Journal of Educational Psychology

Strategy Training Eliminates Sex Differences in Spatial Problem Solving in a STEM Domain

Mike Stieff, Bonnie L. Dixon, Minjung Ryu, Bryna C. Kumi, and Mary Hegarty

Online First Publication, November 4, 2013. doi: 10.1037/a0034823

CITATION

Stieff, M., Dixon, B. L., Ryu, M., Kumi, B. C., & Hegarty, M. (2013, November 4). Strategy Training Eliminates Sex Differences in Spatial Problem Solving in a STEM Domain. *Journal of Educational Psychology*. Advance online publication. doi: 10.1037/a0034823

Journal of Educational Psychology

0022-0663/13/$12.00 DOI: 10.1037/a0034823

Strategy Training Eliminates Sex Differences in Spatial Problem Solving in a STEM Domain

Mike Stieff
University of Illinois at Chicago

Bonnie L. Dixon, Minjung Ryu, and Bryna C. Kumi
University of Maryland, College Park

Mary Hegarty
University of California, Santa Barbara

Poor spatial ability can limit success in science, technology, engineering, and mathematics (STEM) disciplines. Many initiatives aim to increase STEM achievement and degree attainment through selective recruitment of high-spatial students or targeted training to improve spatial ability. The current study examines an alternative approach to increasing achievement that includes problem-solving strategy training. In this study, we examined how training in multiple problem-solving strategies affects science achievement and its relations to sex and spatial ability. We compared 3 interventions that trained either mental imagery strategies, analytic problem-solving strategies, or their combination in the context of a college chemistry course. As predicted, students adopted more analytic strategies after analytic training, and women used significantly more analytic strategies than men after instruction. Training in the combined use of mental imagery and analytic strategies eliminated sex differences in achievement, but training in a single type of strategy resulted in a male achievement advantage. Our work demonstrates that achievement is dependent not only on spatial ability but also on strategy choice, and that strategy training offers a viable route to improving the performance of female students.

Keywords: spatial ability, problem solving, spatial reasoning

Science, technology, engineering, and mathematics (STEM) professionals routinely engage in tasks that involve identifying spatial information and reasoning about dynamic spatial relations, such as planetary orbits and geologic structures. Indeed, such spatial problem solving is a fundamental component of learning and practice in many STEM disciplines (Gilbert, 2005; National Research Council, 2006). Given the extensive role of spatial problem solving required in STEM fields, it is unsurprising that students who perform well on spatial ability measures are more likely to enjoy STEM courses, score higher on STEM assessments, and pursue STEM careers (Wai, Lubinski, & Benbow, 2009). In light of correlations between spatial ability and STEM achievement, some have called for academic institutions to use spatial ability to select students for participation in STEM degree programs or to offer courses that attempt to improve students' spatial abilities (e.g., Lubinski, 2010; Sorby, 2009).

Controversially, spatial ability is debated as a causal factor that explains the underrepresentation of females in STEM disciplines (Halpern et al., 2007), as females have been found to underperform males on measures of spatial ability (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). Because of these sex differences, it has been suggested that women may be disadvantaged in STEM fields due to spatial ability and that even high-spatial females will be unlikely to pursue scientific careers (Humphreys, Lubinski, & Yao, 1993; Shea, Lubinski, & Benbow, 2001; Wai et al., 2009). In response, several programs have explored the potential of educational interventions that target spatial ability to increase the achievement and retention of women in STEM degree programs (Hsi, Linn, & Bell, 1997; Miller & Halpern, 2013; Sorby, 2001). Although a recent meta-analysis has demonstrated that spatial ability can be improved reliably with training (Uttal et al., 2013), there is not strong evidence that spatial ability training transfers to STEM learning such that it reliably improves STEM achievement for women or students in general. Thus, questions remain about whether spatial ability offers an adequate explanation for sex differences in STEM achievement and STEM career choice.

One explanation for the failure of spatial ability training to transfer to STEM achievement is that such training fails to attend to sex differences in strategy preferences and neglects the role of analytic strategies in STEM problem solving. Outside STEM disciplines, men and women are often seen to employ qualitatively

---

Mike Stieff, Department of Chemistry and Learning Sciences, University of Illinois at Chicago; Bonnie L. Dixon, Department of Chemistry and Biochemistry, University of Maryland, College Park; Minjung Ryu, Department of Curriculum and Instruction, University of Maryland, College Park; Bryna C. Kumi, Department of Chemistry and Biochemistry, University of Maryland, College Park; Mary Hegarty, Department of Psychological and Brain Sciences, University of California, Santa Barbara.

This work was supported by National Science Foundation (NSF) Grant DRL-0723313 to Mike Stieff and Bonnie L. Dixon and NSF Grant DRL-0722333 to Mary Hegarty. We gratefully acknowledge the support of Timothy Tasker regarding data analysis.

Correspondence concerning this article should be addressed to Mike Stieff, Department of Chemistry and Learning Sciences, University of Illinois at Chicago, 1007 West Harrison Street, Chicago, IL 60607. E-mail: mstieff@uic.edu

different strategies to solve spatial problems. Males display a preference and aptitude for using imagistic strategies, whereas females employ a preference and aptitude for using analytic strategies (Heil & Jansen-Osmann, 2008; Lawton, 2010; Maccoby & Jacklin, 1974; Pezaris & Casey, 1991). Moreover, spatial problem solving in STEM routinely involves strategies that do not rely on spatial imagery: In fact, a range of strategies yield correct solutions to problems found in engineering (Hsi et al., 1997; Schwartz & Black, 1996), chemistry (Stieff, 2007, 2011; Stieff & Raje, 2010), and mathematics (Hegarty & Kozhevnikov, 1999; Lowrie & Kay, 2001). These strategies include various heuristics, algorithms, and other analytic strategies that often involve systematic modification of diagrams to achieve solutions. Using these learned strategies, problem solvers capitalize on the formalisms of discipline-specific representations to gain insight into relevant spatial information or in some cases to generate problem solutions without attending to spatial information at all.

Although strategy variation has been established in protocol studies with small numbers of participants, there have been no systematic investigations of the relations among strategy choice, sex, spatial ability, and achievement in the context of extended STEM instruction. Questions remain about which students are more likely to adopt different strategies during routine problem solving or benefit from targeted strategy instruction. For example, men and women may employ and benefit from different strategies in light of established sex differences in strategy use, and low-spatial students may not apply a specific strategy as successfully as high-spatial students. Without a clear understanding of the relations among strategy choice, sex, spatial ability, and achievement, it is not obvious whether equitable learning outcomes are more likely from interventions that attempt to improve spatial ability, teach a specific strategy, or train students to use multiple strategies. The present study focused on the relative effectiveness of targeted strategy training for men and women in the context of the STEM discipline of organic chemistry. The results of our study indicate that targeted strategy instruction can decrease the magnitude of sex differences in STEM achievement on authentic achievement assessments and that strategy choice, in addition to spatial ability, predicts achievement in organic chemistry.

Sex Differences in Strategy Use

Detailed analyses of spatial problem solving indicate that men and women do not necessarily apply the same strategies to solve spatial problems. Males display a preference and aptitude for using imagistic strategies, whereas females employ a preference and aptitude for using analytic strategies (Heil & Jansen-Osmann, 2008; Lawton, 2010; Maccoby & Jacklin, 1974; Robert & Chevrier, 2003). *Imagistic strategies* involve reasoning by generating and transforming internal visual–spatial images; *analytic strategies* involve decomposing a visual–spatial image and reasoning via abstracted rules that are often applied to external representations of spatial information. Regardless of the degree of difference in these strategies, both are effective routes to successful spatial problem solving for men and women.

Sex differences in strategy choice have been studied most extensively in the literature on spatial ability (outside the context of STEM problem solving). Self-reports of strategy use on tests of mental rotation (Freedman & Rovegno, 1981; Peters, Chisholm, & Laeng, 1995), field (in)dependence (Allen & Hogeland, 1978), and wayfinding (Bosco, Longoni, & Vecchi, 2004; Dabbs, Chang, Strong, & Milun, 1998; Lawton, 1994, 2010) indicate not only that men and women apply multiple strategies, but that the sexes differ in strategy preference. On each of these measures, females are more likely to report strategies that permit piecemeal analysis of spatial information represented in task stimuli, use verbal labels to rerepresent spatial information, and leverage external tools to offload spatial information, such as gesture, diagrams, and concrete artifacts. In contrast, males are more likely to report strategies that involve the construction of holistic visual–spatial images, mental simulation of spatial transformations, and other processes that operate on nonverbal representations.

In addition to self-reports, response time and accuracy data also indicate a preference for imagistic strategies among men and analytic strategies among women. When comparing the mental rotation speed of men and women, Heil and Jansen-Osmann (2008) observed that the response time of women, but not men, is dependent upon the complexity of the mental rotation stimulus. Mental rotation speed slows for more complex objects among women, which suggests that women selectively apply an analytic strategy that involves deconstructing complex stimuli to reason about individual components. In contrast, men rotate both simple and complex stimuli at equal speeds, suggesting that men apply an imagistic strategy that involves rotating and comparing stimuli as whole objects. Interestingly, as the spatial complexity of a task increases, sex differences in strategy use become increasingly pronounced (Halpern et al., 2007), suggesting that men and women may employ different cognitive resources purposefully to reduce task difficulty. Furthermore, sex differences in spatial task performance are reduced when time limits are not imposed (Voyer, 2011), suggesting that these differences may be due more to the relative efficiency of male and female strategies for these tasks than to the eventual success of these strategies.

Functional magnetic resonance imaging studies of mental rotation offer additional evidence that men and women routinely apply different strategies when engaged in spatial problem solving (Hugdahl, Thomsen, & Ersland, 2006; Jordan, Wustenberg, Heinze, Peters, & Jäncke, 2002; Weiss et al., 2003). When completing a classic test of mental rotation, parietal activation predominates in males; however, females show additional activation in the frontal lobe that is not observed in males. These patterns of activation indicate that men and women are engaged in different processing mechanisms. The localized pattern of activation in the parietal lobe suggests that men reason about spatial images as visual gestalts, in keeping with an imagistic strategy. Conversely, activation of the inferior frontal gyrus among women implies that women utilize language functions to apply verbal labels for representing abstracted spatial information, in keeping with an analytic strategy.

Alternative Strategies for Problem Solving in Science

The alternative strategies preferentially employed by the sexes are highly congruent with the alternative strategies employed by novices and experts in STEM disciplines. As noted above, the various strategies available for spatial problem solving in STEM disciplines can also be differentiated into those that are more imagistic and those that are more analytic (Stieff, Hegarty, &

Dixon, 2010; Stieff, Ryu, & Dixon, 2010). *Spatial-imagistic strategies* involve extensive reasoning about scientific phenomena via the generation, inspection, and manipulation of internal visual–spatial images and include mental rotation, perspective taking, and spatial visualization processes. *Spatial-analytic strategies* involve the application of algorithms and heuristics to external representations that preserve and transform spatial information without invoking internal visual–spatial representations. Importantly, both categories of strategies involve reasoning about spatial relationships important in STEM disciplines and thus exclude those strategies that operate on nonspatial information, such as verbal or mathematical information.

Figure 1 illustrates the application of both types of strategies to solve a common assessment item in undergraduate organic chemistry, which requires the problem solver to compare two diagrams to determine whether both represent the same molecule or different molecules that are mirror images of each other. The task is computationally equivalent to items found on classical tests of mental rotation that typically include generic three-dimensional shapes. In the figure, the top pathway represents a spatial-imagistic strategy that involves mentally simulating the rotation of one diagram into congruence with the other to verify they superimpose perfectly (and are thus identical molecules). The bottom pathway represents a spatial-analytic strategy that involves annotating each diagram according to disciplinary rules that allow the problem solver to compare the relative spatial orientation of subcomponents in the diagram: If the relative orientation is identical, the diagrams represent identical molecules. Of note, such spatial-analytic strategies have also been observed on classic tests of mental rotation that include generic three-dimensional shapes (Just & Carpenter, 1985).

Analyses of problem solving by experts in various STEM disciplines indicate that experts tend to apply spatial-analytic strategies that rely on the manipulation of domain specific representations that are abstracted from the spatial information given in a problem (Schwartz & Black, 1996; Stieff, 2007; Stieff & Raje, 2008, 2010) rather than the ability to transform mental images, which spatial ability tests typically measure (Hegarty & Waller, 2006). In contrast, students often employ spatial-imagistic strategies as a first step in problem solving; however, as expertise develops they begin to employ more spatial-analytic strategies (Schwartz & Black, 1996; Stieff, 2011). Importantly, spatial-analytic strategies can be induced spontaneously as well as learned from direct instruction (Lowrie & Kay, 2001; Schwartz & Black, 1996; Stieff, 2007; Stieff, Ryu, Dixon, & Hegarty, 2012). In some cases, students learn to employ spatial-analytic strategies after extended practice (Schwartz & Black, 1996). In other cases, direct instruction from an expert can cause students to switch immediately from a spatial-imagistic to spatial-analytic strategy (Stieff, 2007).

Many spatial-analytic strategies are heuristic in nature and are applicable to a narrow range of problems, so they have to be used in conjunction with spatial-imagistic strategies in the broader problem-solving situation. The cooperative use of both types of strategies is readily observed in STEM disciplines where novices employ imagistic strategies as a first approach to solve spatial problems with analytic strategies being induced or adopted later, and both experts and novices tend to fall back on spatial-imagistic strategies when their spatial-analytic heuristics no longer apply (Schwartz & Black, 1996; Stieff, 2007). For example, when Schwartz and Black (1996) asked people to solve gear chain problems (determining the direction of rotation of gears in a gear chain), participants' initially used a spatial-imagistic strategy to simulate the motion of the individual gears, but on the basis of these simulations, discovered the simple rule that any two interlocking gears must move in opposite directions and switched to a spatial-analytic strategy. Both groups reverted to the mental simulation strategy, however, when given a novel type of gear problem in which the gears formed a ring. Results such as these suggest that effective problem solving in spatial domains involves an interplay between imagistic and analytic strategies (Hegarty, 2004, 2010) and that analytic strategies may help bootstrap students into using imagistic strategies more effectively as expertise grows.

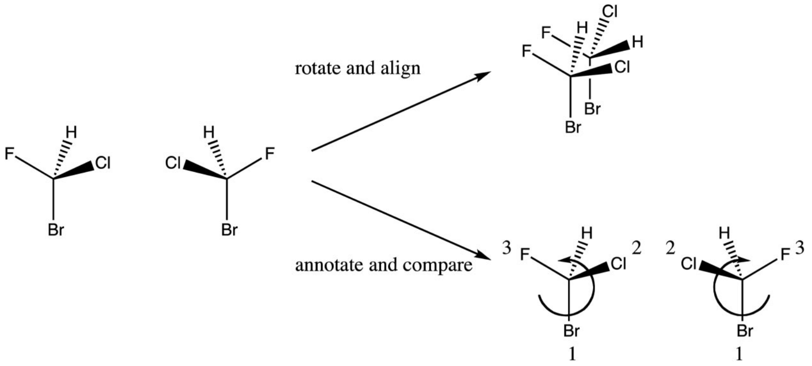

Figure 1. Alternative strategies for making identity judgments about molecular diagrams. The top pathway illustrates a mental rotation strategy that involves aligning each diagram by rotating around the picture plane. The bottom pathway illustrates an algorithm that involves annotating the position of the atoms and comparing their relative orientations. Dashed lines represent objects projecting behind the page, and dark lines represent objects projecting above the page. From "Sex Differences in the Mental Rotation of Chemistry Representations," by M. Stieff, 2013, *Journal of Chemical Education*, 90. Copyright 2013 by the American Chemical Society.

Attenuating Sex Differences With Targeted Strategy Training

Established sex differences in spatial ability and strategy suggest that men and women may not benefit equally from strategy training in STEM disciplines. Meta-analyses indicate that the male spatial ability advantage is robust: The majority of interventions reliably improve the spatial ability of both sexes yet do not eliminate the achievement gap between men and women (Uttal et al., 2013). Sex differences in strategy use may provide one explanation for the failure of training interventions to eliminate sex differences. Previous attempts to train spatial ability have focused primarily on training and practice using male-preferred imagistic strategies, such as visualizing spatial transformations. Thus, if it is effective in transferring to STEM achievement, training that focuses on teaching imagistic strategies is likely to improve STEM achievement for both sexes, thereby maintaining the sex difference in STEM achievement.

Furthermore, effects of targeted training in imagistic strategies may not be enduring, and there is limited evidence that it transfers to STEM achievement. Whereas Sorby (2009) found positive effects of targeted imagistic training on performance in engineering graphics with self-selected students, a recent study that used random assignment demonstrated that a supplemental course using identical training was unable to produce lasting improvements in STEM achievement for men or women (Miller & Halpern, 2013). In this study, sex differences on spatial ability measures were narrowed after training, and spatial training improved physics problem solving in the short term for both sexes; however, when tested 8 months later, participants displayed the same large sex differences in spatial ability observed at pretest and the effects of training on physics achievement had dissipated. Problematically, the intervention failed to eliminate sex differences in STEM achievement, as men performed better than women on tests of physics knowledge regardless of whether they had received spatial training.

An alternative approach to attenuating sex differences in STEM achievement is to train students to use analytic strategies, which have been found to be readily adopted by experts in STEM disciplines, and are ostensibly the preferred strategies of females. As reviewed above, analytic strategies support successful problem solving in STEM disciplines; however, these strategies are often overlooked by reform efforts that urge training male-preferred imagistic strategies (National Research Council, 2006; Sorby, 2009). Given women's preference for and success in using analytic strategies for spatial problem solving, it is reasonable that targeting training to their strengths in using analytic strategies may selectively increase female STEM achievement by helping them better apply analytic strategies for disciplinary problem solving. Previous approaches assume that women's preference for analytic strategies for spatial problem solving is causally responsible for sex differences in achievement. In contrast, we consider that it is the lack of instruction in how to make use of analytic strategies in STEM classrooms that disadvantages women.

We know of only one previous study that compared analytic and imagistic strategy training for spatial problem solving. Kyllonen, Lohman, and Snow (1984) compared the effectiveness of these two forms of training for performance on the Paper Folding Test (Ekstrom, French, Harman, & Derman, 1976). Their results revealed complex interactions of strategy training with fluid and crystallized ability and with item difficulty such that they could not confirm an overall advantage of either form of training. The most striking result of the study by Kyllonen et al. was that individuals varied their strategy across problems, suggesting that although men and women may prefer to use one type of strategy, they readily employ multiple strategies in response to task demands. This finding has also been observed in organic chemistry by Stieff and Raje (2010), who documented that both male and female chemists employed spatial-imagistic and spatial-analytic strategies in tandem to reason about spatial relationships in molecules.

Thus, it is also possible that training in the use of both imagistic strategies and discipline-specific analytic strategies together in one course might be more effective at eliminating sex differences than training that focuses exclusively analytic approaches. There are two reasons to expect that combined strategy training may eliminate sex differences in STEM achievement. First, instructional interventions that train women to relate multiple strategies may provide women with a larger range of strategies to apply. As discussed, men may be able to reach greater levels of achievement in analytic instructional conditions because they are more likely to be already proficient in the use of imagistic strategies and will be better able to relate them to analytic strategies. By explicitly training women to relate analytic and imagistic strategies, they may reach comparable levels of performance to men, who make such relations spontaneously. Second, given STEM experts' tendencies to transition from imagistic to analytic strategies with experience (Schwartz & Black, 1996; Stieff, 2007) and to use both types of strategies in tandem on complex problems (Stieff & Raje, 2010), training that helps students, particularly women, to relate imagistic and analytic strategies together may yield the largest gains in achievement. To our knowledge no studies have directly addressed sex differences in spatial problem solving through targeted training in analytical strategies or the use of analytic strategies together with imagistic strategies in the context of a STEM discipline.

Present Study

Here we conducted targeted strategy training interventions in imagistic strategies, analytic strategies, and combined strategies to examine whether men and women employ different problem-solving strategies in the STEM discipline of college organic chemistry and how training can influence strategy use and achievement. Organic chemistry provides an opportune context to study the differential effects of single and combined strategy training on STEM achievement. Among the STEM disciplines, organic chemistry is a highly spatial domain and one in which men and high-spatial students typically outperform women and low-spatial students (Harle & Towns, 2011). A significant component of instruction in organic chemistry involves teaching about the spatial relationships within and between molecules to explain chemical and physical properties of various substances. In a typical organic chemistry course, students must learn to discriminate between organic molecules that contain identical atoms that are arranged in differing ways, to predict how atoms move through space over time in a chemical reaction, and to explain how the spatial relationships within a molecule relate to other concepts, such as polarity, bonding, and reactivity. Given the degree of spatial information that students must learn and use to problem solve during the course of instruction, the discipline is highly suited for studies that examine the relationship between spatial ability, achievement, sex, and strategy training.

First, we predicted that students would employ primarily spatial-imagistic strategies prior to instruction, but that these strategies would be replaced by more spatial-analytic strategies after disciplinary instruction (Hypothesis 1). Second, we expected that training in spatial-analytical strategies would lead students to employ more spatial-analytic strategies after instruction in the Analytic and Combined Interventions (Hypothesis 2). Finally, we predicted that both analytic training (reflecting female-preferred strategies) and combined training (reflecting expert practice) would reduce sex differences in achievement (Hypothesis 3). To further investigate the effect of the interventions, we also examined the relations among strategy use, sex, spatial ability, and achievement using structural equation modeling methods. We tested a model that proposes that spatial ability predicts strategy use, spatial ability and strategy use independently predict course achievement, and together spatial ability and strategy mediate sex differences in course achievement (Hypothesis 4).

Method

Participants

Participants were recruited from three sequential semesters of an introductory organic chemistry course at an East Coast research-extensive university. Approximately two thirds (62.5%) of the enrolled students were women, which is typical for this course. The study sample consisted of 372 students (147 men and 225 women) who (a) were over 18 years of age, (b) completed a spatial ability battery at the beginning of the semester, (c) self-reported their SAT composite scores, and (d) took the Chemistry Spatial Problem Solving and Strategy Preference Assessment (CSPSSPA) at both the beginning and end of the semester. One hundred and six students (48 male, 58 female) received the Analytical Strategy Intervention, 124 students (42 male, 82 female) completed the Imagistic Strategy Intervention, and 142 students (57 male, 85 female) completed the Combined Strategy Intervention. Participants who completed all measures comprised 70% of all students enrolled in the Analytic Intervention, 84% of all students enrolled in the Imagistic Intervention, and 76% of all students enrolled in the Combined Intervention. Participants were enrolled in a raffle each semester in which three participants were awarded an iPod Touch valued at $150.

Procedure

Students received one of three designed interventions over sequential semesters: In each semester, strategy training was delivered via three 1-hr workshops that were supplementary to normal curriculum activities and emphasized in lecture. Spatial ability and academic aptitude were assessed prior to each intervention. Chemistry achievement and strategy choice were assessed with the 12-item CSPSSPA during the first and last week of each semester.

Interventions

The interventions differed in how class lectures were taught and in targeted strategy training (i.e., analytic heuristics, mental imagery, and combined) in workshops attended by all students. All courses were taught by a female instructor (with 10 years of experience) and a male teaching assistant (with 1 year of experience) prior to participating in the intervention. The instructor (an author of this article) asserted that all strategies were equally useful and that she had observed students to vary in their ability to apply each type of strategy. An independent researcher observed workshop and lecture activities to confirm that the instructor adhered to the pedagogical methods detailed here.

Lectures. Strategies were first introduced to students in a lecture setting by the course instructor and teaching assistant. In the Analytic Intervention, imagery language and spatial gestures were minimized during teaching, and imagistic strategies were not presented. Problem-solving strategies centered solely on the use of diagrams for problem solving using disciplinary heuristics. The instructor and teaching assistant used no concrete models in instruction, and students were not encouraged to purchase or to use models. In the Imagistic Intervention, the instructor and teaching assistant used both concrete models and spatial gestures to depict spatial relationships during instruction. Students were not taught any disciplinary heuristic strategies for solving spatial tasks and were encouraged to mentally visualize the spatial relationships depicted in disciplinary diagrams while problem solving. Students were encouraged to purchase models to practice visualizing spatial relationships. In the Combined Intervention, both disciplinary heuristics and imagistic strategies were taught, and concrete models were employed. After using a disciplinary heuristic to solve a problem, the instructor would verify the accuracy of the solution by illustrating the spatial relationships in a concrete model. Both strategy types were offered as equally useful, and students were encouraged to use the strategy (or strategies) they believed was more effective. Necessarily, the amount of time devoted to each type of strategy was less than in the other two interventions. Students were encouraged to purchase models and to use models, and instruction emphasized relating analytic strategies to imagistic strategies relevant to a given problem.

Workshops. Three 1-hr workshops were enacted during each intervention in mandatory recitation sections of 20 students. The instructor taught each workshop following the week in which novel spatial concepts were introduced in the lecture. The goal of the workshops was to revisit problem-solving strategies presented in lecture and practice applying the strategies to novel problems in small groups. Each workshop followed the same schedule: The first 15 min reviewed specific problem-solving strategies related to the given intervention that were presented in lecture the previous week. In the remaining 45 min, students worked in small groups to complete a worksheet that allowed the students to practice applying taught strategies.

Homework sets. Three homework sets were assigned for credit during each intervention to allow students to practice the application of the strategies presented in each workshop independently. Students were permitted to work directly with the instructor outside class to complete the homework using the strategies presented in the relevant workshop.

Measures

Spatial ability. Students were tested on three measures of spatial ability, the Vandenberg Mental Rotation Test (Vandenberg

& Kuse, 1978), the Paper Folding Test (Ekstrom et al., 1976), and a modified version of Guay's Visualization of Views Test (Guay & McDaniel, 1976).

In the Vandenberg Mental Rotation Test, participants view a depiction of a three-dimensional criterion figure and four test figures. Their task is to determine which test figures are rotations of the criterion figure, as quickly and accurately as possible. The test is administered in two sections, with 10 items and a time limit of 3 min for each section. A participant can score up to 4 points for each question, so the maximum possible score is 80. In the Paper Folding Test, participants are shown two standard figures representing a piece of folded paper with circles indicating where the paper was punched. To the right of these figures are five squares with different configurations of circles. The participant's task is to identify the pattern of holes that would result when the standard figure was unfolded. There are two sections with 10 items and a time limit of 3 min for each. The maximum score possible is 20. In the Visualization of Views Test, a three dimensional object is depicted in the center of a transparent cube. The same object from a different viewpoint is depicted below the cube. The task is to indicate the corner of the cube from which the new view of the object is taken (24 items, 8 min).

The tests were administered via a website that presented the standard instructions for the test and displayed the test items. In the Vandenberg Mental Rotation Test and Paper Folding Test, participants indicated their responses by mouse clicks on boxes placed under each of the answer choices. In the case of the Visualization of Views Test, participants responded by a mouse click on one of the corners of the cube. At the end of the time limit for all tests and test sections, the test items disappeared and were replaced by a message stating that the time allotted was over.

Academic aptitude. Academic aptitude was measured with self-reported composite SAT scores, which have been found to be a good proxy for actual SAT performance (Mayer et al., 2007).

Chemistry spatial problem solving and strategy preference. Spatial problem solving in chemistry and strategy preference were measured by the CSPSSPA, which consisted of 12 items that assessed student understanding of spatial relationships relevant to organic molecules and organic transformations. Collectively, the 12 items assessed concepts of stereochemistry, stereoselectivity, and regiochemistry covered in the organic chemistry curriculum. Each item required students to identify spatial relationships between molecules and substituents within a molecule. All items were fixed-choice response, and all were scored for correctness (1 = correct, 0 = incorrect). Spatial problem-solving ability on the assessment was determined by calculating the total number of correct answers. Previously, the assessment was piloted with cognitive interviews and field tests to establish the content validity of each item (Stieff & Raje, 2010; Stieff et al., 2012). Criterion validity of the achievement test was established via correlation with final course grade, $r(372) = .62$, $p < .001$, demonstrating that the test was representative of student achievement in organic chemistry.

Participants were asked to report the strategy they used to solve each item by selecting from a fixed list of strategies applicable to each problem. Participants were allowed to choose more than one strategy and to write in their own if they believed that none of the choices matched their strategy. Each list of strategies for individual problems was developed using cognitive interviews, previously reported by Stieff, Ryu, and Dixon (2010). Each strategy was coded according to a priori categories of strategy type. Briefly, categories included those strategies that relied more extensively on reasoning via mental imagery (spatial-imagistic), rules and heuristics that operated on spatial information (spatial-analytic), construction of novel diagrams (spatial-diagrammatic), and rules and heuristics that operated on nonspatial information (algorithmic). Participants could also indicate if they knew the answer to a problem (recall) or if they randomly guessed (guessing); responses were not included in the analysis. For the present study, only spatial-imagistic and spatial-analytic strategy use was analyzed. Multiple spatial-analytic and spatial-imagistic strategies are applicable to each item on the assessment and are included on the fixed-choice list for each item. For a complete analysis of strategy reports on the assessment, see Hegarty, Stieff, and Dixon (2013).

The construct validity and reliability of the self-report strategy items was determined in a separate study (Stieff et al., 2012) in which the 103 organic chemistry students completed the strategy items via personal response devices ("clickers"). Students reported that at least one of the listed self-report strategies represented their applied strategy on 98% of the items. In only two cases did a student report being unable to find his strategy on the list; analysis of his write-in strategy revealed that he reported using mental rotation to problem solve, but described mental rotation using personal language different from the language used on the survey.

Chemistry achievement. In addition to performance on the research-designed assessment, chemistry achievement was also measured with five (three exams and two quizzes) instructor-designed content assessments administered during the normal course of instruction. Departmental teaching assistants under the supervision of the instructor scored all assessments. Assessments were administered at approximately the same time point during each course, and each assessed the domain content knowledge and problem-solving ability related to the course curriculum. Assessment items differed between semesters but were isomorphic to one another and differed only in the molecule referenced. Students could receive a maximum of 150 on each quiz and 300 on each exam for a total possible combined score of 1,200 points, which was used to determine a final letter grade in the course.

Results

Descriptive Statistics

Descriptive statistics for ability measures are included in Table 1. Students in the three interventions did not differ significantly on the Mental Rotation Test, $F(2, 369) = 0.69$, $p = .50$; the Paper Folding Test, $F(2, 369) = 0.18$, $p = .83$; or the Visualization of Views Test, $F(2, 369) = 0.99$, $p = .37$. Men were observed to outperform women on the Mental Rotation Test, $F(1, 370) = 62.95$, $p < .001$, $\eta_p^2 = .15$, and the Visualization of Views Test, $F(1, 370) = 49.46$, $p < .001$, $\eta_p^2 = .12$, but not on the Paper Folding Test, $F(1, 370) = 2.18$, $p = .14$, which is consistent with previous research (Linn & Peterson, 1985; Voyer, Voyer, & Bryden, 1995).

SAT Composite score did not differ among students in the three interventions, $F(2, 363) = 1.34$, $p = .26$, or between men and women, $F(1, 364) = 0.926$, $p = .25$, indicating that all groups were equivalent in academic aptitude prior to participating in the interventions. Achievement and strategy use (i.e., self-reports of spatial-imagistic and spatial-analytic strategies) are reported in

Table 1
Average Spatial Ability and SAT Scores (Standard Deviation) Among the Three Interventions

Intervention	Imagistic	Analytic	Combined	Men	Women
Visualization of Views	11.1 (6.7)	10.6 (6.9)	11.5 (6.7)	14.6 (6.6)	9.8 (6.1)
Mental Rotation	30.4 (18.5)	29.8 (18.4)	31.3 (18.8)	40.8 (17.8)	27.3 (14.7)
Paper Folding	11.5 (4.1)	11.4 (3.9)	10.7 (4.3)	12.0 (4.2)	11.3 (3.7)
SAT Composite	1964 (212)	1919 (212)	1932 (234)	1932 (202)	1949 (249)

Table 2 and fully analyzed below. A more complete analysis of strategy learning in each intervention can be found in Hegarty et al. (2013).

Do Students Change Strategies as a Result of Instruction?

We predicted that students would employ primarily spatial-imagistic strategies prior to instruction, but that these strategies would be replaced by more spatial-analytic strategies after disciplinary instruction (Hypothesis 1). We also predicted that this would be particularly true for students in the Analytic and Combined Interventions, because these interventions included training in analytic strategies (Hypothesis 2). We first examined the raw number of spatial-imagistic and spatial-analytic strategies reported by students. With all intervention groups combined, participants reported using a plurality of spatial-imagistic strategies (48% of all reported strategies) and relatively few spatial-analytic strategies (3% of all reported strategies) prior to instruction. (The remaining 49% of reported strategies primarily involved guessing.) At pretest there were no significant sex differences in the use of either spatial-analytic strategies, $F(1, 366) = 0.392$, $p = .68$, or spatial-imagistic strategies, $F(1, 366) = 0.965$, $p = .38$. The three intervention groups also did not differ significantly in use of spatial-analytic strategies, $F(2, 366) = 1.55$, $p = .20$, at pretest. However, students differed in the use of spatial-imagistic strategies at pretest among the three interventions, $F(2, 366) = 3.69$, $p = .03$, $\eta_p^2 = .02$. Students in the Analytic Intervention reported using more spatial-imagistic strategies ($M = 3.20$, $SD = 3.4$) than students in both the Imagistic ($M = 2.34$, $SD = 2.0$), $F(1, 228) = 4.79$, $p = .03$, $\eta_p^2 = .021$, and Combined Interventions ($M = 2.34$, $SD = 2.3$), $F(1, 246) = 4.94$ $p = .03$, $\eta_p^2 = .02$. Spatial-imagistic strategy use did not differ between the Combined and Imagistic Interventions, $F(1, 264) = 0.001$, $p = .98$.

To assess changes in strategy use, we first analyzed how the number of reported spatial-imagistic and spatial-analytic strategies changed with a 2 (sex: male, female) × 3 (intervention: imagistic, analytic, combined) × 2 (time: pretest, posttest) repeated-measures analysis of variance (ANOVA). As shown in Table 2, use of both types of strategies increased after instruction. First, there was a large increase in the average number of spatial-imagistic strategies reported by students from pre- to posttest, $F(1, 366) = 223.29$, $p < .001$, $\eta_p^2 = .38$. A main effect of intervention was also observed, $F(2, 364) = 10.66$, $p = < .001$, $\eta_p^2 = .06$. Planned contrasts show that students in the Combined Intervention ($M = 3.92$, $SD = 3.6$) reported adopting more spatial-imagistic strategies after instruction than students in both the Imagistic ($M = 2.70$, $SD = 3.5$), $F(2, 262) = 7.65$, $p = .006$, $\eta_p^2 = .03$, and Analytic Interventions ($M = 1.81$, $SD = 3.8$), $F(1, 242) = 19.99$, $p < .001$, $\eta_p^2 = .08$. Although not statistically significant, a trending difference suggested that students adopted more spatial-imagistic strategies in the Imagistic Intervention than in the Analytic Intervention, $F(1, 226) = 1.66$, $p = .08$. A main effect of sex on increased use of spatial-imagistic strategy use was also observed, $F(1, 366) = 5.58$, $p = .019$, $\eta_p^2 = .02$; however, no interaction between sex and intervention was present, $F(2, 366) = 0.82$, $p = .44$. Regardless of intervention, men ($M = 3.42$, $SD = 3.8$) adopted more spatial-imagistic strategies than women ($M = 2.58$. $SD = 3.6$).

Second, there was a large increase in the average number of spatial-analytic strategies reported by students from pre- to posttest, $F(1, 366) = 545.98$, $p < .001$, $\eta_p^2 = .60$. A main effect of intervention was also observed, $F(2, 366) = 7.08$, $p = < .001$, $\eta_p^2 = .04$. Planned contrasts show that students in the Analytic Intervention ($M = 3.83$, $SD = 2.5$) adopted more spatial-analytic strategies than students in either the Combined Intervention ($M = 2.36$, $SD = 2.12$), $F(1, 244) = 13.24$, $p < .001$, $\eta_p^2 = .05$, or the Imagistic Intervention ($M = 2.60$, $SD = 2.1$), $F(1, 226) = 6.80$,

Table 2
Descriptive Statistics for Strategy Use and Achievement Measures

Intervention	Imagistic		Analytic		Combined	
	Men	Women	Men	Women	Men	Women
Spatial-imagistic strategies pretest	2.31 (2.0)	2.22 (2.0)	2.81 (2.3)	3.41 (3.8)	2.42 (2.3)	2.45 (2.5)
Spatial-imagistic strategies posttest	5.5 (3.2)	4.87 (2.7)	5.92 (3.3)	4.19 (3.1)	6.75 (3.4)	5.67 (3.1)
Spatial-analytic strategies pretest	0.17 (0.38)	0.20 (0.50)	0.23 (0.42)	0.37 (0.75)	0.17 (0.48)	0.27 (0.75)
Spatial-analytic strategies posttest	2.40 (1.8)	3.08 (2.1)	2.79 (2.1)	4.36 (2.2)	2.08 (2.1)	2.94 (2.2)
CSPSSPA pretest	2.8 (1.3)	2.99 (1.2)	2.64 (1.4)	2.75 (1.5)	3.03 (1.43)	2.95 (1.5)
CSPSSPA posttest	7.11 (2.1)	6.23 (2.3)	7.7 (2.1)	6.94 (2.1)	7.2 (2.3)	7.55 (2.2)
Course achievement	473 (112)	446 (106)	488 (105)	441 (120)	491 (136)	488 (103)

Note. Values indicate average number of reported strategies on the strategy questionnaire and total points on achievement measures, respectively. Standard deviation is indicated in parentheses. CSPSSPA = Chemistry Spatial Problem Solving and Strategy Preference Assessment.

p = .01, η_p^2 = .03. No differences were observed between the Combined Intervention and the Imagistic Intervention groups, F(2, 262) = 0.622, p = .43. A main effect of sex on increased frequency of spatial-analytic strategy use was also observed, F(1, 366) = 12.21, p = .001, η_p^2 = .03; however, no interaction between sex and intervention was present, F(2, 366) = 0.54, p = .58. Regardless of intervention, women (M = 3.02, SD = 2.3) adopted more spatial-analytic strategies than men (M = 2.27, SD = 2.2).

To determine the change in student strategy preference, we next examined changes in the ratio of spatial-imagistic and spatial-analytic strategies reported relative to the total number of strategies reported (see Figure 2) with a 2 (sex) × 3 (intervention) × 2 (time) repeated-measures ANOVA. In keeping with Hypothesis 1, students in general reported using proportionally more spatial-analytic strategies after instruction than before instruction, F(1, 366) = 360.39, $p < .001$, η_p^2 = .50, concomitant with proportionally fewer spatial-imagistic strategies, F(1, 366) = 40.00, $p <$ .001, η_p^2 = .10. In keeping with Hypothesis 2, there was a trending effect of intervention on the proportionate decrease of spatial-imagistic strategies reported, F(2, 366) = 2.77, p = .06. Planned contrasts show that students in the Analytic Intervention (M = −0.19, SD = 0.33) used relatively fewer spatial-imagistic strategies after instruction than students in the Combined Intervention (M = −0.07, SD = 0.42), F(1, 244) = 5.27, p = .02, η_p^2 = .02. No differences were observed between the Imagistic Intervention (M = −0.13, SD = 0.36) and the Combined Intervention, F(1, 262) = 1.55, p = .21, or the Analytic Intervention, F(1, 226) = 1.17, p = .28. Neither a main effect of sex, F(1, 366) = 0.48, p = .49, nor an interaction between sex and intervention was observed, F(2, 366) = 0.55, p = .58.

Also in keeping with Hypothesis 2, there was a main effect of intervention on the proportionate increase of spatial-analytic strategies reported, F(2, 366) = 3.56, p = .03, η_p^2 = .02. Planned contrasts show that students in the Analytic Intervention (M = 0.18, SD = 0.16) adopted more spatial-analytic strategies than students in the Combined Intervention (M = 0.13, SD = 0.14), F(1, 244) = 6.64, p = .01, η_p^2 = .05. No differences were observed between the Imagistic Intervention (M = 0.17, SD = 0.16) and the Combined Intervention, F(1, 262) = 3.35, p = .07, or the Analytic Intervention, F(1, 226) = 0.47, p = .49. There was also a main effect of sex, F(1, 366) = 4.05, p = .045, η_p^2 = .01; however, no interaction between sex and intervention was observed, F(2, 366) = 0.09, p = .91. Regardless of intervention, women (M = 0.17, SD = 2.3) adopted proportionately more spatial-analytic strategies than men (M = 0.14. SD = 2.2).

These findings are in line with our first hypothesis that disciplinary instruction will increase the tendency of all students (male and female) to rely more on spatial-analytic strategies than spatial-

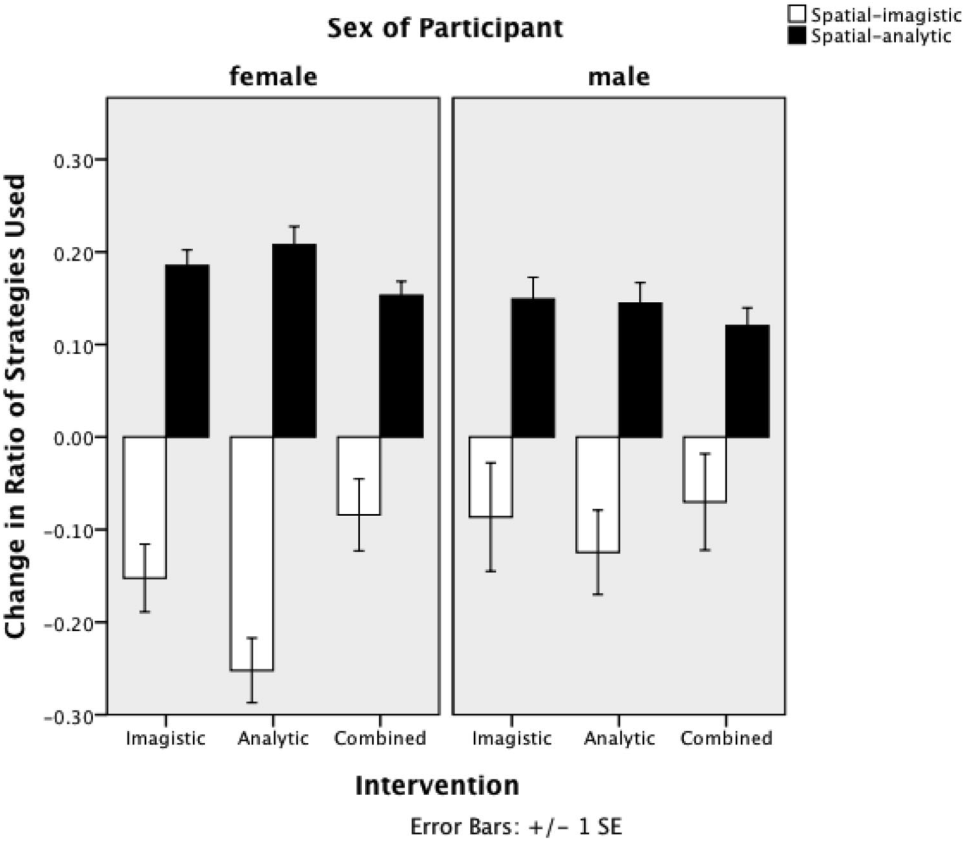

Figure 2. Changes in ratio of spatial-imagistic and spatial-analytic strategies to all strategies reported by men and women in the three interventions. All students reported using relatively more spatial-analytic strategies after instruction. Positive values indicate an increase in the ratio of strategies used. Negative values indicate a decrease in the ratio of strategies used.

imagistic strategies. Although all students increased the use of both spatial-imagistic and spatial-analytic strategies as they replaced guessing strategies, the relative ratio of spatial-imagistic strategies decreased with a concomitant increase in spatial-analytic strategies after instruction in all settings. Any instruction yielded a reduced reliance on spatial-imagistic strategies with an increased reliance on spatial-analytic strategies.

These findings also partially support our second hypothesis that training in spatial-analytical strategies leads students to employ more spatial-analytic strategies. In keeping with this hypothesis, the number of spatial-analytic strategies increased more with analytic training, and this increase was larger in the Analytic Intervention than the Combined Intervention, as predicted. Given that students in the Combined Intervention received instruction in multiple strategies, the ratio of spatial-analytic strategies reported was larger in the Analytic Intervention where only one type of strategy was trained. Although students in the Imagistic Intervention spontaneously adopted some analytic strategies after instruction (despite the absence of instruction to use these strategies), they used spatial-analytic strategies less frequently than the other groups. From these results, we can conclude that instruction in authentic contexts in alternative problem-solving strategies directly influences students to favor the use of taught strategies.

Perhaps more interesting, we observed strong sex differences in the relative frequency of using either spatial-imagistic or spatial-analytic strategies. Regardless of intervention, women reported using spatial-analytic strategies more frequently, and men reported using spatial-imagistic strategies more frequently. Women also reported using proportionally more spatial-analytic strategies than men, suggesting that females may have a bias for using spatial-analytic strategies in isolation, whereas males are more likely to apply both types of strategies. These findings are consistent with prior results indicating reliable sex differences in strategy preferences, and the tendency of males to relate strategies together more reliably than females (Friedman & Miyake, 2000; Heil & Jansen-Osmann, 2008).

Does Combined Training in Analytic Strategies Reduce Sex Differences in Achievement?

We hypothesized that both analytic and combined training would reduce sex differences in achievement (Hypothesis 3). To test this hypothesis, we analyzed student performance on both the CSPSSPA and total achievement points gained on authentic course assessments. First, for the CSPSSPA, pretest achievement did not differ in the three interventions, *F*(2, 366) = 0.43, *p* = .65, or between men and women, *F*(1, 366) = 0.249, *p* = .72, and, not surprisingly, students in all three interventions improved from pretest to posttest, *F*(1, 363) = 1,011, *p* < .001, $\eta_p^2 = .73$. To compare the differences in improvement, we analyzed the gain scores from pre- to postinstruction with a 2 (sex: male, female) × 3 (intervention: imagistic, analytic, combined) repeated-measures ANOVA. Gain scores for male and female students in each intervention are illustrated in Figure 3.

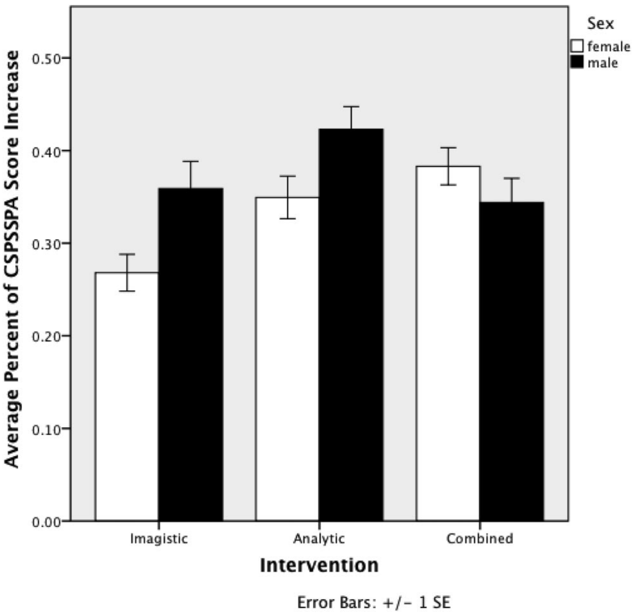

Figure 3. Average gain on the Chemistry Spatial Problem Solving and Strategy Preference Assessment (CSPSSPA) for men and women at posttest in each intervention. Men outperformed women in both the Imagistic and Analytic Interventions, but no differences were observed in the Combined Intervention. Men did not display significant differences in achievement among the interventions; however, women performed best in the Combined Intervention and worst in the Imagistic Intervention.

There was a trending main effect of intervention on achievement gain, *F*(2, 366) = 2.96, *p* = .08, and there was no main effect of sex, *F*(1, 364) = 0.131, *p* = .72. However, there was a significant interaction between intervention and sex, *F*(2, 366) = 5.36, *p* = .005, $\eta_p^2 = .028$. Men outperformed women in the Imagistic and Analytic Interventions, but not in the Combined Intervention. The average gain for men did not differ among the three interventions, *F*(2, 144) = 1.50, *p* = .24. In contrast, the average gain for women did differ as a function of intervention, *F*(2, 222) = 8.31, *p* = .001, $\eta_p^2 = .07$. Planned contrasts show that the Combined Intervention increased female gains (*M* = 0.40, *SD* = 0.20) compared to the Imagistic Intervention (*M* = 0.28, *SD* = 0.22), *F*(1, 165) = 15.4, *p* < .001, $\eta_p^2 = .09$, and the Analytic Intervention (*M* = 0.34, *SD* = 0.21), *F*(1, 141) = 4.02, *p* < .047, $\eta_p^2 = .028$. Although a trend was evident, female gains did not differ significantly between the Imagistic and Analytic Interventions, *F*(1, 138) = 2.77, *p* = .09. The gains for women in the Combined Intervention was about 1/3 of a standard deviation higher than in the Analytic Intervention (*d* = 0.31) and 1 standard deviation higher than in the Imagistic Intervention (*d* = 1.02).

Next, we examined student performance on the total achievement points gained on authentic course assessments using a 2 (sex) × 3 (intervention) ANOVA (see Figure 4). There was a main effect of sex, *F*(1, 350) = 2.05, *p* = .05, $\eta_p^2 = .011$, and a trending effect of intervention, *F*(2, 364) = 2.42, *p* = .09, on achievement. A trend in the interaction between sex and intervention was also observed, *F*(2, 350) = 2.11, *p* = .12. Planned comparisons revealed that male achievement was equivalent across all three interventions, *F*(2, 135) = 0.02, *p* = .97; however, female achievement differed, *F*(2, 215) = 5.41, *p* = .005, $\eta_p^2 = .048$. Women in the Combined Intervention group scored more points on the course assessments (*M* = 497, *SD* = 103) than those in either the Analytic (*M* = 440, *SD* = 126), *F*(1, 138) = 8.15, *p* = .005, $\eta_p^2 = .06$, or Imagistic (*M* = 445, *SD* = 108), *F*(1, 158) = 7.88,

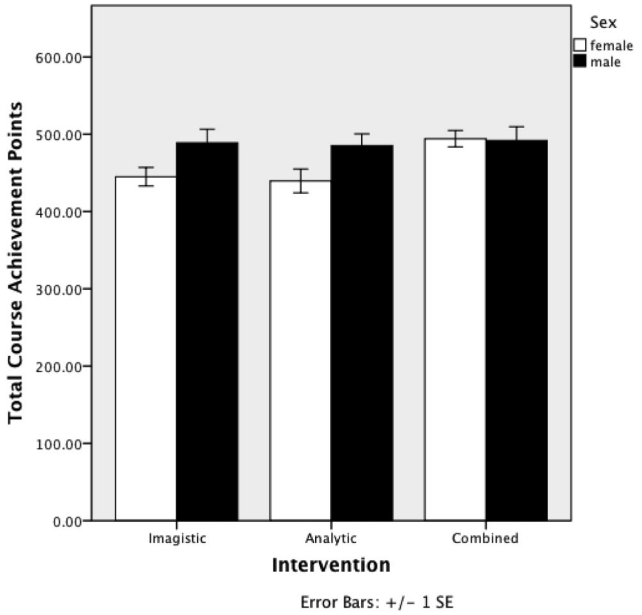

Figure 4. Average total accumulated points on all course assessments administered in each intervention. Male achievement did not differ among interventions, but women in the Combined Intervention outperformed women in the other two interventions.

$p = .006$, $\eta_p^2 = .05$, Intervention. Female achievement did not differ between the Analytic and Imagistic Interventions, $F(1, 132) = 0.17$, $p = .68$. Achievement in the Combined Intervention was about 1/2 of a standard deviation higher than the Imagistic Intervention ($d = 0.47$) and the Analytic Intervention ($d = 0.44$).

Thus, the data partially support our third hypothesis. Combined training in spatial-analytic disciplinary heuristics together with spatial-imagistic selectively benefits females and attenuates sex differences in achievement. Notably, the benefit of analytic instruction was only observed under instructional methods that involved relating spatial-analytic strategies to spatial-imagistic strategies. Instruction and encouragement to use spatial-analytic strategies in isolation did not narrow the sex difference. We observed this effect not only on a researcher-designed achievement measure, but also on authentic achievement measures designed by a course instructor. In contrast, we observed no detectable impact of the interventions on male achievement, which was consistent across all types of instruction.

What Is the Relationship Between Course Achievement, Sex, Spatial Ability, and Strategy Use?

To examine the contributions of sex, spatial ability, and spatial-analytic strategy use to course achievement, we tested a structural equation model of relations among these variables. On the basis of known sex differences in spatial ability and strategy preferences, we proposed that sex predicts both spatial ability and spatial-analytic strategy use, each of which independently predicts achievement. Additionally, we proposed that spatial ability and strategy use fully mediate sex differences in achievement (Hypothesis 4).

Table 3 shows overall correlations among the variables included in the model. The measure of spatial ability was the average of the *z* scores of the Vandenberg Mental Rotation Test, Visualization of Views Test, and Paper Folding Test given moderate correlations between these measures and preliminary analyses, which established that the three scores were collinear in all models. As discussed previously, sex was significantly correlated with spatial ability and the use of spatial-analytic strategies. Spatial ability was positively correlated with course achievement and negatively correlated with the use of spatial-analytic strategies. The use of spatial-analytic strategies was related to course grade, but this relationship was marginally significant.

The model, shown in Figure 5, shows that sex differences in course grade are indeed mediated by both sex differences in spatial ability and sex differences in spatial-analytic strategy use choice. Because sex was coded such that female = 1 and male = 2, positive regression coefficients indicate higher scores for men and negative coefficients indicate higher scores for women for paths originating from the sex variable. As indicated by the regression coefficients, females are more likely to adopt spatial-analytic strategies, which impacts total achievement positively, as does spatial ability. Importantly, observed correlations between sex and course achievement were fully mediated by spatial ability and strategy choice such that the direct path between sex and achievement was nonsignificant after controlling for these variables. The model is consistent with Hypothesis 4 and satisfactorily accounts for the pattern of observed correlations in Table 3, $\chi^2(1) = 0.997$, *ns*; root-mean-square error of approximation = 0.0, comparative fit index = 1.0, standardized root-mean-square residual = .005.

Discussion

This study demonstrates that strategy training is one viable approach to eliminate sex differences in achievement in the STEM discipline of organic chemistry. Importantly, our study demonstrates

Table 3
Correlation Matrix Among Study Variables

Variable	Sex	1	2	3	4	5	6
1. Visualization of Views	$.35^{**}$	—					
2. Mental Rotation	$.36^{**}$	$.55^{**}$	—				
3. Paper Folding	$.08^{\dagger}$	$.41^{**}$	$.49^{**}$	—			
4. Spatial ability composite	$.33^{**}$	$.80^{**}$	$.83^{**}$	$.77^{**}$	—		
5. Number of analytic strategies	$-.22^{**}$	$-.18^{**}$	$-.13^{**}$	$-.10^{*}$	$-.19^{**}$	—	
6. Course achievement	$.11^{*}$	$.18^{**}$	$.18^{**}$	$.23^{**}$	$.26^{**}$	$.07^{\dagger}$	—

$^{\dagger} p = .07$. $^{*} p < .05$. $^{**} p < .01$.

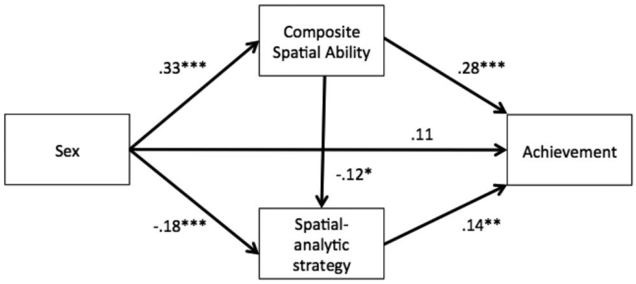

Figure 5. Path diagram illustrating regression coefficients for spatial ability and strategy use. Sex differences in achievement are fully mediated by spatial ability and spatial-analytic strategy use. Spatial ability is higher among men, but spatial-analytic strategy use is higher among women. Spatial ability and spatial-analytic strategy both positively affect achievement. * $p < .05$. ** $p < .01$. *** $p < .001$.

that achievement, at least in organic chemistry, is dependent not only spatial ability but also on strategy choice and that targeted instruction can significantly impact strategy choice. In fact, we observed that sex differences in achievement were fully mediated by both spatial ability and strategy choice. All students improved from pre- to posttest after instruction; however, we found that men achieved higher scores than women in instructional conditions that trained either spatial-imagistic or spatial-analytic strategies in isolation. Men and women did not differ in achievement in the combined intervention that related both types of strategies. Thus, our results show that combined strategy training selectively improves the performance of female students and eliminates a sex difference in achievement.

These findings might help explain some of the variance in the reported correlations between spatial ability and science achievement in previous studies. Failure to account for differences in strategy use may produce correlations that overestimate the predictive validity of spatial ability. Although the effect of strategy use in the present study was modest compared to spatial ability, alternative pedagogies that improve or increase the use of analytic strategies over a sustained period may produce larger effects. The relationships observed here suggest that spatial ability can partially account for sex differences in achievement on spatial problems in STEM disciplines, but the effect of other factors, such as strategy choice, should be included in models that aim to relate sex differences in spatial ability to STEM achievement (Ceci & Williams, 2010; Ceci, Williams, & Barnett, 2009; Halpern, 2007; Spelke, 2005).

It is important to note that sex differences in achievement were eliminated only in our combined intervention. We argue that this observation is due to the fact that effective spatial problem solving in chemistry (and likely other STEM disciplines) involves an interplay between imagistic and analytic approaches as seen among chemistry experts (Stieff & Raje, 2010). Although imagistic methods are often adopted spontaneously by novices, they place high demands on visual–spatial working memory (Shah & Miyake, 1996), such that they may be effective only for those with good spatial abilities. In contrast, analytic strategies are less cognitively demanding, but are more situation-specific, and involve abstraction from the spatial situation described in a problem. However, analytic strategies may be applied effectively only if students first have an appreciation of the full spatial situation, as is emphasized in holistic imagery approaches. Thus, analytic strategies might be most effective when they are taught to students who are already good at applying imagistic strategies or when they are taught in conjunction with imagistic strategies.

We argue that imagistic training produced the usual male advantage in performance, because on average, males have better spatial imagery abilities (Linn & Petersen, 1985; Voyer et al., 1995), so that these strategies are more successful and less effortful for males than for females. In line with this interpretation, men were observed to use more imagistic strategies at posttest regardless of instruction, whereas women were more likely to adopt taught strategies. It is perhaps more surprising that men were also advantaged under purely analytic training. We suggest that this advantage occurred because, in general, men were more likely to begin with well-developed imagistic strategies, to which they could relate the analytic strategies emphasized in this instruction. In contrast, women in this intervention adopted analytic strategies with much greater frequency and may have neglected to apply them in tandem with imagistic strategies.

We propose that combined training eliminated the sex difference found with other forms of training because it directly taught the interplay between imagistic and analytic strategies that we believe is central to success in this STEM domain. This intervention did not just train analytic strategies, but taught students how these more abstract strategies are related to imagistic strategies, which represent the spatial information in a problem in a more holistic way. As we might expect, given evidence from laboratory studies of spatial problem solving (Heil & Jansen-Osmann, 2008; Lawton, 2010; Maccoby & Jacklin, 1974; Robert & Chevrier, 2003), women employed more analytic strategies than men, but more critically, they employed analytic strategies and imagistic strategies more consistently and effectively in the Combined Intervention to attain higher levels of performance. This outcome is consistent with the interpretation that analytic strategies are most effective when they build on or are used in conjunction with imagistic strategies.

It is also important to acknowledge and address some of the limitations of this study. First, because we conducted our interventions in the authentic context of a college chemistry class, the design of our study was necessarily quasi-experimental, with different cohorts of students being assigned to the different interventions. However, we showed that the students in these cohorts were equivalent in terms of both spatial ability and general academic achievement, and we controlled all aspects of the instructional setting, including the instructor, so that we can be relatively confident that our results reflect differences in the interventions rather than spurious factors. Second, our method of assessing students' strategies was based on self-reports. Although self-reports can be unreliable, self-reports of problem-solving strategy in chemistry have been observed to be highly consistent with more objective measures (Stieff, 2007; Stieff, Hegarty, & Deslongchamps, 2011). Third, we contrasted students' strategy reports before and after they received instruction, and one might argue that before instruction students may not have had sufficient knowledge to even attempt the problems. However, in another study, we examined strategies both immediately after students learned the relevant content knowledge and at the end of an organic chemistry class and found the same result as observed here: Students initially adopted primarily imagistic strategies and transitioned to using more analytic strategies by the end of the course (Stieff et al., 2012). Finally, the instructional interventions were relatively short, and we examined problem solving in just one domain. Future studies might investigate whether males and females respond differently to more intensive and lasting interventions than those employed here.

It will be important to examine whether our effects replicate in other STEM domains. The specific spatial-analytic strategies included in the interventions here are clearly applicable to problem solving in organic chemistry only. However, the pedagogic strategies (i.e., analytic, imagistic, and combined training) are transferable to other STEM disciplines. For example, the parity rule documented by Schwartz and Black (1996) is a clear example of a spatial-analytic strategy applicable only to mechanical engineering tasks. To generalize the findings of the present study to another STEM discipline would require a concerted effort to identify the available spatial-analytic strategies and develop an intervention that trains students to use these strategies. We believe that this work can be accomplished through the collaborative work of STEM content experts and cognitive or learning scientists.

It is also unclear whether specific types of strategies are more or less appropriate for different types of organic chemistry problems. In previous work Stieff and colleagues (Stieff, 2011; Stieff & Raje, 2010; Stieff, Ryu, & Dixon, 2010) have shown that variations in problem goals elicit different types of strategies from problem solvers. Unfortunately, the structure of our strategy assessment, which includes only one example of each problem type, does not permit us to determine whether individuals systematically apply the same strategies to solve a problem of a given kind. Moreover, the variance in each participant's strategy choice across tasks does not allow us to determine whether some individuals are better able to apply one strategy over another. Additional research is needed to determine the relative effectiveness of each strategy type for solving a range of problems in the discipline.

It is clear that the ability to reason about spatial relationships (both simple and complex) is critical to success in STEM disciplines, and new programs and interventions to improve STEM achievement are sorely needed. Existing approaches depend on the assumption that spatial ability is a prerequisite for success and suggest either selecting high-spatial students for STEM education or providing courses that train spatial ability directly. The present study suggests an alternative pathway for recruiting and training students in STEM disciplines. Namely, new programs should explore training in multiple strategy use. We have shown that though spatial ability correlates with achievement in organic chemistry, the use of analytic strategies is similarly important. Thus, sex differences in organic chemistry achievement are not intransigent; targeted training in problem solving can reduce or eliminate expected sex differences predicted by disparities in spatial ability. By changing the way we teach STEM courses, we can improve the performance of all students, including those who may be currently underrepresented in the STEM disciplines.

References

Allen, M. J., & Hogeland, R. (1978). Spatial problem-solving strategies as functions of sex. *Perceptual and Motor Skills, 47,* 348–350. doi:10.2466/pms.1978.47.2.348

Bosco, A., Longoni, A. M., & Vecchi, T. (2004). Gender effects in spatial orientation: Cognitive profiles and mental strategies. *Applied Cognitive Psychology, 18,* 519–532. doi:10.1002/acp.1000

Ceci, S. J., & Williams, W. M. (2010). *The mathematics of sex: How biology and society conspire to limit talented women and girls*. New York, NY: Oxford University Press.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135,* 218–261. doi:10.1037/a0014412

Dabbs, J. M., Chang, E., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Human Behavior, 19,* 89–98. doi:10.1016/S1090-5138(97)00107-4

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

Freedman, R. J., & Rovegno, L. (1981). Ocular dominance, cognitive strategy, and sex differences in spatial ability. *Perceptual and Motor Skills, 52,* 651–654. doi:10.2466/pms.1981.52.2.651

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General, 129,* 61–83. doi:10.1037/0096-3445.129.1.61

Gilbert, J. K. (Ed.). (2005). *Visualization in science education*. Dordrecht, the Netherlands: Springer. doi:10.1007/1-4020-3613-2

Guay, R. B., & McDaniel, E. D. (1976). *The visualization of viewpoints*. West Lafayette, IN: Purdue Research Foundation.

Halpern, D. F. (2007). Science, sex, and good sense: Why women are underrepresented in some areas of science and math. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 121–130). Washington, DC: American Psychological Association. doi:10.1037/11546-010

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8,* 1–51. doi:10.1111/j.1529-1006.2007.00032.x

Harle, M., & Towns, M. (2011). A review of spatial ability literature, its connection to chemistry, and implications for instruction. *Journal of Chemical Education, 88,* 351–360. doi:10.1021/ed900003n

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences, 8,* 280–285. doi:10.1016/j.tics.2004.04.001

Hegarty, M. (2010). Components of spatial intelligence. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 52, pp. 265–297). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(10)52007-3

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91,* 684–689. doi:10.1037/0022-0663.91.4.684

Hegarty, M., Stieff, M., & Dixon, B. L. (2013). Cognitive change in mental models with experience in the domain of organic chemistry. *Journal of Cognitive Psychology, 25,* 220–228. doi:10.1080/20445911.2012.725044

Hegarty, M., & Waller, D. A. (2006). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121–169). Cambridge, England: Cambridge University Press.

Heil, M., & Jansen-Osmann, P. (2008). Sex differences in mental rotation with polygons of different complexity: Do men utilize holistic processes whereas women prefer piecemeal ones? *Quarterly Journal of Experimental Psychology, 61,* 683–689. doi:10.1080/17470210701822967

Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education, 86,* 151–158. doi:10.1002/j.2168-9830.1997.tb00278.x

Hugdahl, K., Thomsen, T., & Ersland, L. (2006). Sex differences in visuo-spatial processing: An fMRI study of mental rotation. *Neuropsychologia, 44,* 1575–1583. doi:10.1016/j.neuropsychologia.2006.01.026

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. doi:10.1037/0021-9010.78.2.250

Jordan, K., Wustenberg, T., Heinze, H.-J., Peters, M., & Jäncke, L. (2002). Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia, 40,* 2397–2408. doi:10.1016/S0028-3932(02)00076-3

Just, M. A., & Carpenter, M. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92,* 137–172. doi:10.1037/0033-295X.92.2.137

Kyllonen, P. C., Lohman, D. F., & Snow, R. E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology, 76,* 130–145. doi:10.1037/0022-0663.76.1.130

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles, 30,* 765–779. doi:10.1007/BF01544230

Lawton, C. A. (2010). Gender, spatial abilities, and wayfinding. In J. C. Chrisler & D. R. McCreary (Eds.), *Handbook of gender research in psychology: Vol. 1. Gender research in general and experimental psychology* (pp. 317–341). New York, NY: Springer. doi:10.1007/978-1-4419-1465-1_16

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: A meta-analysis. *Child Development, 56,* 1479–1498. doi:10.2307/1130467

Lowrie, T., & Kay, R. (2001). Relationship between visual and nonvisual solution methods and difficulty in elementary mathematics. *Journal of Educational Research, 94,* 248–255. doi:10.1080/00220670109598758

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences, 49,* 344–351. doi:10.1016/j.paid.2010.03.022

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

Mayer, R. E., Stull, A. T., Campbell, J., Almeroth, K., Bimber, B., Chun, D., & Knight, A. (2007). Overestimation bias in self-reported SAT scores. *Educational Psychology Review, 19,* 443–454. doi:10.1007/s10648-006-9034-z

Miller, D. I., & Halpern, D. F. (2013). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences, 26,* 141–152. doi:10.1016/j.lindif.2012.03.012

National Research Council. (2006). *Learning to think spatially*. Washington, DC: National Academies Press.

Peters, M., Chisholm, P., & Laeng, B. (1995). Spatial ability, student gender, and academic performance. *Journal of Engineering Education, 84,* 69–73. doi:10.1002/j.2168-9830.1995.tb00148.x

Pezaris, E., & Casey, M. B. (1991). Girls who use "masculine" problem-solving strategies on a spatial task: Proposed genetic and environmental factors. *Brain and Cognition, 17,* 1–22. doi:10.1016/0278-2626(91)90062-D

Robert, M., & Chevrier, E. (2003). Does men's advantage in mental rotation persist when real three-dimensional objects are either seen or felt? *Memory & Cognition, 31,* 1136–1145. doi:10.3758/BF03196134

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science, 20,* 457–497. doi:10.1207/s15516709cog2004_1

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125,* 4–27. doi:10.1037/0096-3445.125.1.4

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology, 93,* 604–614. doi:10.1037/0022-0663.93.3.604

Sorby, S. A. (2001). A course in spatial visualization and its impact on the retention of female engineering students. *Journal of Women and Minorities in Science and Engineering, 7,* 153–172.

Sorby, S. A. (2009). Education research in developing 3-D spatial skills for engineering students. *International Journal of Science Education, 31,* 459–480. doi:10.1080/09500690802595839

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science: A critical review. *American Psychologist, 60,* 950–958. doi:10.1037/0003-066X.60.9.950

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction, 17,* 219–234. doi:10.1016/j.learninstruc.2007.01.012

Stieff, M. (2011). When is a molecule three-dimensional? A task-specific role for imagistic reasoning in advanced chemistry. *Science Education, 95,* 310–336. doi:10.1002/sce.20427

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction, 29,* 123–145. doi:10.1080/07370008.2010.507318

Stieff, M., Hegarty, M., & Dixon, B. (2010). Alternative strategies for spatial reasoning with diagrams. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 115–127). Berlin, Germany: Springer. doi:10.1007/978-3-642-14600-8_13

Stieff, M., & Raje, S. (2008). Expertise and spatial reasoning in advanced scientific problem solving. In P. A. Kirschner, J. J. G. van Merriënboer, & T. de Jong (Eds.), *Proceedings of the Eighth International Conference of the Learning Sciences* (Vol. 2, pp. 366–373). Mahwah, NJ: Erlbaum.

Stieff, M., & Raje, S. (2010). Expert algorithmic and imagistic problem solving strategies in advanced chemistry. *Spatial Cognition and Computation, 10,* 53–81. doi:10.1080/13875860903453332

Stieff, M., Ryu, M., & Dixon, B. L. (2010). Students' use of multiple strategies for scientific problem solving. In S. R. Goldman, J. Pellegrino, K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Proceedings of the Ninth International Conference of the Learning Sciences* (Vol. 1, pp. 765–772). Mahwah, NJ: Erlbaum.

Stieff, M., Ryu, M., Dixon, B. L., & Hegarty, M. (2012). The role of spatial ability and strategy preference for spatial problem solving in organic chemistry. *Journal of Chemical Education, 89,* 854–859. doi:10.1021/ed200071d

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139,* 352–402. doi:10.1037/a0028446

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47,* 599–604. doi:10.2466/pms.1978.47.2.599

Voyer, D., (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: A meta analysis. *Psychonomic Bulletin & Review, 18,* 267–277. doi:10.3758/s13423-010-0042-0

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. doi:10.1037/0033-2909.117.2.250

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101,* 817–835. doi:10.1037/a0016127

Weiss, E., Siedentopf, C. M., Hofer, A., Diesenhammer, E. A., Hoptman, M. J., Kremser, C., . . . Delazer, M. (2003). Sex differences in brain activation pattern during a visuospatial cognitive task: A functional magnetic resonance imaging study in healthy volunteers. *Neuroscience Letters, 344,* 169–172. doi:10.1016/S0304-3940(03)00406-3

Received August 23, 2012
Revision received July 8, 2013
Accepted July 15, 2013 ■