasks. The goal of our study was to explore ways of improving 5-year-olds' performance on a spatial transformation task and to examine the strategies children use to solve this task. We found that boys performed better than girls before training and that both boys and girls improved with training, whether they were given explicit instruction or just practice. Regardless of training condition, the more children gestured about moving the pieces when asked to explain how they solved the spatial transformation task, the better they performed on the task, with boys gesturing about movement significantly more (and performing better) than girls. Gesture thus provides useful information about children's spatial strategies, raising the possibility that gesture training may be particularly effective in improving children's mental rotation skills.

Keywords: spatial development, gesture, strategies, mental rotation, sex differences

Children and adults frequently encounter spatial tasks such as navigating through the environment, interpreting maps and diagrams, and following assembly directions for furniture or toys. Spatial skills are also associated with success in math and the sciences, leading to particular careers that require spatial reasoning (Humphreys, Lubinski, & Yao, 1993). For example, doctors use spatial skills to read x-rays, and architects use spatial skills to design new buildings. Despite the relevance of spatial skills, spatial development has received relatively little attention compared with domains such as language development. In particular, we do not yet know which types of input help children develop spatial skills and whether the same input is equally effective for boys and girls.

The first goal of our study was to expose children to several types of inputs and examine their effectiveness in helping children develop mental transformation skills. A second goal was to examine the strategies children use to solve mental transformation

problems, as these strategies might provide hints about the kinds of inputs that could promote spatial skill development. Because evidence has suggested that gestures play an important role in spatial understanding (e.g., Krauss, 1998; Lavergne & Kimura, 1987), we paid particular attention to the gestures children used while explaining their strategies.

Sex Differences in Spatial Skills

Spatial skills have become an important research topic in psychology and education, especially given recent evidence that links spatial skills to achievements in the sciences, math, and engineering (i.e., Humphreys, Lubinski, & Yao, 1993). For example, Casey, Nuttall, Pezaris, and Benbow (1995) found relations between spatial and mathematical skills. In fact, one study has shown that mental rotation abilities are better predictors of performance on the math section of the Scholastic Assessment Test (SAT) than levels of math anxiety and math self-confidence (Casey, Nuttall, & Pezaris, 1997). In addition, Johnson (1984) found that spatial abilities are linked to college students' mathematical problem solving, above and beyond the effects of their math SAT scores. Differences in spatial abilities also seem to have an effect on people's choice of occupation (Benbow, Lubinski, Shea, & Eftekhari-Sanjani, 2000). Underrepresentation of women in the science, technology, engineering, and mathematics fields may therefore be related to their lower levels of spatial skills. Identifying the presence, and onset, of sex differences in spatial abilities thus becomes an important area of investigation.

Existing research has demonstrated that males, on average, perform better than females on many spatial tasks. Two meta-analyses revealed that this male advantage is highly robust on mental rotation tasks (those that involve imagining what a figure will look like in a new orientation; M. C. Linn & Peterson, 1985; Voyer, Voyer, & Bryden, 1995). In 1974, Maccoby and Jacklin stated that the male advantage on spatial transformation tasks first appears around puberty and remains throughout adulthood. More recently, however, several studies have challenged this conclusion,

documenting sex differences well before adolescence. For example, Johnson and Meade (1987) found a consistent and significant sex difference between boys and girls on a battery of spatial tests by age 10. McGuinness and Morley (1991) found that by 4 years of age, children display sex differences in the speed with which they build a three-dimensional building. Most recently, Levine, Huttenlocher, Taylor, and Langrock (1999) found sex differences by age 4½ on a simplified two-dimensional spatial task involving mental transformations.

Environmental Input and Training

Several accounts have attributed sex differences in spatial skills to biological differences between males and females (Geary, 1996; Halpern, 1992; Kimura, 1999. See, in addition, literature on the relation between spatial skills and people with congenital adrenal hyperplasia: Berenbaum, Korman, & Leveroni, 1995; Hampson, Rovet, & Altmann, 1998; Resnick, Berenbaum, Gottesman, & Bouchard, 1986.) These biological differences are likely to interact with environmental factors, such as schooling and child-initiated activities, to exacerbate spatial differences (Baenninger & Newcombe, 1995; Huttenlocher, Levine, & Vevea, 1998; Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). As an example of an environmental factor, Huttenlocher et al. examined improvement in spatial skills in children from kindergarten through first grade. They found that improvement between October and April of each year (the school year) was significantly greater than between April and October (tail end of the school year and summer vacation). Input provided in school thus has an effect on the development of spatial skills. Therefore, even though biology contributes to spatial skills, input also plays a key role.

Despite the fact that existing studies have highlighted the importance of input, we have yet to learn which types of input improve children's spatial learning and whether these inputs affect boys and girls equally. One objective of the present study was therefore to examine whether young boys' and girls' mental transformation skills can be differentially enhanced by various types of training (see Newcombe, Mathason, & Terlecki, 2002, for an overview of the importance of this line of research). Although there has been little research examining this question in children, the effect of input on spatial skills has been examined in adults. Baenninger and Newcombe's (1989) meta-analysis of spatial training effects concluded that spatial skills can be enhanced, but that brief training is no more effective than practice. Most important, adult men and women improve equally after training, resulting in the status quo—a male advantage in spatial skills. We do not know, however, whether these training effects hold for children.

Space and Gesture

One approach to understanding how to enhance spatial skills is to explore how people communicate about space, paying particular attention not only to their words but also to their gestures. Previous research has shown that gesture frequently accompanies talk (Goldin-Meadow, 2003; Kendon, 1980; McNeill, 1992). Spontaneous gestures are so robust that they are used by congenitally blind speakers who have never seen anyone gesture (Iverson & Goldin-Meadow, 1998) and by sighted speakers when they know that a listener cannot see them (Rimé, 1982). Moreover, sponta-

neous gestures often convey information that is not found in speech (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Goldin-Meadow, Alibali, & Church, 1993; Perry, Church, & Goldin-Meadow, 1988).

Even more important, gesture is well suited to capturing spatial information (McNeill, 1992; Kita & Özyürek, 2003) and is frequently produced when talking about space (Lavergne & Kimura, 1987). Gestures occur more often when people are defining spatial words than nonspatial words (Krauss, 1998) and are frequent when people are describing how they navigate through space (Emmorey, Tversky, & Taylor, 2000; Schaal, Uttall, Levine, & Goldin-Meadow, 2005). Moreover, when speakers use spatial words and are prevented from gesturing, their speech is slower and contains more dysfluencies than when they are prevented from gesturing while using nonspatial words (Rauscher, Krauss, & Chen, 1996). Thus, there is growing evidence that gesture and spatial thinking are linked to one another.

Nonetheless, the role of gesture in the development of spatial skills has not been explored. We address this issue in our study by examining the gestures children produce when explaining how they solved mental transformation problems. In addition, we examine whether there are differences in the speech and gesture produced by boys and girls in an attempt to understand early sex differences in spatial skills.

The Present Study

The current research builds on Levine et al.'s (1999) study, which found a male advantage in children's spatial skills. Levine et al. used two-dimensional stimuli that were divided in half by the vertical line of symmetry. The two halves were shown to children either rotated or translated apart. Participants were shown the divided shape and asked which of four whole shapes the two pieces would make if put together (see Figure 1). Results revealed that boys performed significantly better than girls on this task at all ages from 4 years 6 months to 6 years 11 months. In addition, children of all ages performed better on translation items than on rotation items and better on bilaterally symmetrical items than on vertically symmetrical items.

The present study used the same task as Levine et al. (1999) and examined the effects of training on children's performance. Our study contained three conditions: two training conditions and a third practice condition in which children were given extra practice items but no instruction. The two training conditions were designed to give the child experience with transforming the pieces. In the imagine movement condition, children were asked to imagine the pieces moving together. In the observe movement condition, children were asked to watch an experimenter move the pieces together.

As previously mentioned, the adult literature has suggested that brief training is no more effective than practice in improving spatial transformation skills (Baenninger & Newcombe, 1989). Our study asks whether this result holds for children. If so, we would expect boys to maintain their advantage over girls after our interventions. However, other outcomes are possible. For example, in adults mental practice has been found to help those who are not proficient in a skill, but to hinder performance in those who are already proficient (Lutz, Landers & Linder, 2001). If children respond as adults do, girls might be expected to improve more in

Choice Array

(a) Direct translation (b) Diagonal translation (c) Direct rotation (d) Diagonal rotation

Four types of Pieces Cards for above Choice Array
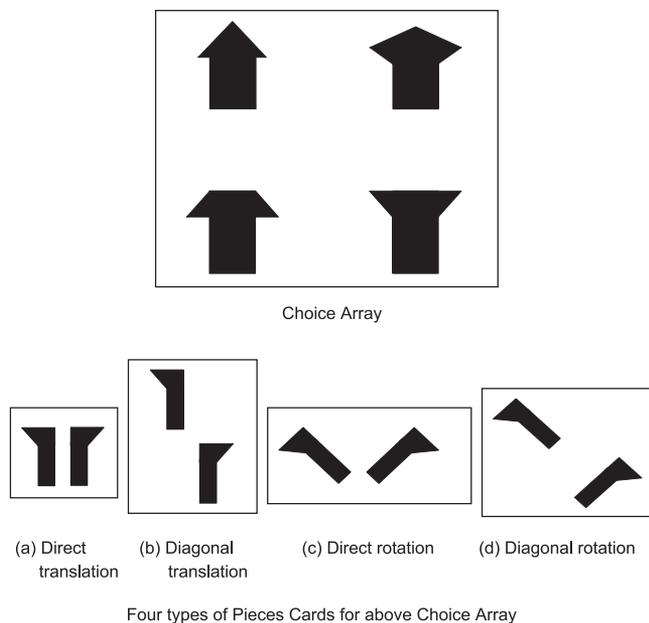
Figure 1. Spatial transformation items: example of the four problem-type configurations (using a shape with vertical symmetry). These items also exemplify the four test forms: A is the target in Form A; B is the target in Form B; C is the target in Form C; and D is the target in Form D. The positions of the pieces were counterbalanced over the four forms.

our training conditions than boys (particularly in the imagine movement condition, which involves mental imagery). Another possibility is that our training conditions will help the rich get richer; that is, it will improve the boys' performance more than the girls'. For example, adult men and women seem to approach mental rotation tasks differently—men tend to mentally rotate the entire object holistically, whereas women tend to use a piecemeal approach in which they rotate the object part by part (Kail, Carter, & Pellegrino, 1979), suggesting that women have more difficulty in visualizing spatial transformations than men. If this pattern holds for children, we might expect boys to benefit more from our training conditions than girls.

In addition to exploring the effects of training, we obtained information about children's strategies by asking them, after the posttest, to explain how they arrived at their answers to the mental transformation problems. If young boys and girls approach mental rotation tasks using different strategies (Kail et al., 1979), we would expect their explanations to differ. We examined this by coding children's explanations for strategies produced in speech and for strategies produced in gesture. As children often express information in their gestures that they do not express in their speech (Goldin-Meadow, Alibali, & Church, 1993; Goldin-Meadow, 2003), they might use their hands to convey an understanding of the spatial transformation task not found in their spoken explanations.

In summary, the current study sought to (a) replicate the findings of Levine et al. (1999); (b) examine whether specific training conditions are more effective in improving performance than practice; and (c) explore the strategies boys and girls express in speech and gesture when describing how they solve these transformation problems.

Method

Participants

Eighty 5-year-old children participated in the study (42 boys and 38 girls, $M = 67.12$ months, $SE = 10.52$ days). For the intervention portion of the study, children were matched (within sex) across the three conditions on the basis of their pretest scores, narrowing our pool of children to 63 for the analyses of training effects (33 boys and 30 girls, $M = 66.80$ months, $SE = 12.27$ days). Preschools in the greater Chicago area were recruited through phone calls. Teachers who agreed to have their class participate were given consent forms to send home to parents. All children included in the study had a returned consent form from a parent and assented to the study before participating. Ethnicity and socioeconomic status data were not collected, but the population from which the sample was drawn was predominantly Caucasian and middle to upper middle class.

Materials and Stimuli

Participants were given 16 pretest items and 16 posttest items (all items differed from one another). Each item consisted of two target pieces on the "pieces card" and a 2×2 "choice array" that included the target shape (that could be formed by the two target pieces) and three foils. Half of the problems contained shapes that were symmetrical along the vertical axis, and the other half contained shapes that were symmetrical along the horizontal axis.[1] Pieces on the pieces card were formed by dividing the target shape in half along the axis of symmetry.

The child's task was to select the whole shape from among four choices in a 2×2 array that could be formed from the halves. The location of the target shape in the choice array was randomly varied across trials, with the constraint that it did not appear in the same position on more than two consecutive trials.

Each participant was shown four different types of problems that varied with respect to the relative positioning of the two pieces. The pieces were (a) translated perpendicular to the line of symmetry (direct translation), (b) translated and then moved diagonally apart (diagonal translation), (c) rotated 45° outward from the line of symmetry (direct rotation), or (d) rotated and then moved diagonally apart (diagonal rotation). See Figure 1 for an example item type.

For both the pretest and the posttest, participants received each type of item (direct translation, diagonal translation, direct rotation, diagonal rotation) four times—twice with vertically symmetrical items and twice with horizontally symmetrical items, resulting in a total of 16 items. The bottom of Figure 1 displays the four positions in which the pieces were arranged. For diagonal translation and diagonal rotation configurations, half of the vertically symmetrical items had the right piece higher, and the other half had the left piece higher. (Similarly, half of the horizontally symmetrical items had the upper piece to the right, and half of them had the upper piece to the left.) The order of the problem types was randomized across trials, with the constraint that the same problem type was not presented twice in a row.

Four different forms of each test were used during each phase of the study (pretest, training, and posttest). Each form was given to roughly one quarter of the participants. The forms varied in the positions of the pieces for a particular target shape but were identical in the order of the choice array cards (see Figure 1 caption). For example, Choice Array Card 1 was the same across the four forms, but in Form A, the pieces were displayed in the direct translation configuration; in Form B, they were displayed in

---

[1] Bornstein and Stiles-Davis (1984) found that young children were more sensitive to vertical symmetry than to horizontal symmetry. We therefore altered the stimuli from the Levine et al. (1999) study to explore whether there was a difference between vertically versus horizontally symmetrical items, but did not find such an effect (see the Results section).

the diagonal translation configuration; in Form C, they were displayed in the direct rotation configuration; and in Form D, they were displayed in the diagonal rotation configuration. Last, the 16 questions that composed the pretest for half of the participants composed the posttest for the other half of the participants, and vice versa.

During the training portion of the study, children were given 12 additional transformation items that were similar to the pretest and posttest stimuli. The only difference was that children were shown actual pieces (made from black-colored wood) placed on white cards rather than pictures of the pieces. The choice arrays were identical to the pretest and posttest phase. Training test forms were similar to pretests and posttests, and each included three direct translations, three diagonal translations, three direct rotations, and three diagonal rotations. Finally, eight probe questions were asked after the completion of the posttest to elicit children's explanations of how they solved the task. During this phase, the pieces cards and choice arrays were presented in the same manner as the pretest and posttest items.

Design and Procedure

All testing was video recorded with consent from children's parents. Children were tested twice, roughly 1 week apart, in a quiet area in their school. During the first testing session, children were given a pretest followed by the Wechsler Preschool and Primary Scales of Intelligence—Revised (WPPSI–R) Vocabulary subtest (Wechsler, 1989) to ensure that any sex differences found on the spatial task could not be attributed to general cognitive differences between the boys and girls in our sample. During the second session, children were randomly assigned to and given one of three experimental intervention conditions (one of the two training conditions or practice), were given a posttest, and were asked a set of probe questions. Children were later matched across the three experimental conditions (within each sex) based on their pretest scores. Two experimenters tested each child: One administered the pretest, posttest, and WPPSI–R Vocabulary subtest, and the other administered the training trials and probe questions to every child.

Pretest and posttest. The pieces card and the choice array were simultaneously placed on a table in front of the child, with the pieces card closest to the child. On the first trial, the experimenter instructed the child to "look

at these pieces" (on the pieces card); "Now, look at these shapes" (on the choice array); "If you put the pieces together they can make one of the shapes. Point to the shape the pieces make." On subsequent trials, the experimenter only said, "Point to the shape the pieces make." No feedback was given on any item.

Experimental conditions. Procedures and instructions for the two training conditions and the practice condition are shown in Table 1.

Probe questions. Each participant was presented with a set of eight transformation items that were randomly selected from items children received during their pretest. (Although the choice card was the same as on the pretest, the pieces card differed; e.g., if the pieces were arrayed in a direct translation configuration on the pretest, they would be arrayed in one of the other three configurations on the probe question.) On each item, the child was prompted to choose an answer and was then asked to explain how she or he arrived at that answer ("Tell me how you got that").

Results

Pretest

Pretest scores on the spatial task were examined to see how children performed before training. There were no performance differences on the basis of form (A, B, C, or D) or the particular test set children received for their pretest (recall that half of the children received one set of 16 questions on the pretest, and the other half of the children received another set of 16 questions on the pretest). Data were therefore collapsed over the two test sets and the four forms for the remainder of the analyses. We performed t tests on the percentage of correct scores on the spatial task for all 80 children tested, with sex as the between-subjects factor. There was a significant effect of sex, $t(78) = 2.665$, $p = .009$, two-tailed, $d = .59$, with boys correctly solving 67.73% ($SE = 2.37\%$) of the items and girls correctly solving 56.76% ($SE = 3.44\%$). We then ran a t test on the WPPSI–R Vocabulary scaled scores and found no effect of sex ($p = .395$, ns), indicating

Table 1
Instructions and Procedure for the Three Intervention Conditions

Intervention condition	Instructions and procedure
Imagine movement	Experimenter to child: "Look at these pieces. If you put the pieces together, they will make a shape. In your mind move the pieces together and then move them back apart. Remember, *don't* move them with your fingers, move them in your mind."
	Experimenter waited 5 s, placed the choice card in front of the child, and said, "Point to the shape the pieces make."
	Subsequent trials: "In your mind, move the pieces together and then move them back apart."
Observe movement	Experimenter to child: "Look at these pieces. If you put the pieces together, they will make a shape. Watch me move the pieces together and then move them back apart. Remember, *don't* move them with your fingers, watch me move them."
	Then experimenter placed the choice card in front of the child and said, "Point to the shape the pieces make."
	Subsequent trials: "Watch me move the pieces together and then move them back apart."
Practice (same as pretest and posttest)	Experimenter to child:(Both the pieces and choice card are out from the start of a trial.) "Look at these pieces, but *don't* move them with your fingers. Now, look at these shapes. If you put the pieces together, they will make a shape. . . . Point to the shape the pieces make."
	Subsequent trials: "Point to the shape the pieces make."

Table 2
Mean Percentage Correct on Pretest and Posttest for Boys and Girls in the Three Conditions

	Imagine movement			Observe movement			Practice		
	Pretest[a]	Posttest	Probe	Pretest[a]	Posttest	Probe	Pretest[a]	Posttest	Probe
Boys									
M	70.47	64.23	73.86	66.51	69.91	78.75	67.06	76.16	78.41
SE	4.05	7.11	7.04	3.88	3.65	3.25	5.60	4.52	5.35
Girls									
M	56.90	71.90	72.50	57.53	64.40	77.50	56.26	65.64	62.50
SE	6.35	5.21	4.86	6.03	7.34	6.12	7.10	5.12	8.94

Note. These scores are not for the entire sample of 80 children, but only for the 63 children who were matched across conditions.
[a] There were no significant differences on pretest scores between conditions (within sex).

that the male advantage on the spatial pretest was not a reflection of a general cognitive advantage of the boys in our sample.

A mixed-model analysis of variance (ANOVA) examined children's performance on different item types. Symmetry (vertical or horizontal), transformation (translation or rotation), and relative location of the pieces (direct or diagonal) were within-subjects factors, with sex as the between-subjects factor. Results showed a main effect of sex, $F(1, 78) = 7.10$, $p = .009$, $d = .59$, with boys performing better than girls (as found in the previously reported t test), and a main effect of transformation, $F(1, 78) = 15.95$, $p < .001$, $\eta_p^2 = .17$, with children performing better on translation items ($M = 5.44$ out of 8 questions, $SE = 0.21$) than rotation items ($M = 4.56$, $SE = 0.19$). There were no effects of symmetry or location, and no interaction effects.

Training: Pretest to Posttest Scores

Sixty-three participants were included in the analysis of pretest to posttest scores after we matched children in the three experimental conditions on the basis of their pretest scores. These participants showed a similar pretest pattern as the full group of 80 children, with a significant effect of sex ($p = .01$, $d = .63$). A mixed-model ANOVA examined score changes from the pretest to the posttest for each individual. Time was the within-subject variable (pretest score, posttest score), with sex (male, female) and experimental condition (observe movement, imagine movement, practice) as between-subjects variables. A significant main effect of time indicated that children attained higher scores on the posttest than on the pretest, $F(1, 57) = 7.68$, $p < .01$, $d = .34$ (pretest: $M = 62.72\%$, $SE = 2.3\%$; posttest: $M = 68.78\%$, $SE = 2.3\%$). There were trends toward a sex effect, $F(1, 57) = 3.08$, $p = .08$, $d = .45$ (boys: $M = 69.06\%$, $SE = 2.73\%$; girls: $M = 62.11\%$, $SE = 2.87\%$) and a Time × Sex interaction, $F(1, 57) = 3.41$, $p = .07$, $d = .46$.

The marginally significant effects of sex on the posttest warranted further exploration as there was a significant sex difference on the pretest. Examining each condition, we found that boys and girls improved at the same rate in the observe movement and practice conditions; boys thus performed better on the posttest than girls in these two conditions (see Table 2). However, the imagine movement condition affected boys and girls differently—that is, there was a significant interaction between time (growth between pretest and posttest) and sex in this condition, $F(1, 19) = 8.69$, $p <$

$.01$, $\eta_p^2 = .31$. Girls improved their performance with training, whereas boys did not. (Boys actually performed worse on the posttest than on the pretest; see Table 2.) This interaction is what caused the overall sex difference to decrease from the pretest to the posttest.

Probe Scores

We next investigated performance on the eight probe problems. Probe scores were marginally significantly better than posttest scores, $F(1, 56) = 3.51$, $p = .07$, with no interactions with sex or condition.[2] An examination of the probe scores (independent of posttest scores) also exhibited no effect of condition and no Condition × Sex interaction. Although not significant, boys performed better than girls (boys: $M = 76.95\%$, $SE = 3.13\%$; girls: $M = 70.83\%$, $SE = 3.99\%$). As performance was not affected by condition, we collapsed the data across the three training conditions when examining the strategies the children expressed on the probe questions.

Strategy Explanations

Four major codes were created to categorize children's strategy explanations during probe questions: movement, perceptual features, perceptual whole, and alignment. Seventy-four percent of children's responses fell into these four categories. The remaining responses were vague, were impossible to see or hear, or fell into a rare "other" strategy category. Table 3 presents definitions and examples of each coding category in speech and gesture. A particular strategy was coded only once per question (i.e., regardless of how many times a child referred to perceptual features within a given response, the child received credit for perceptual features only once for that response). However, a given response to a question could, and often did, contain more than one type of strategy. In these cases, children were given credit for every strategy type expressed during their explanation. A second coder independently coded speech and gesture explanations for 12 of the 63 participants. Agreement between coders was 93% for picking out strategies from the speech stream, 96% for picking out strat-

---

[2] This finding is not surprising given the practice effects associated with this task, as shown in this study and in Levine et al. (1999).

Table 3
Coding Scheme for Strategies

Category	Definition	Speech examples	Gesture examples
Movement	Any indication of movement of the pieces	"Just slide them together and then it looks like that."	Miming moving the pieces together with the hands
Perceptual features	Any indication that the child is focusing on a particular feature of either the piece(s) or the whole	"Because there is a little bend in here and a point thing here."	Pointing to a specific feature (e.g., a corner) or tracing an outline of a specific feature (e.g., a curve)
Perceptual whole	Any indication that the child is seeing the pieces as a whole	"It looks like an arrow!"	(None produced)
Alignment	Placing the piece(s) on top of the corresponding portion of the whole	"This fits on top of this like this."	Placing fingers over one of the pieces and maintaining that shape as the hand is placed on the corresponding portion of the whole shape
Vague	An attempt to express a strategy, but the coder is unable to clearly understand the explanation	"Because I looked at that carefully, and I looked at the differences."	Waving gestures above the pieces that do not seem to indicate any specific strategy
Other	Any strategy other than one listed here	"And here it's like half of it. But so and two halves make a whole."	Using the hand to form a straight line through the middle of the whole shape to represent the line of symmetry
No strategy	Any indication that the child does not know an answer	"I don't know."	Flip hands palm up in the air to indicate uncertainty

egies from the gesture stream, 96% (Cohen's κ = .93, p < .001) for identifying specific strategies in speech, and 85% (Cohen's κ = .78, p < .001) for identifying specific strategies in gesture. For disagreements, coders reviewed the tapes and agreed on a code.

We looked first at the strategies children expressed in their speech. Children referred to movement most often (M = 3.48 questions out of 8), followed by perceptual features (M = 2.30 questions). The remaining two strategies—perceptual whole and alignment—were relatively infrequent in the children's speech (Ms = 0.76 and 0.10, respectively). Children showed the same pattern in their gestures, referring to movement most often (M = 4.06), followed by perceptual features (M = 2.02), with perceptual whole (M = 0.00) and alignment (M = 0.03) occurring not at all or very infrequently.[3]

Next, we examined whether coding gestures provided additional strategy information not found when we coded speech on its own. Children often conveyed the same strategy in speech and in gesture on a given problem. However, at times they expressed one strategy in speech and a different strategy in gesture (e.g., speech referred to perceptual features, whereas gesture indicated movement of the pieces). Figure 2 displays the mean number of problems on which children conveyed each of the four strategies in (a) speech alone (the child either produced no gesture on that problem or produced speech with a different strategy in the accompanying gesture), (b) both speech and gesture (the child produced the same strategy in both gesture and speech), or (c) gesture alone (the child gestured and did not speak at all on the problem or produced gesture with a different strategy in the accompanying speech). Note that children often produced movement or perceptual features in gesture without producing the strategy in speech. Indeed, we found that children expressed the movement strategy on an average of 4.99 problems (out of 8 possible) if we count strategies produced in speech or gesture (i.e., all categories in Figure 2) compared with only 3.48 if we ignore gesture and count only strategies produced in speech (i.e., the speech + gesture and speech only categories in

Figure 2), $t(62)$ = 5.848, p < .001. Similarly, we found that children expressed perceptual features on 2.93 problems if we count strategies produced in speech or gesture, but did so on only 2.30 problems if we count only strategies produced in speech, the lens through which children's knowledge of strategies is traditionally examined, $t(62)$ = 5.27, p < .001.

Relation Between Strategies and Performance on the Spatial Problems

We next conducted analyses to examine the relation between the strategies children expressed on the probe problems and their performance on those particular items. We found that correct scores on the probe problems correlated significantly with three strategy categories: (a) a positive correlation with movement expressed in either speech or gesture (i.e., any reference to movement regardless of modality; r = .272, p = .032); (b) a positive correlation with movement expressed in gesture (with or without speech; r = .299, p = .018); and (c) a negative correlation with unclear strategies expressed in speech (r = −.262, p = .040). To determine which modality was responsible for the positive relation between the movement strategy and performance on the probe problems, we entered three mutually exclusive variables into a regression analysis: (a) movement expressed in speech only (Sp-only), (b) movement expressed in both speech and gesture (Sp + G), and (c) movement expressed in gesture only (G-only). Putting all three factors into the regression did not result in a good fit (p > .05; Sp-only = .037, Sp + G = .137, G-only = .247). More important, removing movement expressed in speech only improved the overall fit of the model, $F(2, 61)$ = 3.247, p = .046,

---

[3] We found no differences in the strategies children produced in speech and gesture based on whether they were solving translation or rotation problems, suggesting that the children were not using different strategies to solve the two types of problems.
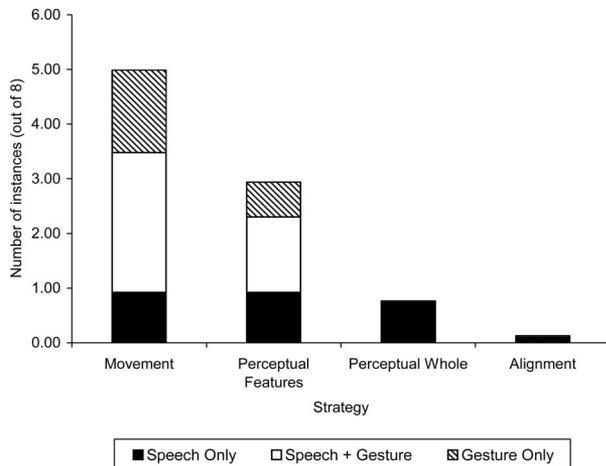
Figure 2. The number of times that each strategy was mentioned on the eight probe questions in speech only, speech + gesture, and gesture only.

$R^2 = .099$, semipartials: Sp + G = .195, G-only = .284. Thus, explanations lacking gesture (i.e., Sp-only) were not correlated with performance on the probe problems ($r = .079$, $p = ns$), whereas explanations containing gesture (i.e., G-only and Sp + G) were ($r = .299$, $p = .018$).

Although this provides useful information about which strategies are linked to better performance overall, we were unsure whether children were gesturing about movement on the specific questions they were answering correctly. Therefore, we ran separate correlations examining the strategies children expressed when answering a question correctly versus when answering a question incorrectly. Talk about movement (ignoring gesture) was correlated with the number of problems children answered correctly ($r = .358$, $p = .004$). However, it was also correlated with the number of problems children answered incorrectly ($r = .416$, $p = .002$). Similarly, talk about movement combined with gesture about movement was correlated with both the number of problems answered correctly ($r = .286$, $p = .024$) and incorrectly ($r = .299$, $p = .030$). Interestingly, talk about perceptual features was correlated only with the number of problems answered incorrectly ($r = .392$, $p = .004$), as was gesture about perceptual features ($r = .499$, $p < .001$). Only one type of response was exclusively related to answering the problems correctly—gesture about movement. Gesturing about moving the pieces was correlated with the number of problems answered correctly ($r = .461$, $p < .001$), but it was not correlated with the number of problems answered incorrectly ($r = .202$, ns). Thus, gesturing about moving the pieces together was uniquely related to correct performance, whereas talking about moving the pieces was not.

The children explained how they solved the problems after actually solving them and therefore may have expressed strategies that they did not use when performing the task (cf. Ericsson & Simon, 1980, but see Goldin-Meadow, Nusbaum, Garber, & Church, 1993, for evidence that gestures produced in post hoc explanations can reflect on-line processes). Although there is no way for us to be certain, if we were to find a correlation between the strategies the children expressed on the probe questions and their performance on the pretest, this would provide support for

our claim that children's explanations reflected strategies they actually used when solving the problems. In fact, we found that several types of explanations were positively related to pretest scores: (a) movement expressed in either modality ($r = .404$, $p = .001$); (b) movement expressed in both speech and gesture at the same time ($r = .259$, $p = .042$); and (c) any gesture about movement, whether or not it was accompanied by speech about movement ($r = .375$, $p = .003$). Note that what all of these types of explanations have in common is movement expressed in gesture. Indeed, we found that speaking about movement without gesturing was not related to pretest performance ($r = .037$, ns). In addition, regression analyses mirror those performed with probe scores. The best fit model relating strategies to performance included gesture-only and speech + gesture, $F(2, 62) = 4.580$, $p = .014$, $R^2 = .132$, semipartials: Sp + G = .304, G-only = .279. In other words, expressing movement in gesture but not in speech was significantly related to children's performance on the pretest, suggesting (although the argument is admittedly post hoc) that the strategies children expressed on the probe questions were likely to reflect how they actually solved the problems during test phases.

Sex Differences in Expressing the Movement Strategy

Because children who performed well on the spatial transformation problems were more likely to refer to movement in gesture, we examined whether boys and girls differed in how often they expressed this type of explanation. We broke down movement strategies into those expressed in speech alone, in both speech and gesture, and in gesture only. Interestingly, there were no differences in how many times boys and girls referred to movement in speech alone ($Ms = 0.82$ and 1.03, respectively) or in both speech and gesture ($Ms = 2.85$ and 2.23, respectively). However, boys and girls differed significantly in the number of times they indicated movement in gesture alone ($Ms = 2.09$ and 0.87, respectively, $p = .015$, $d = .62$; see Figure 3), despite the fact that they did not differ in how often they used other kinds of gestures (means eliminating movement gestures = 3.82 and 3.03, respectively, ns). In addition, Figure 4 displays the proportion of boys and girls who gestured about movement (with or without speaking about movement) on 0–8 questions (total possible = eight times, once on each probe question). Note that 23% of girls did not gesture about movement on any of the eight questions, compared with only 6% of boys. In contrast, 27% of boys gestured on all eight questions, compared with only 3% of girls. Thus, there appears to be a clear difference in how frequently boys and girls gesture about movement, a difference that could either be a reflection of their performance on the spatial task or possibly a driving force behind their performance on the task.

Discussion

We found that boys performed better than girls on our spatial transformation task, thus replicating Levine et al. (1999). Although a number of studies have concluded that sex differences do not emerge before age 10 (Johnson & Meade, 1987; M. C. Linn & Peterson, 1985; Voyer et al., 1995), this study joins other recent research in showing that sex differences are present on at least some spatial tasks as early as 5 years of age (Levine et al., 1999; McGuinness & Morley, 1991). Consistent with previous studies of
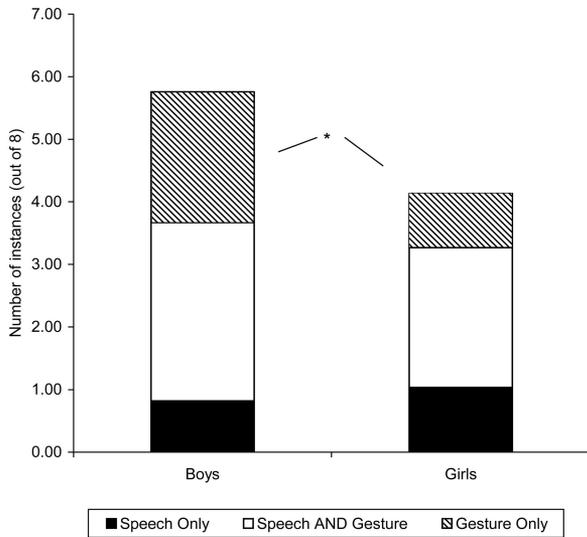
Figure 3. The number of movement strategies produced by boys and girls on the eight probe questions in speech only, speech + gesture, and gesture only. *p < .02, comparing gesture only for boys and girls.

children and adults, there was considerable overlap in the distributions of scores for boys and girls, but a clear male advantage overall (Kimura, 1999; Levine et al., 1999; Peters et al., 1995).

One of our goals was to improve children's performance on the spatial transformation task. Indeed, our interventions did improve performance; however, as a whole our training conditions were no better than practice in raising children's scores. Our findings thus replicate research with adults (Baenninger & Newcombe, 1989) in that the effects of brief training were comparable to practice. Previous researchers have used manipulations to instruct adults that are similar to our training conditions. However, they have also used other techniques, such as computer games (e.g., Tetris) and virtual environments, that need to be tried with children before we can conclude that children are as nonresponsive to short-term explicit instruction on spatial transformation tasks as adults.

Interestingly, a closer look at our imagine movement condition indicated that asking children to imagine the movement of the pieces had a negative effect on the boys' performance (the girls improved in this condition, but not appreciably more than they improved in the observe movement and practice conditions). The imagine movement instructions may have interfered with an effective strategy that the boys were already using. Several lines of research have suggested that encouraging explicit practice when an effective implicit strategy is already in use can disrupt performance (Beilock, Carr, MacMahon, & Starkes, 2002; Lutz et al., 2001). However, because we did not obtain strategy explanations from children before they were trained, we cannot directly test this hypothesis. More specific work examining strategies used before and after instruction is needed to explain the pattern of results found in our imagine movement condition.

We found that children in our study frequently conveyed strategies in gesture that were not expressed in the accompanying speech (see Figure 2). These gestured strategies occurred either

when the children did not talk at all or when they gestured about a strategy different from the one they expressed in speech (e.g., they gestured about moving the pieces but talked about the points on a shape). More important, children who performed better on the spatial transformation task often referred to movement in their gestures and not in their speech. Furthermore, although boys and girls did not differ in their talk about movement, boys gestured significantly more often about movement than girls and performed better on the spatial transformation tasks than girls.

Where do these sex differences come from and how can they be changed? Several lines of research have suggested that sex differences in spatial performance may be affected by differences in experience. Studies have shown that boys play more with "masculine" toys, which tend to be more spatial (i.e., blocks and other manipulatives, puzzles, balls, transportation toys; Connor & Serbin, 1977; Servin, Bohlin, & Berlin, 1999) and that these differences may be related to performance on spatial tasks (Fagot & Littman, 1976; Serbin & Connor, 1979). Similar findings have been reported for exposure to video games requiring spatial skills, such as Tetris. Boys report spending more time playing video games than girls, both in the home and in arcades (Dominick, 1984; S. Linn & Lepper, 1987; Quaiser-Pohl, Geiser, & Lehmann, 2006; Subrahmanyam & Greenfield, 1994). Moreover, several studies have found links between exposure to video games relying on spatial skills (both naturally and in experimental settings) and performance on spatial tasks (De Lisi & Wolford, 2002; Dorval & Pepin, 1986; McClurg & Chaillé, 1987; Quaiser-Pohl et al., 2006; Subrahmanyam & Greenfield, 1994). These findings have led researchers to suggest that the male advantage on certain spatial tasks may be affected by their differential play with spatial toys and video games. However, it is also possible that a preexisting spatial advantage leads males to engage more in spatial activities, which in turn enhances this advantage (Casey, Nuttall, & Pezaris, 1999).

Although this study does not speak to the origins of sex differences in spatial skills, it does suggest a possible means of enhancing spatial skills. Just as exposure to certain video games may help
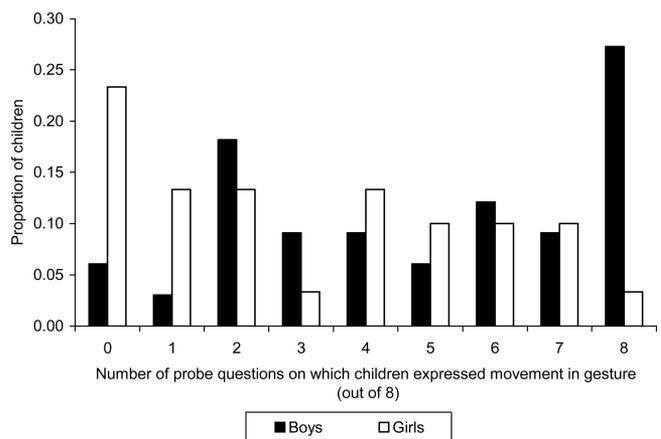
Figure 4. The proportion of boys and girls producing the movement strategy in gesture on the eight probe questions. The modal response for the girls was to produce movement in gesture on none of the eight questions; the modal response for the boys was to produce movement in gesture on all eight questions.

improve spatial skills, our findings suggest that gesture might be an excellent tool for teaching children how to solve spatial transformation tasks. Gesture has the potential to highlight the importance of moving the pieces in a way that actual movement cannot. In the observe movement condition that we used in our study, the experimenter placed the two pieces together and left them together for 2 s before pulling them back apart. The children may have focused on what the pieces looked like after the movement—that is, on the outcome of the movement—rather than on the movement per se. After all, noticing the shape the pieces make when put together is helpful in choosing a correct answer on the problem. Because gesture can depict movement without involving a resulting shape, it may help focus children's attention on the transformation itself and, in so doing, improve their skill at mentally transforming spatial information. Gesturing movement might therefore be a more effective instructional technique than demonstrating movement.

Our future studies will examine the effects of gesture on improving children's mental rotation skills. We believe this kind of input may be effective for several reasons. First, there is evidence that both watching someone else's gestures and producing one's own gestures can improve children's learning. Specifically, children learn from watching their teacher's gestures, whether those gestures convey different information from their speech (Singer & Goldin-Meadow, 2005) or the same information (Church, Ayman-Nolley, & Mahootian, 2004; Valenzeno, Alibali, & Klatzky, 2003). In addition, children who incorporate problem-solving strategies conveyed in their teacher's gestures into their own gestures are particularly likely to master the problem (Cook & Goldin-Meadow, 2006). In fact, neurological research has suggested that producing and observing gestures activates overlapping brain regions (i.e., portions of the premotor areas; Rizzolatti & Arbib, 1998). Neural priming may thus occur while either watching someone gesture or while gesturing oneself and may foster mental imagery of that type of movement. Second, there seems to be a strong link between spatial understanding and the use of gesture. Increasingly, researchers examining spatial understanding (including navigation) report that when people talk about space, they use their hands to display iconic gestures about moving through the space (Emmorey et al., 2000; Schaal et al., 2005). The time is ripe for a study examining whether adding gesture to instruction enhances children's ability to mentally manipulate spatial information.

In sum, we have confirmed earlier reports that boys are better than girls at spatial transformation tasks, even at age 5. Although our attempts at improving children's mental transformation skills using explicit instruction were no more effective than practice, our interviews provided hints about the strategies children may be using on the mental transformation task—and, at the same time, underscored the importance of examining the modalities through which children express their strategies. Children, particularly boys, frequently gestured about movement even when they did not talk about it. More important, referring to movement in gesture (but not talk) was associated with successful performance on our spatial transformation tasks. Our findings thus lead us to suggest that using gesture to instruct children may have a profound and positive impact on the development of early spatial skills, particularly mental transformation skills.

References

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology, 25,* 468–523.

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

Baenninger, M., & Newcombe, N. (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learning and Individual Differences, 7,* 363–379.

Beilock, S., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied, 8,* 6–16.

Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning ability at age 13: Their status 20 years later. *Psychological Science, 11,* 474–480.

Berenbaum, S. A., Korman, K., & Leveroni, C. (1995). Early hormones and sex differences in cognitive abilities. *Learning and Individual Differences, 7,* 303–321.

Bornstein, M. H., & Stiles-Davis, J. (1984). Discrimination and memory for symmetry in young children. *Developmental Psychology, 20,* 637–649.

Casey, M. B., Nuttall, R., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33,* 669–680.

Casey, M. B., Nuttall, R. L., & Pezaris, E. (1999). Evidence in support of a model that predicts how biological and environmental factors interact to influence spatial skills. *Developmental Psychology, 35,* 1237–1247.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31,* 697–705.

Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism, 7,* 303–319.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23,* 43–71.

Connor, J. M., & Serbin, L. A. (1977). Behaviorally based masculine- and feminine-activity-preference scales for preschoolers: Correlates with other classroom behaviors and cognitive tests. *Child Development, 48,* 1411–1416.

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7,* 211–232.

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology, 163*(3), 272–282.

Dominick, J. R. (1984). Videogames, television violence and aggression in teenagers. *Journal of Communication, 34,* 136–147.

Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills, 62,* 159–162.

Emmorey, K., Tversky, B., & Taylor, H. A. (2000). Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation, 2,* 157–180.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87,* 215–251.

Fagot, B., & Littman, I. (1976). Relation of preschool sex-typing to intellectual performance in elementary school. *Psychological Reports, 39,* 699–704.

Geary, D. C. (1996). Sexual selection and sex differences in mathematical abilities. *Behavioral and Brain Sciences, 19,* 229–284.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think.* Cambridge, MA: Harvard University Press.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100,* 279–297.

Goldin-Meadow, S., Nusbaum, H., Garber, P., & Church, R. B. (1993). Transitions in learning: Evidence for simultaneously activated strategies. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 92–107.

Halpern, D. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hampson, E., Rovet, J. F., & Altmann, D. (1998). Spatial reasoning in children with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Developmental Neuropsychology, 14,* 299–320.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology, 78,* 250–261.

Huttenlocher, J., Levine, S., & Vevea, J. (1998). Environmental input and cognitive growth: A study using time-period comparisons. *Child Development, 69,* 1012–1029.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature, 396,* 228.

Johnson, E. S. (1984). Sex differences in problem solving. *Journal of Educational Psychology, 76,* 1359–1371.

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

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

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–228). The Hague, the Netherlands: Mouton.

Kimura, D. (1999). *Sex and cognition.* Cambridge, MA: MIT Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48,* 16–32.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science, 7,* 54–60.

Lavergne, J., & Kimura, D. (1987). Hand movement asymmetry during speech: No effect of speaking topic. *Neuropsychologia, 25,* 689–693.

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

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

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

Linn, S., & Lepper, M. (1987). Correlates of children's usage of video games and computers. *Journal of Applied Social Psychology, 17,* 72–93.

Lutz, R., Landers, D. M., & Linder, D. E. (2001). Procedural variables and skill level influences on pre-performance mental practice efficacy. *Journal of Mental Imagery, 25,* 115–134.

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

McClurg, P. A., & Chaillé, C. (1987). Computer games: Environments for developing spatial cognition? *Journal of Educational Computing Research, 3,* 95–111.

McGuinness, D., & Morley, C. (1991). Sex differences in the development of visuo-spatial ability in pre-school children. *Journal of Mental Imagery, 15,* 143–150.

McNeill, D. (1992). *Hand and mind.* Chicago: University of Chicago Press.

Newcombe, N. S., Mathason, L., & Terlecki, M. (2002). Maximization of spatial competence: More important than finding the cause of sex differences. In A. McGillicuddy-De Lisi & R. De Lisi (Eds.), *Biology, society, and behavior: The development of sex differences in cognition* (pp. 183–206). Westport, CT: Ablex.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3,* 359–400.

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

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

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7,* 226–231.

Resnick, S. M., Berenbaum, S. A., Gottesman, I. I., & Bouchard, T. J. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology, 22,* 191–198.

Rimé, B. (1982). The elimination of visible behaviour from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology, 12,* 113–129.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21,* 188–194.

Schaal, A. G., Uttall, D. H., Levine, S. C., & Goldin-Meadow, S. (2005, May). *Children's gestures provide insight into their representations of space.* Poster presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Serbin, L. A., & Connor, J. M. (1979). Sex-typing of children's play preferences and patterns of cognitive performance. *Journal of Genetic Psychology, 134,* 315–316.

Servin, A., Bohlin, G., & Berlin, L. (1999). Sex differences in 1-, 3-, and 5-year-olds' toy choice in a structured play-session. *Scandinavian Journal of Psychology, 40,* 43–48.

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16,* 85–89.

Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology, 15,* 13–32.

Valenzeno, L., Alibali, M. W., & Klatzky, R. L. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28,* 187–204.

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

Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scales of Intelligence—Revised.* San Antonio, TX: Psychological Corporation.