Journal of Science Education and Technology, Vol. 15, No. 1, March 2006 (© 2006)
DOI: 10.1007/s10956-006-0361-0

Real-Time Data Display, Spatial Visualization Ability, and Learning Force and Motion Concepts

Maria Kozhevnikov[1,3] and Ronald Thornton[2]

In this study, we examined how students' levels of spatial visualization ability interact with learning physics in a microcomputer-based laboratory (MBL) environment. Undergraduate students who had taken an introductory physics course based on MBL tools were pre- and posttested at the beginning and at the end of the semester on spatial visualization ability and their conceptual understanding of mechanics. The results showed that while spatial visualization is a reliable predictor for students' performance on physics conceptual evaluation tests before MBL instruction, the relation is not significant after the instruction. Furthermore, as a result of MBL instruction, students' levels of spatial visualization increased significantly. In addition, a group of science teachers presented with different types of MBL activities also showed a significant increase in spatial visualization ability.

KEY WORDS:

INTRODUCTION

Educational research has consistently revealed poor qualitative understanding of physics among introductory physics students (e.g., Champagne *et al.*, 1980; Hestenes, 1995; Hestenes *et al.*, 1992; Thornton and Sokoloff, 1990). Numerous studies in physics education repeatedly pointed out that students did not exhibit any significant improvement on physics conceptual evaluation tests after traditional lecture-based instruction (e.g., Cummings *et al.*, 1999; Halloun and Hestenes, 1985; Hestenes, 1995; Hestenes *et al.*, 1992; Sokoloff and Thornton, 1997; Thornton, 1999a,b). In the area of mechanics, for instance, students exhibit a consistent set of difficulties interpreting kinematics graphs (see Beichner, 1994; McDermott *et al.*, 1987) and misunderstanding force and motion concepts (e.g., Halloun and Hestenes, 1985; Thornton, 1997; Trowbridge and McDermott, 1981).

To address students' difficulties in learning physics concepts, a wide variety of instructional techniques have been developed, many of them based on the use of computer technologies (e.g., Beichner, 1996; Rosenquist and McDermott, 1987; Thornton and Sokoloff, 1990; White, 1993). One of the promising techniques to promote students' conceptual understanding of physics concepts has been the use of microcomputer-based labs (MBLs) for real-time experimental data graphing (e.g., Laws, 1997; Mokros and Tinker, 1987; Sokoloff and Thornton, 1997). MBL tools involve probes interfaced with a computer and are used to gather and graphically display data from science experiments. One example of MBL-based curricula is Workshop Physics (Laws, 1997), in which formal lectures were eliminated and replaced by experimental activities based on MBL tools to record, display, and analyze the data. Another example is RealTime Physics (Thornton and Sokoloff, 1992), an active discovery-based curriculum, in which microcomputer-based interactive demonstrations are integrated in the lecture/laboratory structure. With these tools and curricula, it has been possible to bring about sig-

[1]Department of Psychology, Rutgers University, 101 Warren St., Smith Hall, Newark, New Jersey.
[2]Center for Science and Mathematics Teaching, Tufts University, Medford, Massachusetts.
[3]Currently, Maria Kozhevnikov is working at National Science Foundation; e-mail: mkozhevn@nsf.gov

nificant changes in students' conceptual understanding of mechanics concepts (Sokoloff and Thornton, 1997).

Researchers speculated that the success of MBL activities in promoting students' conceptual change is due to the heavy emphasis MBL curricula place on visual–spatial representations (e.g., Linn *et al.*, 1987; Mokros and Tinker, 1987). Students use graphical representations as a central means of communications within the context of student-controlled experiments. They are asked to predict and discuss results in terms of graphs, and if there are any inconsistencies between the graphs of the observation and the predictions, the students must realize this and correct their experimental procedure or their predictions. The focus of the current study is to investigate how spatial visualization ability (i.e., the ability to understand, manipulate, and interpret visual/spatial representations) interacts with learning physics in microcomputer-based laboratories involving real-time graphing. We suggest that the strong emphasis MBL curricula place on visual/spatial representations has the potential not only to facilitate students' understanding of physics concepts, but also to enhance their spatial visualization skills.

Spatial Ability and Learning Physics

A modern view of working memory (Baddeley, 1992) has proposed that working memory consists of separate processors for verbal and visual/spatial information. The verbal part of working memory (called the phonological loop) specializes in processing verbal information, while the visual part of working memory (called the visuo-spatial sketchpad) specializes in processing visual/spatial information. According to this model, visual/spatial processes such as generating, manipulating, or interpreting visual/spatial images take place in the visuo-spatial sketchpad. A basic assumption of this theory is that both verbal and visual/spatial processors have a limited processing capacity, and that proper allocation of cognitive resources is critical to learning (Mousavi *et al.*, 1995). Experimental research has suggested that spatial ability tests reflect visual–spatial working memory capacity, and that people who differ in spatial abilities also differ in performance on tasks that involve visual/spatial processing (e.g., Carpenter *et al.*, 1999; Just and Carpenter, 1985). Spatial ability tests usually involve judgments about the identity of a pair of stimuli presented at different angles (speeded mental rotation tasks) or more complex spatial visualization tasks that require the subjects to generate, maintain, and coordinate information during spatial transformations. For instance, for one of the spatial visualization tests—paper folding test—students view an illustration of a piece of paper being folded and then having a hole punched through it. Students then attempt to pick the correct unfolded configuration (Ekstrom *et al.*, 1976). Researchers (e.g., Salthouse *et al.*, 1990; Shah and Miyake, 1996) proposed that individual differences on spatial visualization tests arise because of the simultaneous processing (i.e., mentally folding the paper) and storage demands (i.e., maintaining in memory the partially folded representation) that tax the supply of visual/spatial working memory resources.

Spatial visualization ability tests, like the paper folding test, have been shown to be reliable predictors of student achievement in a wide range of technical and engineering subjects (e.g., McGee, 1979). There are also reports that physicists have uniformly high spatial but not verbal ability (Dictionary of Occupational Title, 1991; Roe, 1953), indicating that the spatial, but not the verbal, component of intelligence is especially important for solving physics problems. Also, educational studies have found that science students (physics majors in particular) possess more highly developed spatial ability skills than do nonscience students (Lord and Nicely, 1997; Siemankowski and MacKnight, 1971). Since problem solving in physics often involves visualizing complex spatial processes and mentally manipulating graphs and diagrams, it is not surprising that low spatial ability students have been shown to experience difficulties in learning physics, in particular while solving problems that require visualizing abstract physics concepts and/or interpreting graphs (Kozhevnikov *et al.*, 2002; Pallrand and Seeber, 1984). There is also evidence that as kinematics events become more complex, competing visual/spatial processing demands are more likely to arise. For instance, Isaak and Just (1995) found that students' susceptibility to incorrect judgments about rolling motion is related to their spatial visualization ability. Isaac and Just proposed that the simultaneous processing of the rotation and translation components of motion overloaded available visual–spatial working memory.

The above findings raise questions about the possibilities and ways to train and enhance spatial skills. Although there have been numerous attempts to improve spatial ability by having students perform numerous mental rotation tasks, most training efforts have been unsuccessful (see Lohman, 1979, 1988, for a review). The basic problem is that while training has improved performance on the specific spatial task (e.g., rotating letters), generalization to a new and even very similar task (e.g., rotating cards) has been limited. For example, Bethell-Fox and Shepard (1988) found that practice with feedback in mentally rotating particular stimuli did not transfer to new stimuli, and that with continued practice on particular stimuli, rotation time became independent of the complexity of the stimuli for most subjects. On the basis of this result, Bethell-Fox and Shepard proposed that students might represent a stimulus and imagine its rotation piece by piece when the stimulus is unfamiliar, but represent a stimulus as a whole if it is sufficiently well learned. Although this switch to holistic representations is not an improvement of mental rotation ability itself, it may account for much of the improvement shown on spatial skills tests. Furthermore, Lohman and Nichols (1990) showed that practice in solving three-dimensional mental rotation problems can lead to significant gains on two-dimensional rotation tests (e.g., card and figure rotation tests) and spatial visualization tasks (e.g., paper folding). However, it appeared that along with the increased number of correctly solved items on spatial ability tests, the number of incorrectly solved items increased significantly, so that gains observed were due to increases in speed but not due to reductions of errors.

Although studies on directly training spatial abilities have not been particularly successful, there are several studies that found significant increases in students' performance on spatial visualization tests after taking geometry or science courses (e.g., Lord, 1985; Lord and Holland, 1997; Lord and Rupert, 1995; Pallrand and Seeber, 1984). Additionally, it has been found that integrating computer graphics with teaching mathematics, engineering, or science improves students' performance on spatial tests (e.g., Barnea and Dori, 1999). These data suggest that spatial abilities can be trained by presenting students with a variety of visual/spatial stimuli, such as scientific graphs and diagrams. The problem with the above studies, however, is that they did not report a pre–post change in the number of incorrectly solved items, and never considered the possibility mentioned by Lohman and Nichols (1990) that the overall gains observed could be the results of increases in speed of solving spatial tasks and not due to a reduction of errors.

In Study 1, we examined how students' levels of spatial visualization ability affect students' performance on physics conceptual evaluation tests, and specifically students' performance on different types of physics problems, before and after MBL instruction. In Study 2, we examined in detail how the amount of MBL instruction and students' active engagement with MBL tools while learning physics concepts affects students' spatial visualization skills. In Study 3, we investigated whether participants with a prior physics background would improve their spatial visualization ability as a result of using MBL tools.

STUDY 1

The group of participants consisted of 76 undergraduate students who took a general noncalculus physics course at Tufts University. The students were exposed to MBL instruction during the introductory physics lecture class, which included a series of four 40-min lectures on the topics of kinematics and dynamics. These lectures were taught using the *Interactive Lecture Demonstrations* (ILD) method supported by real-time MBL tools (Sokoloff and Thornton, 1997; Thornton and Sokoloff, 1990, 1992). In addition, the students had two 2-h kinematics laboratories where they worked in small groups in an MBL environment.

Interactive Lecture Demonstration

Table I outlines four sequences of ILDs in mechanics that were used during the course lectures. In an ILD session, students were given a "prediction sheet" with space to write predictions and answer questions. Students engaged in small-group discussions with their nearest neighbors and then recorded their final predictions. An instructor elicited common student predictions from the whole class, and then carried out the demonstration with MBL measurements (e.g., the graph of velocity, acceleration, or position vs. time) displayed. Students were also given an essentially identical "results sheet" that they completed.

As an example, the Newton's First and Second Law ILD sequence is shown in Fig. 1. These short descriptions are taken from the ILD teacher materials. An excerpt from the student Prediction Sheet showing the first few predictions for this sequence is

Table I. Mechanics Interactive Lecture Demonstration Sequences

ILD sequence	Contents
Kinematics 1: Motion with carts	Kinematics of constant velocity and uniformly accelerated motion using a motion detector to display motion of a low friction cart pushed along by a fan unit. Relationship between velocity and acceleration
Kinematics 2: First and second laws	Dynamics using a force probe and motion detector to measure forces applied to low- and high-friction carts, and the resulting velocity and acceleration. Relationship among velocity, acceleration, and force
Kinematics 3: Newton's third law	Using two force probes allows students to examine interaction forces between two objects during fast collision and when one object is in constant contact with another, pushing or pulling
Kinematics 4: Energy of a cart on a ramp	Using a force probe and motion detector to measure kinetic and potential energy of a cart during its motion up and down an incline. Total mechanical energy, kinetic and potential energy and their relationship

shown in Fig. 2. The graphs of a typical set of this sequence (specifically of demonstration 6 presented in Fig. 1) as displayed on a computer screen are shown in Fig. 3. A force probe mounted on the low-friction cart measured the force applied to the cart by a weight attached to a string hung over a pulley (a modified Atwood's machine, see Fig. 1), while a motion detector measured velocity and acceleration. The cart was given a quick push opposite to the force exerted by the hanging weight, and it moved toward the motion detector, slowed down, and returned. The shaded portions of the graphs show the time interval when the cart was moving under the influence of a constant force.

Materials

The materials consisted of tests measuring students' spatial visualization ability levels and their conceptual understanding of mechanics.

Force and Motion Conceptual Evaluation Test (FMCE)

FMCE is a multiple-choice mechanics questionnaire, developed at the Center for Science and Mathematics Teaching at Tufts University, used to evaluate students' conceptual understanding of force and motion concepts (Thornton and Sokoloff, 1992, 1998). The questions for FMCE were developed from the results of science education research in order to identify students' alternative views and models about the most basic concepts of kinematics and dynamics. The test was developed on the basis of extensive classroom and laboratory observation and validated in a large number of studies (Sokoloff and Thornton, 1997; Thornton, 1991, 1992, 1993, 1996, 1999a,b; Thornton and Sokoloff, 1990, 1998).

In this study, we used 32 questions from the original version of FMCE. The questionnaire is presented in Appendix A. The questions were of the following types:

The *Force Sled* questions (Questions 1–7) referred to a sled on ice (negligible friction) pushed by someone. Different motions of the sled were described and students were asked to select the force that could cause each motion. The Force Sled questions concerned only motion along a horizontal line (one-dimensional motion with no gravitational effects). These questions explicitly displayed the source of the forces and were presented in a natural language that required no explicit knowledge of coordinate systems and no knowledge of graphs.

The *CoinToss* questions (Questions 8–10) examined students' understanding of Newton's first two laws. These questions referred to a coin tossed in the air, and asked students to select, among seven choices, the correct description of the forces acting on the coin (1) as it moved upward, (2) when it reached its highest point, and (3) as it moved downward.

Unlike the Force Sled questions, both the *Force Graph* questions (Questions 11–18) and the *Acceleration Graph* questions (Questions 19–23) used graphical representations. The Force Graph questions asked students to pick the appropriate force–time graph (from nine choices) to describe the force that could cause a toy car to move in various ways on

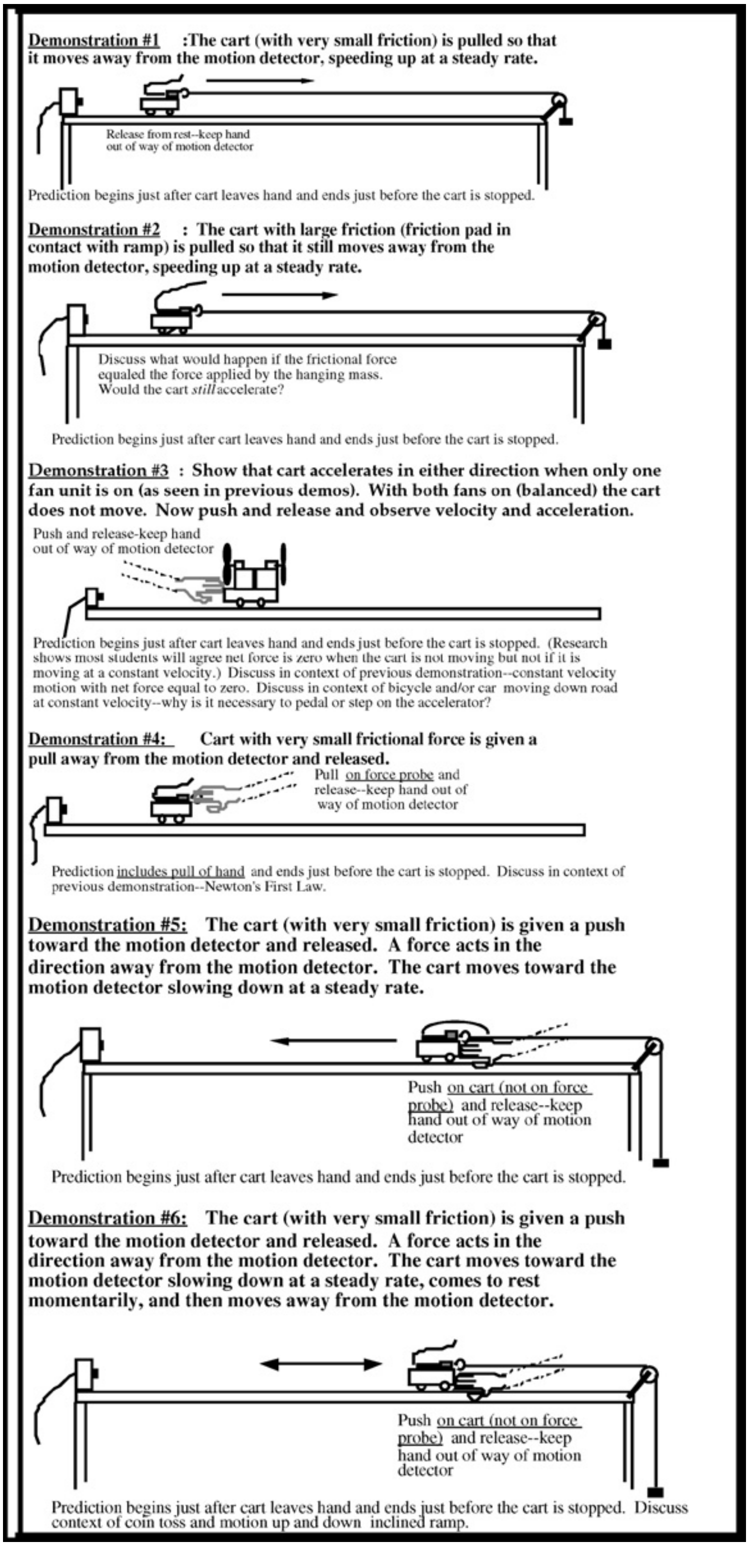


Fig. 1. Newton's first and second Law Interactive Lecture Demonstration sequence. The short descriptions of demonstrations are taken from ILD teacher materials.

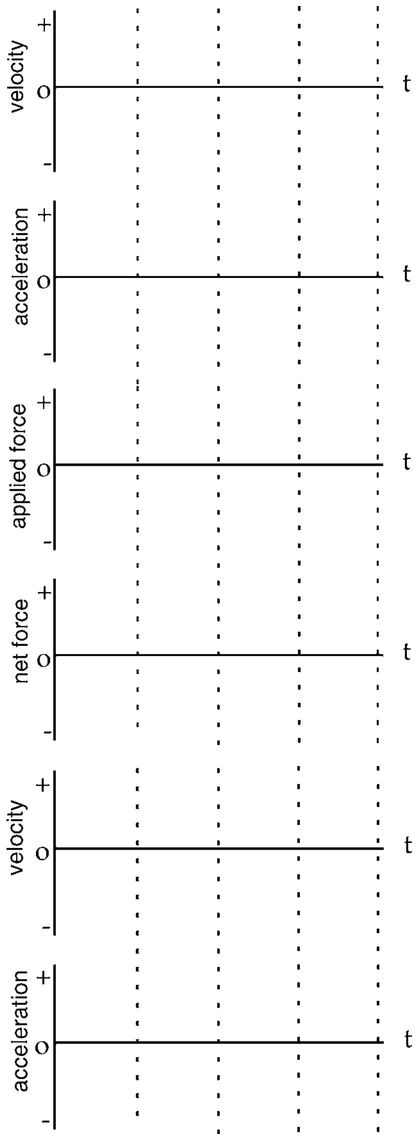


Fig. 2. First part of student prediction sheet for the ILD sequence on Newton's first and second laws. This sheet is collected, and students get credit if it is filled out.

a horizontal surface. The Acceleration Graph questions were similar to the Force Graph questions, except instead of a force graph, students were asked to choose the acceleration graph (from eight choices) that corresponded to a given description.

The *Collisions* questions (Questions 24–32) examined students' understanding of Newton's third law. These questions asked students to evaluate the forces that two colliding objects exerted on one another in different situations (different objects' masses, different initial velocities). Students were asked to choose the correct answer that best described the forces between the colliding objects.

Students were given 40 min to complete this evaluation test. The internal reliability of the FMCE is 0.86 (α-Cronbach).

Spatial Visualization Test

The participants were administered the *Paper Folding Test* (PFT) (Ekstrom *et al.*, 1976). Other analyses (reported by Kozhevnikov *et al.*, 2002) indicated that spatial visualization tests (e.g., PFT, Form Board Test) predict students' performance on physics conceptual evaluation tests, while speeded mental rotation tests (Card Rotation Test, Cube Comparison Test) do not. For this reason, we administered to the subjects only the PFT, one of the most commonly used paper-and-pencil spatial visualization tests. Each item of the PFT shows successive drawings of two or three folds made in a square sheet of paper. The final drawing shows a hole being punched in the folded paper (see an example of the task in Appendix B). The students were asked to select one of the five drawings to show how the punched sheet would appear when fully opened. The PFT consisted of 10 items, and participants were given 3 min to complete the task. The internal reliability of the Paper Folding Test is 0.84.

The total score for the PFT was calculated using a correction for guessing, according to the formula $R - W/(n-1)$ where R is the number correct items, W the number incorrect items, and n the number of response options for each item ($n = 5$ for the PFT).

Procedure

All students were tested at the beginning and at the end of the course on the Paper Folding Test. The FMCE was administered at the beginning and at the end of the semester.[4]

[4]No control group (that received traditional instruction or no instruction at all) was used in Study 1 for the following reason: Numerous studies in physics and science education conducted on a large number of high-school and college students (including also Tufts undergraduate population) consistently pointed out that students did not exhibit any significant improvement on

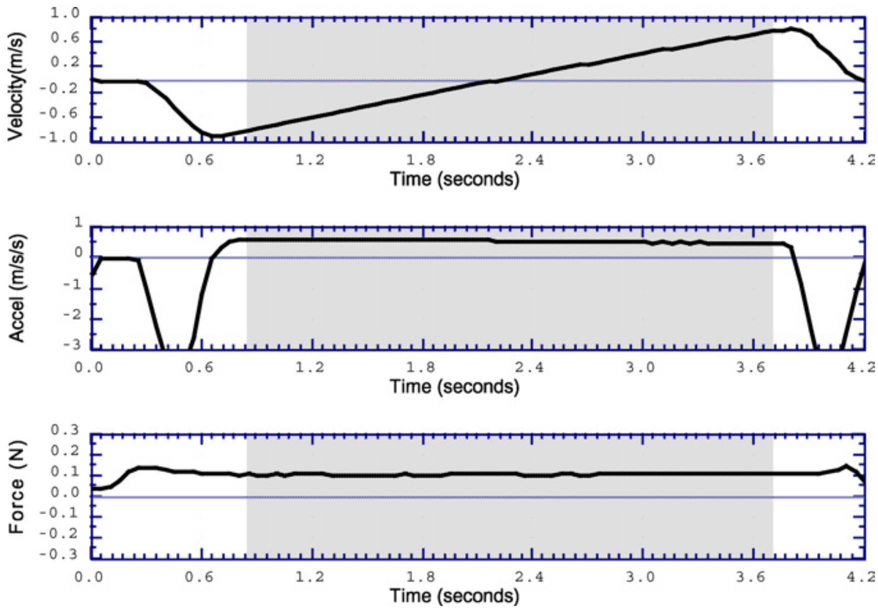


Fig. 3. Typical results from Demonstration 6 of the Newton's first and second Law Sequence (Fig. 1). These are actual data from an ILD given at Tufts.

Results

Paired-sampled t test showed a significant increase in the participants' scores on post-FMCE in comparison with pre-FMCE: ($t(74) = 23.08$, $p < 0.001$) with a posttest mean score on FMCE of 22.61 (SD = 5.01) and a pretest mean score on FMCE of 7.95 (SD = 4.82). The data in Fig. 4 show the improvement of students' scores on each type of FMCE question after MBL instruction. Students showed significant gains for FMCE questions, and these gains were similar to the increases in students' conceptual understandings reported in other studies that used MBL instructions (Sokoloff and Thornton, 1997; Thornton, 1999a; Thornton and Sokoloff, 1990, 1998). The increases were much higher than the increases in students' scores after traditional instruction. For instance, Thornton (1999a,b) and Thornton and Sokoloff (1990) compared the effect of MBL versus traditional instruction and showed a less than 10% gain for FMCE questions when students experienced traditional lecture-based instruction.

physics conceptual evaluation tests, including FMCE, after traditional lecture-based instruction (e.g., Cummings *et al.*, 1999; Halloun and Hestenes, 1985; Hestenes, 1995; Hestenes *et al.*, 1992; Sokoloff and Thornton, 1997; Thornton, 1991, 1992, 1993, 1996, 1999a,b; Thornton and Sokoloff, 1990, 1998).

To examine if students' initial level of spatial visualization ability predicted their performance on FMCE before physics instruction and if the relation continued to be significant after the instruction, we performed the linear regression analysis of pre- and post-FMCE scores on pretest PFT. The linear regression with the pretest PFT as an independent variable and pre-FMCE scores as a dependent variable was reliable: $F(1, 75) = 8.14$, $p < 0.005$. The unstandardized regression coefficient was significant ($B = 0.11$, SD = 0.03, $p < 0.005$). In contrast, the linear regression analysis with the PFT (pretest) as an independent variable and post-FMCE scores as a dependent variable show no significant relation between the two: $F(1, 75) = 1.43$, $p = 0.23$, and the unstandardized regression coefficient was not significant: ($B = 0.05$, SD = 0.04). The results suggest that while spatial visualization was a reliable predictor for students' performance on pre-FMCE, the relation between spatial visualization and FMCE was not significant after the MBL instruction.

Kozhevnikov *et al.* (2002) showed that not all the problems from physics conceptual evaluation tests correlate with spatial visualization ability but mainly those problems that describe complex two-dimensional motion. Thus, to explore more thoroughly the relationships between different types of mechanics questions and spatial visualization abil-

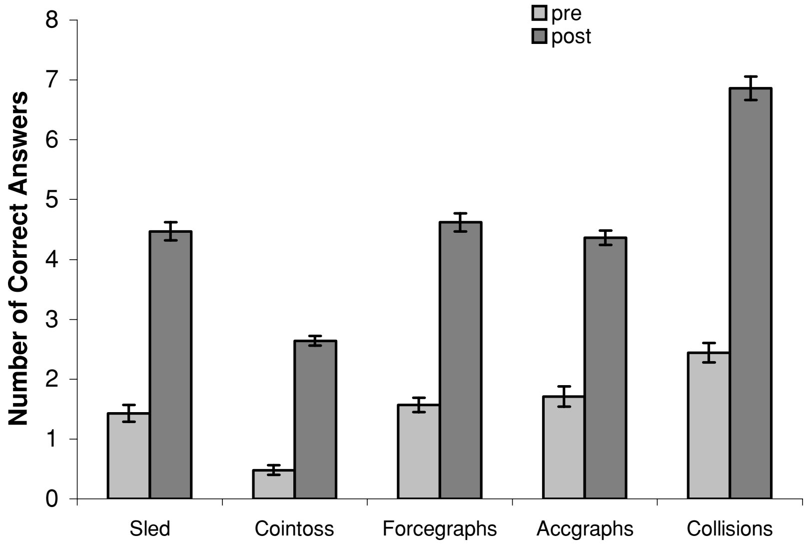


Fig. 4. Performance on the FMCE pretest and posttest by types of questions. Error bars are standard errors.

ity, correlational analysis was performed between the pretest PFT and different types of mechanics problems for pre- and post-FMCE measures. The results of the correlation analysis are presented in Table II. On the pretest, both Force Graph and Acceleration Graph questions, as well as Collisions questions, correlated significantly with spatial visualization ability ($p < 0.01$). On the posttest, pretest PFT correlated significantly only with Acceleration Graph questions ($p < 0.01$). The significant correlations of Force Graph, Acceleration Graph, and Collisions questions with spatial visualization suggest that individual differences in solving these questions before physics instructions were related to the differences in students' spatial visualization ability. Both Force Graph and Acceleration Graph problems required a certain amount of visual–spatial processing to interpret the visual information given in the graph. When no "schema" is available that automatically associates graph features with quantitative relationship (e.g., a curved line indicates accelerating relationship), low-spatial students appear to experience more difficulties than do high-spatial students. Collision problems required students to visualize the action of different forces acting on multiple objects. When no knowledge of Newton's third law is available, spatial visualization ability seems to play an important role in solving this type of problem. In contrast, the Force Sled and Coin Toss questions do not correlate with spatial visualization. In these types of problems, the student was presented with one-dimensional one-body motion that does not seem to require any complex spatial transformations.

As can be seen from Table II, the only type of FMCE question that still correlates significantly with the pretest PFT was Acceleration Graph problems.

Table II. Correlations Between Paper Folding Test and Mechanics Problems in Pre- and Post-FMCE Tests

Task	Correlations between pre-PFT with FMCE pre-test measures	Correlations between pre-PFT with FMCE post-test measures
FMCE (total score)	0.32**	0.13
Force on the sled	0.13	0.09
Coin toss	−0.11	0.18
Force graphs	0.28**	0.01
Acceleration graphs	0.29**	0.28**
Collisions	0.28**	−0.02

**$p < 0.01$.

Acceleration Graph Problems

What is so unique about Acceleration Graph problems in comparison with other types of FMCE questions? To understand why the Acceleration Graph problems still correlated with spatial ability after MBL instructions, we examined characteristic responses of students of different spatial abilities on these types of questions before and after instruction.

Consider, for instance, Question 19. In this question, students were asked to choose an acceleration graph (from eight choices) that corresponds to the following description: "The car moves toward the right, speeding up at a steady rate." Since "steady rate" implies constant acceleration, and direction to the right is defined as positive, the correct graph is the graph that represents the car's acceleration as a horizontal line above the axis of time (choice A). An analysis of students' responses showed that, for Question 19, the characteristic incorrect answer chosen by more that 50% of the students in the pretest was choice E, which represents acceleration as a line going up along the axis of time. Choice E may be considered as "literal" graph misinterpretation, in which students expect the graph to be a picture of the phenomena (that is, they expect the line graph to go up, since the car is speeding up), independent of the graph ordinate.

It appeared that such characteristic literal interpretations were also common for Questions 20 and 22: more than 50% of the students chose these types of literal interpretations in the pretest. Questions 21 and 23, in contrast, did not have explicit literal interpretations; they described the motion at a constant velocity that implies constant (zero) acceleration. In this case, the correct choice (the graph representing a zero horizontal line) and the incorrect characteristic choice (the graph representing a positive horizontal line) had similar horizontal shapes. For these questions, incorrect characteristic responses (i.e., the positive horizontal line) were given less often (less than 50% of all the answers).

We would also like to note that literal graph misinterpretation is a common error made by many students while interpreting kinematics graphs and it has been repeatedly reported in the educational literature (e.g., Beichner, 1994; McDermott *et al.*, 1987; Mokros and Tinker, 1987). The question is whether any relationship exists between students' susceptibility to these literal misinterpretations and their spatial visualization ability level. To examine this question, we compared the number of low- and high-spatial students who gave literal and correct answers for each Acceleration Graph question on pre- and post-FMCE tests (see Table III). On the basis of a median split on performance on the PFT, we classified students as having either low- or high-spatial ability. The low-spatial students ($N = 37$) were those who had scores less than 6.5 (out of a maximum 10) on the PFT. The high spatial students had scores equal or greater than 6.5 ($N = 39$). As seen in Table III, before instruction students who gave the literal misinterpretations on Acceleration graph questions were mostly low-spatial students. Chi-square analysis showed a significant relationship between students' level of spatial ability and the choice of their answer (literal vs. correct interpretation) for Question 19 ($\chi^2(1) = 3.55, p < 0.05$), Question 20 ($\chi^2(1) = 3.61, p < 0.05$) and a marginally significant relationship for Question 22 ($\chi^2(1) = 2.63,$

Table III. Number of Characteristics Incorrect and Correct Answers for Acceleration Graph Problems for Low- and High-Spatial Students on Pre- and Post-tests

	Pretest		Posttest	
Spatial ability	Characteristics incorrect answers	Correct answers	Characteristics incorrect answers	Correct answers
Question 19				
Low	24	9	10	24
High	16	16	0	35
Question 20				
Low	27	10	11	21
High	14	14	0	33
Question 21				
Low	16	11	7	27
High	14	16	0	35
Question 22				
Low	23	9	7	22
High	15	14	1	31
Question 23				
Low	15	9	7	26
High	16	8	0	35

$p = 0.08$).

The data in Table III also show that MBL instruction removed almost all the literal graph misinterpretation error among the high-spatial students. However, fewer low-spatial students overcame these errors. Overall, seven low-spatial students consistently gave "literal" answers to all Acceleration Graph problems in post-FMCE. Further analysis of these students' responses on post-FMCE questions revealed that three of these students gave correct answers on all Sled problems and Coin toss problems, indicating that their difficulties in solving graph problems could be attributed to low spatial ability rather than caused merely by the lack of understanding force and acceleration concepts.

How are Acceleration Graph problems different from Force Graph problems? Why does MBL instruction remove the effect of spatial visualization ability on Force Graph problems, but not on Acceleration Graph problems, although they seem to be very similar? One of the possible explanations for students' difficulties in solving Acceleration vs. Force Graph problems is that most Force Graph problems presented in the FMCE test are less prone to literal misinterpretation than are Acceleration Graph problems. Questions 13, 15, and 16 describe an object "speeding up" or "slowing down," but give explicit verbal instruction that acceleration is constant, which might help students correctly choose the force graph showing the horizontal line**.** In Questions 11–13, the motion of a car is described as having constant velocity, and in this case, the force acting on the object is also constant (zero) force.[5]

In conclusion, the results of Study 1 suggest that prior to physics instruction, students' performance on physics problems that involve graph interpretation as well as on problems that involve two-dimensional motion or multiple objects is related to students' spatial visualization ability. Furthermore, our findings showed that after MBL instruction, spatial visualization ability is not a reliable predictor of students' success in solving most of the force and motion physics problems. The only type of problems that was still related to spatial visualization ability after MBL instruction was Acceleration Graph problems, and further analyses suggested that those graph problems that are prone to literal misinterpretation present the most difficulties to students of low spatial visualization ability.

STUDY 2

Study 2 was conducted to investigate in detail how the amount of MBL instruction and active engagement with MBL tools affects students' spatial visualization ability.

Method

Several groups participated in this study. The first group (Tufts control group) consisted of 61 undergraduate students who took a general non-calculus physics course at Tufts University. This course was taught a year after Study 1 had been conducted by the same instructor as in Study 1 and the sequence of lectures as well as learning materials were identical to those described in Study 1. However, students in this group learned in a more traditional lecture-based environment; that is, they were exposed to only two 40-min Interactive Lecture Demonstrations (Motion with Carts, and Energy of a Cart on the Ramp). Since the exposure to MBL materials was very brief and thus we did not expect a significant increase in spatial visualization ability after only an hour-and-a-half exposure to MBLs,[6]

[5]To explore the above possibility, we conducted an additional study at New Jersey Institute of Technology on a sample of 75 undergraduate students who had taken introductory physics course before the study was conducted. Randomly chosen, half of the students received a version A of Force Graph questions, which included explicit instructions about constant acceleration (as presented in Appendix A). The other half of the students received a version B of Force Graph questions in which such instructions were deleted. All the students received the identical version of Acceleration Graph problems, which did not include any explicit instructions (as presented in Appendix A) as well as the Paper Folding Test. The results showed that the first group of students who received version A outperformed the group of students who received version B ($F(1.74) = 4.56, p < .05$) on the Force Graph questions. No difference was found between the two groups on Acceleration Graph questions ($F(1.74) = 0.11$, $p = 0.74$). Also, we asked four low-spatial students who gave the correct answers to at least one of the Force Graphs Questions but incorrect answers to the similar Acceleration graph problem to explain in a written format their considerations. These students all reported that for the Force Questions (e.g., Question 15), they chose a graph representing a constant force, since Newton's second law explicitly states that "a constant acceleration is caused by a constant force." In contrast, one of the typical explanations for the Acceleration Graph questions (e.g., question 20) was that since "slowing down" implies some change in a motion, the graph describing this motion "should represent a decreasing process."

[6]Tufts groups that received two ILDs was chosen as a control group for the following reason. Although it could be argued that

Table IV. Mean Total Scores on the Paper Folding Pre- and Posttests for the Four Groups

Group	Total score on pretest	Total score on posttest	*t* value
Dickinson college	5.79 (2.87)	7.20 (1.83)	3.1**
Tufts (full ILD sequence)	6.12 (1.85)	6.67 (1.85)	3.65**
Tufts (traditional and 2 ILDs)	6.37 (2.38)	6.94 (1.77)	2.29*
Control	4.31 (2.4)	5.15 (1.62)	1.92

Note. Standard deviations in parentheses.
$^{*}p < 0.05$; $^{**}p < 0.01$.

this group was considered as a control group.

In addition, two other groups participated in Study 2. One group consisted of 35 undergraduate physics majors who took an introductory physics course at Dickinson College. The students studied the topics of kinematics and dynamics based on the *Workshop Physics Activity Guide* (Laws, 1997), which included fifteen 2-h sessions on the topics of kinematics and dynamics based on MBL activities. Students from this group were encouraged to work collaboratively in groups of two and four, depending on the nature of each activity. In general, students were asked to perform experiments to verify theoretical predictions. During their experimental activities they used MBL tools to record, display, and analyze the data. The curriculum covered the topics of one-dimensional motion; gravity and projectile motion, applications of Newton's laws, one- and two-dimensional collision; work and energy, energy conversation, rotation motion; angular momentum and harmonic motion.

The other group consisted of 26 nonscience major students from the Education Department at Tufts University who had not taken any science courses during the semester.

Materials and Procedure

All the participants were pre- and posttested at the beginning and end of the semester using the PFT. We used the same version of the PFT used in Study 1.

it would have been more preferable to find a group of students that received only lecture-based instruction without the use of any technology, such a group would be taught by a different instructor and by different instructional approaches. In addition, it was impossible to find such a group at Tufts University, which has consistently used ILDs for a number of years. Thus, we chose to use a group with two ILDs, which was taught by the same instructor at Tufts as a control group.

Results

The results for the three groups that participated in Study 2 were analyzed together with the results for the group of Tufts students who participated in Study 1. One-way ANOVA indicated that there was significant difference between the four groups in their pretest scores on the PFT, $F(3, 197) = 6.18$, $p < 0.001$. Pairwise comparisons (Turkey post hoc) revealed that the only group that was different in their pretest scores from the three other groups was a group that included nonscience majors. The fact that their spatial visualization scores were significantly lower than those of students from the other groups is consistent with the findings of other researchers (e.g., Lord and Nicely, 1997) that students who preferred science and mathematics do better on spatial tests than students who preferred subjects outside of science and mathematics.

Both Tufts groups (control and experimental) received similar physics instruction: both groups were taught by the same instructors, and the same sequence of lectures and learning materials were used in both classes. As for the Dickinson College group, although their pretest PFT scores were not significantly different from the pre-PFT scores of both Tufts groups, the students in this group were taught by a different instructor and different sequence of materials and different instructional methods were used. For this reason, the direct comparison between Tufts groups and Dickinson College group as well as between Tufts groups and nonscience major group was not possible. Therefore, in our further analysis, we will discuss the Dickinson and nonscience major groups separately.

First, we calculated the total score (using the correction for guessing) for pre- and post-PFT for all four groups. The results are presented in Table IV. Dickinson College and Tufts experimental groups showed a significant increase in their total scores on the PFT ($p < 0.01$). The Tufts control group also showed a significant increase ($p < 0.05$).

Table V. Mean Total Scores on the Paper Folding Tests, the Number of Correctly Solved Items, the Number of Incorrectly Solved Items for Each of the Groups of Pre- and Posttests

Group	Number of correct items on pretest	Number of correct items on posttest	*t* value	Number of incorrect items on pretest	Number of incorrect items on posttest	*t* value
Dickinson	6.48 (2.27)	7.75 (1.52)	5.14**	1.89 (2.05)	1.27 (1.48)	−2.3*
Tufts experimental	6.49 (1.65)	6.91 (1.76)	3.39**	1.51 (1.21)	1.12 (1.13)	−1.92
Tufts control	6.68 (2.17)	7.37 (1.54)	3.18*	1.27 (1.34)	1.72 (1.34)	2.28*
Nonscience majors	4.79 (2.11)	5.66 (1.41)	2.70*	1.41 (1.52)	2.24 (1.64)	2.3*

Note. Standard deviations in parentheses.
$^{*}p < 0.05$; $^{**}p < 0.01$.

The increase for the group of nonscience majors, however, was not significant. In all the statistical analyses, presented earlier, the total score for the PFT was calculated using the correction for guessing: the total score on the test was equal to the number of incorrect items divided by four and then subtracted from the number of correct items. However, the score calculated using this correction for guessing does not eliminate the possibility that the increase in students' scores were due to the increased number of attempted items (both correct and incorrect), and such increases do not automatically imply that a student got fewer items incorrect. For instance, the score of five on the PFT, corrected for guessing, could be a result of five correctly solved items and five unattempted items, or it could be a result of six correctly answered items and four incorrectly answered items. Therefore, we considered separately how the number of correctly and incorrectly answered test items changed from pretest to posttest.

As shown in Table V, all four groups attempted more items (the total amount of attempted items equals the sum of correct and incorrect items) on the posttest in comparison with the pretest. All four groups solved significantly more items on the posttest than on the pretest. As for the number of the incorrectly solved items, the pattern was different: For Tufts control group, the number of incorrectly solved items increased significantly, suggesting the possibility that what changed from pretest to posttest in this group was the number of items attempted, but not the proportion of correctly solved items. For the Tufts experimental group, the number of incorrectly solved items remained the same from the pretest to the posttest, while the number of correctly solved items increased, suggesting that this group, in fact, increased their accuracy in solving the spatial tasks.

The Dickinson College group is the only group that showed a significant decrease in the number of incorrect items from the pretest to posttest ($p < 0.05$) and a significant increase ($p < 0.01$) in the number of correct items. For the group of nonscience majors, the number of incorrectly solved items increased significantly, indicating that what changed from pretest to posttest in this group was the number of items attempted, not the proportion of correctly solved items. The increase in the number of items attempted for nonscience major might be attributed to the test–retest effect reported in other studies (see Lohman, 1988, for a review), which showed that the subjects who take spatial ability tests for the second time usually attempt more items on these tests because of the increased familiarity with the stimuli.

To demonstrate the above-mentioned effect more clearly, we plotted the ratio of correct to incorrect items for all four groups in pretest and posttest in Fig. 5. As shown in the figure, the ratio of correct to incorrect items did not increase in the posttest for the Tufts control group as well as nonscience major group, suggesting that these two groups showed no improvement in their accuracy at solving spatial tasks, while accuracy increased for Tufts experimental and Dickinson groups.[7]

The results clearly indicate that even though students might show an overall tendency to improve their spatial visualization scores as a result of science instruction, the increase in their total scores corrected for guessing could be the result of increases in the number of items attempted but not increases in the proportions of correctly solved items. Thus, it appears that correction for guessing, used as a standard procedure for most psychometric tests, did not eliminate the speed-accuracy trade-off and, as

[7]Since less than 7% of all the students who participated in Study 2 completed all 10 items on the PFT within 3 min time limit, the number of correctly and the number of incorrectly solved items were not confounded variables.

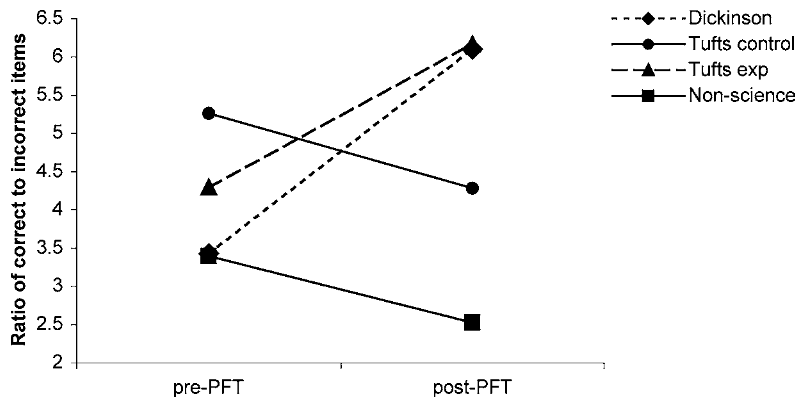


Fig. 5. The ration of correct to incorrect items for all four groups on the paper folding pretest and posttest. Error bars are standard errors.

a result, some of the findings of increased spatial ability reported in the literature may be interpreted incorrectly. Furthermore, the comparison between Tufts experimental and Tufts control group suggests that a full sequence of ILDs is important to produce a significant change in students' spatial visualization ability. Since there is no reason to believe that two ILDs somehow could impair performance on the PFT, our findings suggest that learning physics per se (e.g., in a traditional lecture-based environment) might not be sufficient to cause a significant increase in the accuracy of solving spatial visualization tasks.

STUDY 3

The first goal of Study 3 was to examine whether subjects with a prior physics background would also improve their spatial visualization skills as a result of MBL activities. For this reason, all the participants chosen for Study 3 were middle and high school physics and/or science teachers. The second goal of Study 3 was to replicate the increases in spatial ability found in Study 2 using two independent measures of spatial visualization, the PFT and the three-dimensional Mental Rotation Test.

Method

The participants were 28 middle and high school physics and/or science teachers who participated in a 2-week activity-based summer physics workshop given by West Coast Institute. The goal of the workshop was to expose the teachers to a variety of interactive pedagogies and computer tools, including MBLs, in teaching their courses. The topics discussed in the workshop were not limited to mechanics, but also included electricity and electric circuits as well as general discussions on the role of modeling and guided inquiry in teaching sciences. All the participants were presented with several types of MBL activities: (a) MBL activities in which they were encouraged to work collaboratively in small groups and asked to use MBL tools to record, display, and analyze the data from scientific experiments (overall 9.5 h of the MBL activities with individual sessions ranging from 40 to 140 min); (b) Interactive Lecture Demonstrations (overall 4 h); (c) Open Labs for MBL practice and reflections, where the participants were presented with the possibility to explore and review MBL tools and curricula by themselves and discuss issues with the instructor (overall 4 h).

Materials and Procedure

All the participants were pre- and posttested on the PFT at the beginning and at the end of the physics workshop (with an interval of 2 weeks), and all the participants from the experimental group were pre- and posttested on two measures of spatial visualization ability—the PFT and three-dimensional Mental Rotation Test—at the beginning and at the end of the physics workshop (with an interval of 2 weeks). The same version of the PFT used in Study 1 was presented to the participants in Study 3.

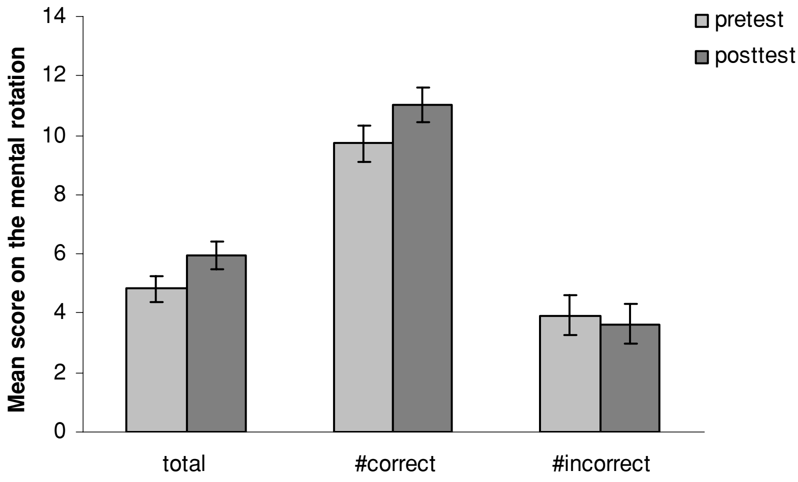


Fig. 6. The total scores, the number of correctly solved items, and the number of incorrectly solved items on the PFT (pretest and posttest). Error bars are standard errors.

Mental Rotation Test

The paper-and-pencil version of the mental rotation test used in this study was developed by Vandenberg and Kuse (1978), on the basis of the stimuli used in the chronometric study of Shepard and Metzler's (1971). The Mental Rotation Test consisted of two parts of 10 items each. Each item consisted of a criterion figure, two correct alternatives, and two incorrect alternatives or "distractors" (see an example in Appendix B). Correct alternatives were always identical to the criterion in structure, but were shown in a rotated position. One half of the distractors were rotated mirror-images of the criterion figures, while the other half were rotated images of one or two of the other criterion figures. Unlike most of the psychometric spatial ability tests, Vandenberg and Kuse eliminated the need to correct for guessing by using the following procedure for scoring: count each item as correct if both correct alternatives are marked by a student, and give no credit otherwise.

Results

Figures 6 and 7 present the total scores corrected for guessing, as well as the number of correctly and incorrectly solved items for the PFT and the Mental Rotation Test, respectively. The paired-samples t test showed a significant increase in the PFT total scores corrected for guessing from posttest vs. pretest ($t(27) = 6.54$, $p < 0.001$) as well as significant increase in the number of correctly solved items ($t(27) = 8.08$, $p < 0.001$) along with a nonsignificant change of the number of incorrectly solved items ($t(27) = 0.9$, $p = 0.37$). As for the Mental Rotation Test, the paired-samples t test showed a significant increase in participants' total scores from pretest to posttest ($t(27) = 2.76$, $p < 0.05$) as well as significant increase of the number of correctly solved items ($t(27) = 2.32$, $p < 0.05$) along with a nonsignificant change of the number of incorrectly solved items ($t(27) = -0.61$, $p = 0.54$). Consistent with the findings of Study 2, the group of science teachers significantly increased their accuracy in solving both the PFT and the Mental Rotation Test problems as a result of MBL exposure.

DISCUSSION

Importance of Spatial Ability in Physics Problem Solving

While many educational and cognitive psychology studies have focused on the importance of the general knowledge of physics laws as well as verbal and analytical skills, surprisingly little attention has been devoted to understanding the use of spatial visualization in physics problem solving. For instance, research on expert-novice problem solving in physics (e.g., Chi and Glaser, 1988; Ericsson and Smith, 1991; Larkin, 1982) has focused mostly on verbal aspects

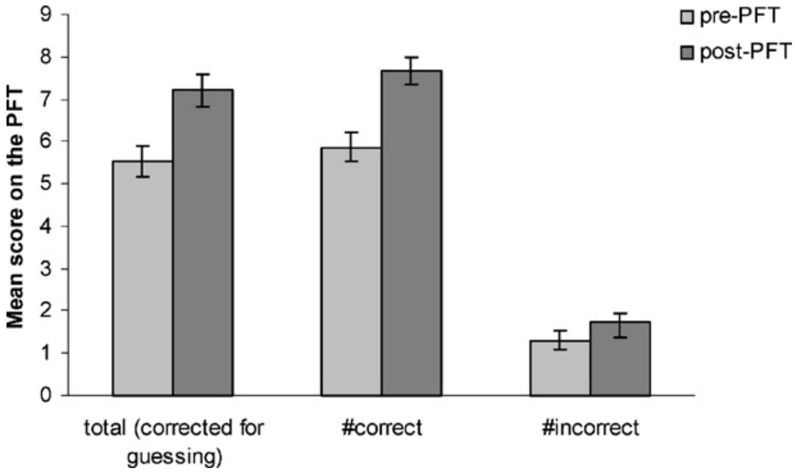


Fig. 7. The total scores, the number of correctly solved items, and the number of incorrectly solved items on the Mental Rotation test (pretest and posttest). Error bars are standard errors.

of problem representations and the semantic knowledge of physics laws, even though the importance of visual/spatial representations has been repeatedly stressed. Similarly, traditional physics curricula have focused mostly on computational aspects of physics problem solving.

The present research shows that spatial visualization ability significantly influences effectiveness of physics instruction but that not all physics problems require the significant use of visual/spatial processing resources. Finding solutions for one-dimensional problems involving judgments about the motion characteristics of only one object (e.g., Force Sled, Coin Toss problems) requires mostly the general knowledge of Newtonian laws and does not make large demands on visual/spatial working memory. In contrast, finding solutions for multidimensional problems (e.g., Force Graph, Acceleration Graph, and Collision problems) seems to require spatial visualization ability, although MBL instruction appears to be able to reduce or even eliminate this relationship. What do multidimensional physics problem solving and spatial visualization ability have in common, and what are the other types of physics problems that might require high visual/spatial resources?

Cognitive research that examined how people predict objects' motions suggests that two factors affect the influence of visual/spatial ability on processing motion events. The first factor is the nature of the overall motion and the second factor is the number of relevant motion parameters or moving objects that the mechanical system contains (Isaak and Just, 1995; Kozhevnikov *et al.*, 2002; Law *et al.*, 1993). Our results showed that correctly answering collision questions that involved judgments about the motion of two colliding objects, at least before the MBL instruction, correlated significantly with spatial visualization ability, and this finding is consistent with the notion that increasing the number of motion parameters and the number of moving objects places greater demands on visual/spatial resources. Furthermore, our finding showed that interpretation of kinematics graphs is another type of physics problems that requires high visual/spatial resources. We found that low-spatial students were more susceptible to "graph-as-picture" interpretations than were high-spatial students. This might have occurred because interpreting such graphs requires translating the abstract graphical representation into a real motion event and this translation process might place large demands on visual/spatial working memory, especially when students do not have appropriate knowledge for associating graph features with quantitative relationships depicted on the graph. As a result, low-spatial students might experience more difficulties than do high-spatial students in graph interpretations and tend to interpret a graph as a picture. One of the limitations of the current study is that the Force and Acceleration graphs given to students were mostly one-interval graphs describing only one particular type of motion (e.g., accelerating motion or constant motion). Future research is needed to explore the relationship between more complex kinematics graphs, MBL instruction, and spatial visualization ability. Presumably, the correct interpretation of more complex graphs will require greater visual/spatial resources, and for low-spatial students, gains from MBL instruction might not be able to compensate for such demands. Additionally, future research is needed to identify other types of mechanics problems that require high-spatial visualization skills to solve. The results of such research will have important instructional implications regarding the development and the use of different visualization aids for solving physics problems.

How Can Educational Technology Facilitate Visual/Spatial Processing?

The finding that spatial reasoning is an important skill in solving many types of physics problems raises questions about the properties of visual displays that should facilitate the use of visual/spatial processing strategies and in particular, help low-spatial students learn from spatial physics concepts and diagrams. Cognitive science research has shown that visualization alone and even dynamic simulations do little to help people understand the dynamics of systems that involve multiple parameters or multiple objects (Kaiser *et al.*, 1992). Moreover, research has shown that low-spatial students have more difficulties than that of high-spatial students in extracting necessary visual information from dynamic animation (Isaak and Just, 1995). It also has been suggested that animation should possess specific features to develop spatial understanding and competence, such as, allowing students to break a system down into a point-particle system with single dynamically relevant parameters or drawing students' attention to a single element of the problem and thus minimizing the influence of other elements (Kaiser *et al.*, 1992).

The main goal of the present research was to determine if MBL environments facilitate students' understanding of force and motion concepts and whether learning in such environments helps to reduce the role of spatial visualization ability on physics performance. The results of the present research revealed that spatial visualization ability was a reliable predictor of students' performance on the physics conceptual evaluation tests before physics instruction. However, after MBL instruction, the correlation between spatial visualization and students' performance on the physics conceptual test was reduced. This change in the relationship between spatial visualization and performance on the physics test might have occurred because during the course of instruction the MBL display presented students with a vast variety of visual graphical representations. Thus, the students were able to develop a set of visual mental templates representing a variety of different classes of motion (Linn *et al.*, 1987). Additionally, the temporal proximity between real motion events and their graphical representations facilitated cognitive linking between the physics concepts and the graphical presentations, and such linking might have decreased cognitive load and the amount of visual/spatial resources needed for graph interpretation by providing students with a mental schemas in which different graphical representations were associated with corresponding types of motion. Furthermore, real-time graphing enabled students to selectively attend to salient points on the graphs representing changes in speed or direction (Brasell, 1987) and thus highlighting the most relevant experiences. This, in turn, might have helped students to learn how to use their visual–spatial resources effectively and how to direct them to the relevant objects and experiences. In addition, students were able to break down the physical system by focusing on the graph of only one of the objects in the experiment or to graph any one of the motion parameters on the screen.

All the above consequences of MBL instruction should facilitate decreasing cognitive load and decreasing the amount of visual/spatial resources necessary to solve a given physics problem. Consistent with our findings, as a result of their 3-year longitudinal study of teaching principles and learning geometry, Lehrer and Chazan (1998) reported that the longitudinal development of spatial visualization ability was unrelated to the development of the mathematical understanding of spatial reasoning (e.g., students' understanding of such concepts as planar figures, angle, length, area, etc.). Thus, it is possible that in the process of learning the mathematics of spatial reasoning, students developed their competence in representing these concepts, and once developed, this competence allowed them to effectively construct and manipulate such mathematical spatial representations independently of the development of their spatial visualization ability.

Furthermore, many of the MBLs provided students the possibility to reflect on their visual graphical representations of a variety of abstract concepts, and this, as well as other features of the MBLs (e.g., meaningful context, student-controlled experiments, physical experience), seemed to help students to improve their performance not only on the physics test but also on spatial visualization tests. Taking into account the results of the previous studies that showed that specific spatial training (e.g., rotating letter) does not transfer to other spatial tasks (e.g., rotating figures), the results of the current study might seem surprising. However, we believe that the variety of visual–spatial graphical representations to which students were exposed during the course of MBL instruction as well as the possibility to interact directly with and explore hypotheses about these representations assisted students in developing their skills to deal with abstract spatial material, which according to Lohman's (1979, p. 126) definition of spatial visualization as an "ability to generate, retain, and manipulate abstract visual images" constitutes the development of spatial ability. Moreover, according to our findings, MBL activities improved spatial visualization skills even for those who already have science background indicating that the experience with visual graphical representations rather than learning physics concepts per se caused the change in spatial visualization skills.

An important implication of the current study is the finding that it is possible to improve students' performance on spatial visualization tests by presenting them with a variety of meaningful abstract visual images and giving them the possibility to manipulate and explore such images. Moreover, the data show that such changes could not be attributed to speed-accuracy trade-off or test–retest effects. However, our findings also pose an important question for future research: In particular, what is the underlying reason for this improvement? Do students' visual–spatial working memory capacity improves, or do students become more proficient with manipulation of spatial representations and understanding spatial relations? We believe that to answer such a question, rigorous experimental studies will be needed. One possibility is to conduct laboratory experiments, based on contemporary visual processing theory, to compare changes before and after MBL instruction in spatial imagery components such as image quality, efficiency of image generation, image rotation, image scanning, adding and subtracting detail in image, as well as integration of images (see, for example, Pellegrino *et al.*, 1985; Poltrock and Brown, 1984). Another possibility is to use different neuroimaging techniques to determine the effects of MBL instruction on brain activation patterns while solving spatial tasks (e.g., whether students show greater activation of specific brain areas after MBL instruction or whether students show new areas of activation). We believe that such studies will be very valuable for understanding what spatial ability is and how it can be enhanced.

APPENDIX A

Force and Motion Conceptual Evaluation

(The correct answers are marked at the beginning of each question)

Directions: Answer questions 1–32 in spaces on the answer sheet. Be sure your name is on the answer sheet. Hand in the questions and the answer sheet.

A sled on ice moves in the ways described in questions 1–7 below. *Friction is so small that it can be ignored.* A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the *one* force (**A** through **G**), which would *keep the sled moving* as described in each statement below.

You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice **J**.

B 1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?

D 2. Which force would keep the sled moving toward the right at a steady (constant) velocity?

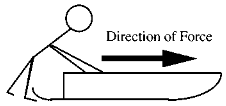
F 3. The sled is moving toward the right. Which force would slow it down at a steady rate (constant acceleration)?

F 4. Which force would keep the sled moving toward the left and speeding up at a steady rate (constant acceleration)?

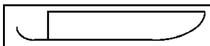
D 5. The sled was started from rest and pushed until it reached a steady (constant) velocity toward the right. Which force would keep the sled moving at this velocity?

B 6. The sled is slowing down at a steady rate and has acceleration to the right. Which force would account for this motion?

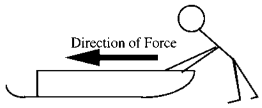
B 7. The sled is moving toward the left. Which force would slow it down at a steady rate (constant acceleration)?



A. The force is toward the **right** and is **increasing** in strength (magnitude).
B. The force is toward the **right** and is of **constant** strength (magnitude).
C. The force is toward the **right** and is **decreasing** in strength (magnitude).



D. No applied force is needed



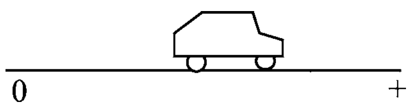
E. The force is toward the **left** and is **decreasing** in strength (magnitude).
F. The force is toward the **left** and is of **constant** strength (magnitude).
G. The force is toward the **left** and is **increasing** in strength (magnitude).

Questions 8–10 refer to a coin, which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (**A** through **G**) to indicate the force acting on the coin for each of the cases described below. Answer choice **J** if you think that none is correct. **Ignore any effects of air resistance.**

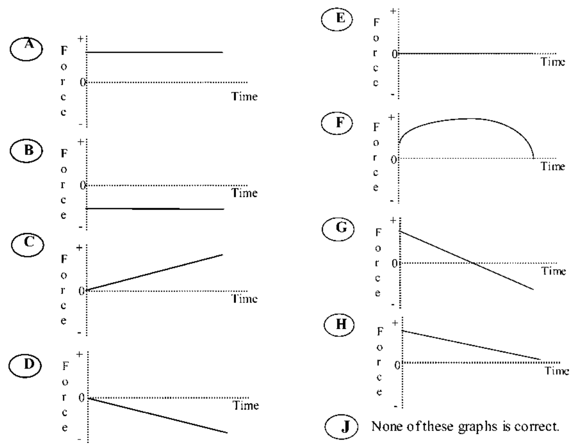
A. The force is **down** and constant.
B. The force is **down** and increasing.
C. The force is **down** and decreasing.
D. The force is zero.
E. The force is **up** and constant.
F. The force is **up** and increasing.
G. The force is **up** and decreasing.

A 8. The coin is moving upward after it is released.
A 9. The coin is at its highest point.
A 10. The coin is moving downward.

Questions 11–18 refer to a toy car, which can move to the right or left along a horizontal line (the positive part of the distance axis).



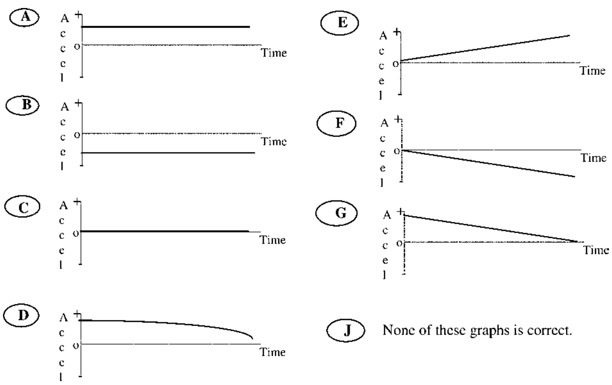
Assume that friction is so small that it can be ignored. A force is applied to the car. Choose the one force graph (**A** through **H**) for each statement below, which could allow the descried motion of the car to continue. You may use a choice more than once or not at all. If you think that none is correct, answer choice **J**.



E 11. The car moves toward the right (away from the origin) with a steady (constant) velocity.
E 12. The car is at rest.
A 13. The car moves toward the right and speeding up at a steady rate (constant acceleration).
E 14. The car moves toward the left (toward the origin) with a steady (constant) velocity.
B 15. The car moves toward the right and slowing down at a steady rate (constant acceleration).
B 16. The car moves toward the left and is speeding up at a steady rate (constant acceleration).
G 17. The car moves toward the right, speeds up and then slows down.
E 18. The car pushed toward the right and then released. Which graph describes the force *after* the car is released?

Different motions of the car are described below. Choose the letter (**A** to **G**) of the **acceleration–time** graph, which corresponds to the motion of the car described in each statement.

You may use a choice more than once or not at all. If you think that none is correct, answer choice **J**.

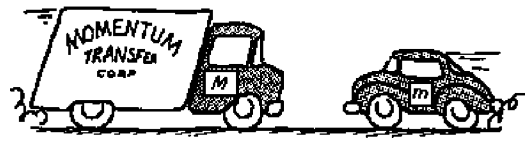


A 19. The car moves toward the right (away from the origin), speeding up at a steady rate.
B 20. The car moves toward the right, slowing down at a steady rate.
C 21. The car moves toward the left (toward the origin) at a constant velocity.
B 22. The car moves toward the left, speeding up at a steady rate.
C 23. The car moves toward the right at a constant velocity.

Questions 24–28 refer to collisions between a car and trucks. For each description of a collision below, choose the one answer from the possibilities **A** though **J** that best describes the forces between the car and the truck.

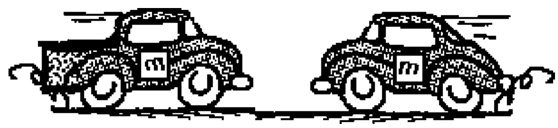
A. The truck exerts a greater amount of force on the car than the car exerts on the truck.
B. The car exerts a greater amount of force on the truck than the truck exerts on the car.
C. Neither exerts a force on the other; the car gets smashed simply because it is in the way of the truck.
D. The truck exerts a force on the car but the car doesn't exert a force on the truck.
E. The truck exerts the same amount of force on the car as the car exerts on the truck.
F. Not enough information is given to pick one of the answers above.
J. None of the answers above describes the situation correctly.

In quaestion 24 through 26 the truck is ***much heavier*** *than the car.*



E 24. They are both moving at the same speed when they collide. Which choice describes the forces?
E 25. The car is moving much faster than the heavier truck when they collide. Which choice describes the forces?
E 26. The heavier truck is standing still when the car hits it. Which choice describes the forces?

In questions 27 and 28 the truck is a small pickup and is the ***same weight*** *as the car.*



E 27. Both the truck and the car are moving at the same speed when they collide. Which choice describes the forces?
E 28. The truck is standing still when the car hits it. Which choice describes the forces?

Questions 29–31 refer to a large truck which breaks down out on the road and receives a push back to town by a small compact car.

Pick one of the choices **A** through **J** below, which correctly describes the forces between the car and the truck for each of the descriptions.

A. The force of the car pushing against the truck is equal to that of the truck pushing back against the car.
B. The force of the car pushing against the truck is less than that of the truck pushing back against the car.

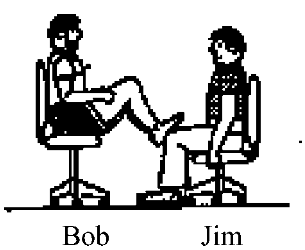
C. The force of the car pushing against the truck is greater than that of the truck pushing back against the car.
D. The car's engine is running so it applies a force as it pushes against the truck, but the truck's engine isn't running so it can't push back with a force against the car.
E. Neither the car nor the truck exert any force on each other. The truck is pushed forward simply because it is in the way of the car.
J. None of these descriptions is correct.

A 29. The car is pushing on the truck, but not hard enough to make the truck move.

A 30. The car, still pushing the truck, is **speeding up** to get to cruising speed.

A 31. The car, still pushing the truck, is at cruising speed when the truck puts on its brakes and causes the car to **slow down**.

E 32. Two students sit in identical office chairs facing each other. Bob has a mass of 95 kg, while Jim has a mass of 77 kg. Bob places his bare feet on Jim's knees, as shown to the right. Bob then suddenly pushes outward with his feet, causing both chairs to move. In this situation, while Bob's feet are in contact with Jim's knees,

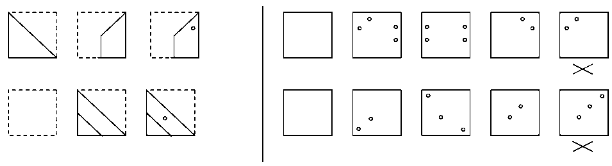


A. Neither student exerts a force on the other.
B. Bob exerts a force on Jim, but Jim doesn't exert any force on Bob.
C. Each student exerts a force on the other, but Jim exerts the larger force.
D. Each student exerts a force on the other, but Bob exerts the larger force.
E. Each student exerts the same amount of force on the other.
J. None of these answers is correct.

APPENDIX B

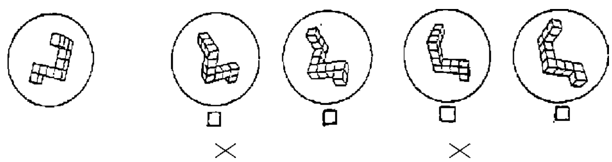
The Example of Items from the Paper Folding Test

In this test you are to imagine the folding and unfolding of pieces of papers. The figures on the left represent a square piece of paper being folded, and the last of these figures has one or two small circles drawn on it to show where the paper has been punched. One of the five figures on the right shows where the holes will be when the paper is unfolded. You are to decide which one of these figures is correct. (These sample problems are done for you).



The Example of Items from the Mental Rotation Test

This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. For each problem there is a primary object on the far left. You are to determine which two of the four objects to the right are the same object given on the far left. In each problem always *two* of the four drawings are the same object as the one on the left. You are to put Xs in the boxes below the correct ones, and leave the incorrect ones blank. (This sample problem is done for you).



ACKNOWLEDGMENTS

This research was partially supported by the National Science Foundation under contract REC-0106760 to the first author and REC-9903309 to the second author. We thank Priscilla Laws for helping to conduct the portion of the study that took place at Dickinson College and Luz-Alterman and

Amy Roster for providing a control group at the Education Department at Tufts University. We also thank Michael Motes and Olessia Blajenkova for their help in analyzing the results.

REFERENCES

Baddeley, A. (1992). Is working memory working? The fifteen Barlett lecture. *The Quarterly Journal of Experimental Psychology* 44A: 1–31.

Barnea, N., and Dori, Y. (1999). High-school chemistry students' performance and gender differences in a computer molecular modeling learning environment. *Journal of Science Education and Technology* 8: 257–271.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics* 62: 750–762.

Beichner, R. J. (1996). The impact of video motion analysis on kinematics graph interpretation skills. *American Journal of Physics* 64: 1272–1277.

Bethell-Fox, C. E., and Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance* 14: 12–23.

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching* 24: 385–395.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., and Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Journal of Cognitive Neuroscience* 11: 9–24.

Champagne, A. B., Klopfer, L. E., and Anderson, J. H. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics* 48: 1074–1079.

Chi, M. T. H., and Glaser, R. (1988). *The Nature of Expertise*, Erlbaum, Hillsdale, NJ.

Cummings, K., Marx, J., Thornton, R., and Kuhl, D. (1999). Evaluating innovation in studio physics. *American Journal of Physics* 67: 38–44.

Dictionary of Occupation Titles (1991). *U.S. Department of Labor, Employment and Training Administration*, U.S. Employment Service, Career Press.

Ekstrom, R. B., French, J. W., and Harman, H. H. (1976). *Manual for Kit of Factor Referenced Cognitive Tests*, Educational Testing Service, Princeton, NJ.

Ericsson, K. A., and Smith, J. (1991). *Toward a General Theory of Expertise*, Cambridge University Press, Cambridge, UK.

Halloun, I. A., and Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics* 53: 1043–1055.

Hestenes, D. (1985). Toward a modeling theory of physics intuition. *American Journal of Physics* 55: 440–454.

Hestenes, D. (1995). What do graduate oral exam tell us? *American Journal of Physics* 63: 1069.

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *American Journal of Physics* 30: 141–154.

Isaak, M. I., and Just, M. A. (1995). Constraints on the processing of rolling motion: The curtate cycloid illusion. *Journal of Experimental Psychology: Human Perception and Performance* 21: 1391–1408.

Just, M. A., and Carpenter, P. A. (1985). Cognitive coordinate processes: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review* 8: 441–480.

Kaiser, M. K., Profitt, D. R., Whelan, S. M., and Helth, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance* 18: 669–690.

Kozhevnikov, M., Hegarty, M., and Mayer, R. E. (2002). Visual/spatial abilities in problem solving in physics. In Anderson, M., Meyer, B., and Olivier, P. (Eds.), *Diagrammatic Representation and Reasoning*, Springer, Berlin.

Larkin, J. H. (1982). The role of problem representations in physics. In Gentner, D., and Stevensm, A. L. (Eds.), *Mental Models*, Hillsdale, NJ, pp. 75–98.

Law, D. J., Pellegrino, J. W., Mitchell, S. R., Fischer, S. C., McDonald, J. P., and Hunt, E. B. (1993). Perceptual and cognitive factors governing performance in comparative arrival-time judgments. *Journal of Experimental Psychology: Human Perception and Performance* 19: 1183–1199.

Laws, P. (1997). *Workshop Physics Activity Guide. The Core Volume with Module 1: Mechanics I*, Wiley, New York.

Lehrer, R., and Chazan, D. (1998). *Designing Learning Environments for Developing Understanding of Geometry and Space*, Erlbaum, Mahwah, NJ.

Linn, M. C., Layman, J., and Nachmias, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Contemporary Educational Psychology* 12: 244–253.

Lohman, D. F. (1979). *Spatial Ability: A Review and Re-Analysis of the Correlational Literature*. Aptitude Research Project, Stanford University School of Education technical report No. 8, Stanford, CA.

Lohman, D. (1988). Spatial abilities as traits, processes, and knowledge. In Sternberg, R. J. (Ed.), *Advances in the Psychology of Human Inteligence*, Vol. 4, Erlbaum, Hillsdale, NJ, pp. 181–248.

Lohman, D. F., and Nichols, P. D. (1990). Training spatial abilities: Effects of practice on rotation and synthesis tasks. *Learning and Individual Differences* 2: 67–93.

Lord, T. R. (1985). Enhancing the visuo-spatial aptitude of students. *Journal of Research in Science Teaching* 22: 395–405.

Lord, T., and Holland, M. (1997). Preservice secondary education majors and visual–spatial perception: An important cognitive aptitude in the teaching of science and mathematics. *Journal of Science Teacher Education* 8: 43–53.

Lord, T., and Nicely, G. (1997). Does spatial aptitude influence science–math subject preferences of children? *Journal of Elementary Science Education* 9: 67–81.

Lord, T. R., and Rupert, J. L. (1995). Visual–spatial aptitude in elementary education majors in science and math tracks. *Journal of Elementary Science Education* 7: 47–58.

McDermott, L. C., Rosenquist, M. L., and van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics* 55: 503–513.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin* 86: 889–918.

Mokros, J., and Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching* 24: 369–383.

Mousavi, S., Low, R., and Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology* 87: 319–334.

Pallrand, G., and Seeber, F. (1984). Spatial ability and achievement in introductory physics. *Journal of Research in Science Teaching* 21: 507–516.

Pellegrino, J. W., Mumaw, R. J., and Shute, V. (1985). Analyses of spatial aptitude and expertise. In Embretson, S. (Ed.), *Test Design: Developments in Psychology and Psychometrics*, Academic Press, New York.

Poltrock, S. E., and Brown, P. (1984). Individual differences in spatial ability. *Intelligence* 8: 93–138.
Roe, A. (1953). *Making of a Scientist*, Dodd, Mead, New York.
Rosenquist, M. L., and McDermott, L. C. (1987). A conceptual approach to teaching kinematics. *American Journal of Physics* 55: 407–415.
Salthouse, T. A., Babcock, R. L., Mitchell, D. R. D., Palmon, R., and Skovronek, E. (1990). Sources of individual differences in spatial visualization ability. *Intelligence* 14: 187–230.
Shah, P., and 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.
Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171: 701–703.
Siemankowski, F., and MacKnight, F. (1971). Spatial cognition: success prognosticator in the science courses. *Journal of College Science Teaching* 59: 1–56.
Sokoloff, D., and Thornton, R. K. (1997). Using interactive lecture demonstrations to create an active learning environment. *The Physics Teacher* 35: 340–346.
Thornton, R. K. (1991). Using the microcomputer-based laboratory to improve student conceptual understanding in physics. *Turkish Journal of Physics* 15: 316–335.
Thornton, R. K. (1992). Enhancing and evaluating students' learning of motion concepts. In Tiberghien, A., and Mandl, H. (Eds.), *Physics and Learning Environments*, NATO ASI Series F, 86, Springer, Berlin, pp. 265–283.
Thornton, R. K. (1993). Changing the physics teaching laboratory: Using technology and new approaches to learning to create an experiential environment for learning physics concepts. In Oblak, S., and Razpet, N. (Eds.), *Proceedings of the Europhysics Study Conference, The Role of Experiment in Physics Education*, Ljubljana, Slovenia.
Thornton, R. K. (1996). Using large-scale classroom research to study student conceptual learning in mechanics and to develop new approaches to learning. In Tinker, R. F. (Ed.), *Microcomputer-Based Labs: Educational Research and Standards*, NATO ASI Series F, 156, Springer, Berlin.
Thornton, R. K. (1997). Using large-scale classroom research to study student conceptual learning in mechanics and to develop new approaches to learning. In *NATO ASI Series*, Spinger, Berlin.
Thornton, R. K. (1999a). Using the results of research in science education to improve science learning. In Proceedings of the *International Conference on Science Education*, Nicosia, Cyprus.
Thornton, R. K. (1999b). Why don't physics students understand physics? Building a consensus, fostering change. In Chaisson, E. J., and Kim, T. C. (Eds.), *The Thirteenth Labor, Improving Science Education*, Gordon and Breach, Amsterdam.
Thornton, R. K., and Sokoloff, D. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics* 58: 858–866.
Thornton, R. K., and Sokoloff, D. (1992). *Tools for Scientific Thinking—Motion and Force Curriculum and Teachers' Guide*, 2nd edn., Vernier Software, Portland.
Thornton, R. K., and Sokoloff, D. (1998). Assessing student learning of Newton's laws: The force and motion conceptual evaluation of active leaning laboratory and lecture curricula. *American Journal of Physics* 66: 338–352.
Trowbridge, D. E., and McDermott, L. C. (1981). Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics* 49: 242–253.
Vandenberg, S., and Kuse, A. R. (1978). Mental rotation, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills* 47: 599–604.
White, B. V. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction* 18: 1–100.