Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2015, Vol. 26(6) 737–749

DOI: 10.1177/0956797615569355
pss.sagepub.com
SAGE

Physical Experience Enhances Science Learning

Carly Kontra[1,2], Daniel J. Lyons[1,2], Susan M. Fischer[3], and Sian L. Beilock[1,2]
[1]Department of Psychology, University of Chicago; [2]Committee on Education, University of Chicago; and [3]Department of Physics, DePaul University

Abstract
Three laboratory experiments involving students' behavior and brain imaging and one randomized field experiment in a college physics class explored the importance of physical experience in science learning. We reasoned that students' understanding of science concepts such as torque and angular momentum is aided by activation of sensorimotor brain systems that add kinetic detail and meaning to students' thinking. We tested whether physical experience with angular momentum increases involvement of sensorimotor brain systems during students' subsequent reasoning and whether this involvement aids their understanding. The physical experience, a brief exposure to forces associated with angular momentum, significantly improved quiz scores. Moreover, improved performance was explained by activation of sensorimotor brain regions when students later reasoned about angular momentum. This finding specifies a mechanism underlying the value of physical experience in science education and leads the way for classroom practices in which experience with the physical world is an integral part of learning.

Keywords
embodied cognition, science education, cognitive neuroscience, STEM learning, fMRI, motor activation, open data, open materials

Received 9/15/14; Revision accepted 1/5/15

There is a rapid shift in education from traditional classrooms to virtual and online learning environments (Allen & Seaman, 2010; Lindgren & Johnson-Glenberg, 2013). Though there are certainly benefits associated with learning that does not require a teacher to be physically present or a student to be in a classroom, one issue that is often ignored—at least in science education—is that during virtual learning, students are not physically experiencing the concepts they are learning about. Even active learning environments centered around small-group collaboration and problem solving often involve students observing phenomena rather than physically experiencing them (Cummings, 2008; Freeman et al., 2014; Singer, Nielsen, & Schweingruber, 2012). In physics classrooms, for instance, observing the consequences of mechanical forces on objects and feeling the consequences of a force directly are both construed as active student engagement (Meltzer & Thornton, 2012). Focusing on the physics of mechanics, in this article we demonstrate that physical experience with the material world leads to the recruitment of sensorimotor brain systems evolved for computing forces, vectors, and trajectories, which aids understanding of complex science concepts involving kinetics. The President's Council of Advisors on Science and Technology (2012) recently identified the adoption of empirically validated teaching practices as critical to the goal of increasing the number of college-level science, technology, engineering, and mathematics graduates by 33%. Physical experience can enhance learning

Corresponding Authors:
Carly Kontra, Department of Psychology, University of Chicago, 5848 S. University Ave., Chicago, IL 60637
E-mail: ckontra@uchicago.edu

Sian L. Beilock, Department of Psychology, University of Chicago, 5848 S. University Ave., Chicago, IL 60637
E-mail: beilock@uchicago.edu

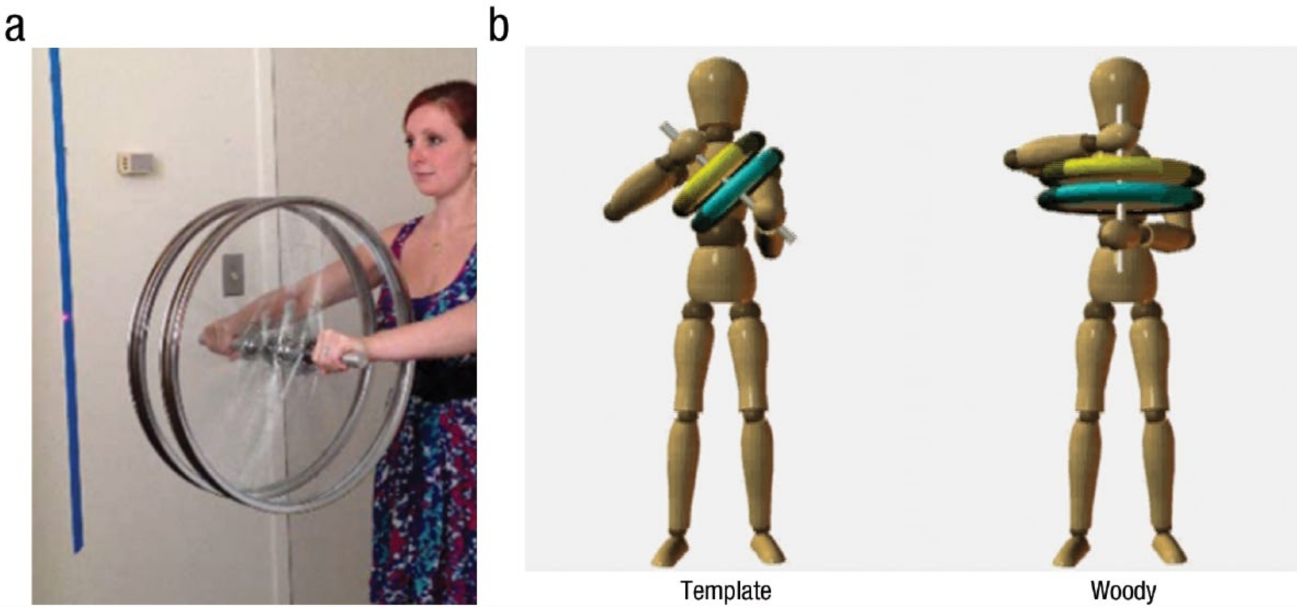

Fig. 1. Materials and stimuli in Studies 1 through 3: (a) the double bicycle wheel (with the red laser dot and blue target line) used for training and (b) a screenshot of the avatars in the pre- and posttests. The yellow and blue stripes on the wheels held by the avatars emphasized the velocities at which the wheels were spinning. See the text for additional details.

and is thus crucial to consider in the design of effective science curricula.

A line of psychological research that falls under the heading *embodied cognition* explains how physical experience influences understanding. According to theories of embodied cognition, thinking—whether one is recalling memories, reasoning, or making inferences—involves activations of the sensory and motor systems initially used to acquire relevant information (Barsalou, Simmons, Barbey, & Wilson, 2003; Niedenthal, 2007). Expert dancers activate sensory and motor regions of the brain involved in dancing more when they watch videos of movements they have practiced in the past than when they watch unfamiliar movements (Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005). And when experienced hockey players hear descriptions of hockey actions (e.g., "The hockey player shot the puck"), they activate the motor system (e.g., dorsal premotor cortex, or dPMc) more than do novice hockey players and fans with only ice-hockey viewing experience. Furthermore, the extent of this motor-system activation predicts how well the hockey scenarios are understood (Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008). Extending these ideas, we hypothesized that interaction with physical quantities should improve students' understanding of relevant concepts in physics by activating sensorimotor brain systems used to execute similar actions in the past, adding kinetic details and meaning to their thinking (Barsalou et al., 2003; Beilock et al., 2008; Glenberg, 1997; Zwann & Taylor, 2006).

Focusing on the physics of mechanics as a test bed, we investigated whether physically experiencing properties of angular momentum aids students' understanding of the concept. In the everyday world, angular momentum is involved when one observes the nose of a football thrown in a tight spiral tip to follow the ball's trajectory, or when one discovers that a moving bicycle is more stable than a stationary one. In physics terms, angular momentum is the product of the moment of inertia of a spinning object and its rotational velocity; the moment of inertia depends on the mass of the object and on the way the mass is distributed. Angular momentum is a vector quantity, which means that it has a magnitude and a direction. Vectors are repeatedly encountered in introductory physics courses and beyond, in more advanced classes and real-world applications of physics.

Our studies used a system of two bicycle wheels that spun independently about a single axle and exploited a consequence of angular momentum's vector nature to allow students to directly feel properties of angular momentum. A laser pointer was mounted in the axle of the wheels, and the red dot of the laser was directed at a blue vertical line on the wall (Fig. 1a). To physically experience the properties of angular momentum, students held the set of spinning bicycle wheels by the axle and were instructed to tilt the axle through space from horizontal to vertical and back while attempting to keep the laser on the target vertical line on the wall. When a wheel spins, the angular-momentum vector points along the axle of the wheel. When the axle is tilted, the direction of this vector changes, and a corresponding torque is felt as a resistive force. This torque was observed as a deflection of the laser away from the vertical line on the wall, such that the laser's red dot traced an ellipse that broadened as the resistive force increased with greater angular momentum and torque. In a control trial, which was repeated throughout the training, neither of the wheels spun, and the red dot of the laser closely followed the vertical line.

In four experiments exploring students' behavior and brain function, we examined whether their understanding of angular momentum differed depending on whether they had (a) physically experienced consequences of manipulating the bicycle wheels or (b) observed consequences of the wheels' changing angular momentum via the laser dot on the wall. We focused on college-level science because it defines science education by modeling for teachers and adults how science should be taught for younger students (Alberts, 2013). We reasoned that physical experience is a key ingredient in enhancing learning about the concept of angular momentum because such experience prompts activation of sensory and motor brain regions when students later think about the concept. Such activation infuses kinetic information into thinking and thereby allows students to better reason about angular momentum and the related concepts of torque and force.

Studies 1 Through 3

Subjects in Studies 1 through 3 were drawn from the University of Chicago community and received either payment or course credit in exchange for their participation. Because we were interested in the initial stages of learning, we limited participation to individuals with no college-level physics experience. The avatar videos used in the pre- and posttests were prepared using Poser (Smith Micro Software, Inc., Santa Cruz, CA) animation software and were displayed to subjects with Presentation software (Neurobehavioral Systems, www.neurobs.com).

Study 1

Subjects in Study 1 were 22 pairs of college-age students (N = 44; 9 males). The sample size was determined by collecting data through the end of spring quarter and then stopping; 2 subjects were dropped because of computer error.

Subjects first read a description of angular momentum and the factors that affect it. They then completed a pretest that gauged their initial understanding of the relations between physical properties of spinning objects (e.g., moment of inertia, angular velocity) and the amount of torque (resistive force) exerted when spinning objects are tilted through space. In each pretest trial, a video displaying two avatars simultaneously was presented (Fig. 1b). Each avatar held a pair of bicycle wheels on an axle and tilted the axle through space from horizontal to vertical and back. All parameters for the avatar and apparatus on the left (Template) stayed the same for each of the 52 trials (and 6 practice trials). The avatar on the right (Woody) held an apparatus that changed on every trial and differed from Template's apparatus in one or more of the following ways: the direction in which each wheel spun, the speed at which each wheel spun, the size of each wheel, and the direction in which the axle was tilted (toward the right or left). The 52 trials tested all manipulated factors, but were not exhaustive (i.e., not fully crossed); for most trials, the two wheels in Woody's apparatus had different sizes and spin speeds. Spin velocity was made perceptually salient by the colored stripes on the wheels (see Fig. 1b). Subjects ended each video with a key press and then answered two forced-choice questions, also with key presses:

- Question 1: "Did Woody experience more or less force than the Template?"
- Question 2: "Are the forces in the same or in different directions?"

Higher accuracy on Question 1, our main focus in this series of studies, indicated better comprehension of the relevant properties of the physical system and their impact on the angular momentum and resulting torque. The analyses of accuracy we report in the main text are based on responses to this question (Question 2 was dropped in Studies 2 and 3).

Following the pretest, one subject in each pair was assigned to the *action group* and was told that he or she would be tilting a set of wheels similar to the set in the videos. The other subject was assigned to the *observation group* and was told to closely observe the tilting and the path of the red laser dot on the wall. Before each of 19 trials, which consisted of tilting the wheels from vertical to horizontal (90°) and back five times in 5 s, the experimenter verbally described to both subjects what was about to occur (i.e., how the wheels' spin direction, spin speed, and size would change from the previous trial). The subject in the action group was told to tilt the wheels so that the red dot of the laser followed the vertical target line on the wall (Fig. 1a). Deflections away from the target line were directly related to the torque, or resistive force, exerted by the apparatus when it was tilted and were visible via the laser pointer's path on the wall. After the 10 min of training, both subjects completed a posttest that was identical to the pretest.

Pretest performance did not differ as a function of group, as revealed by a one-way analysis of variance

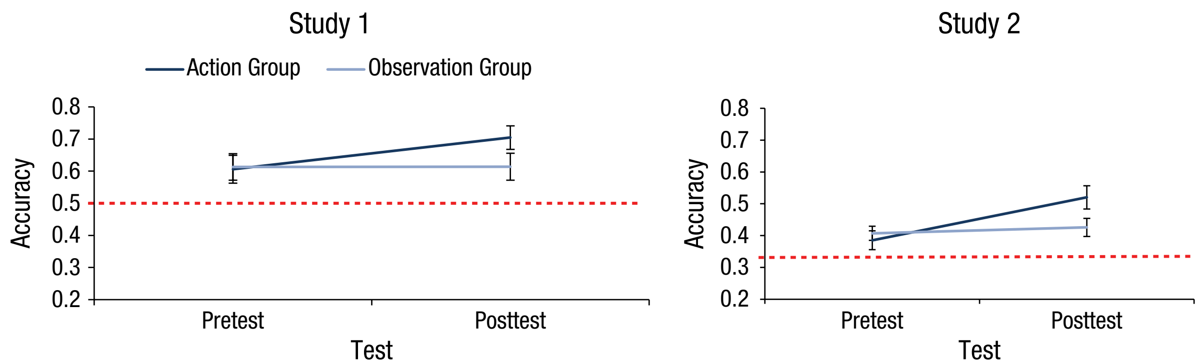

Fig. 2. Accuracy in Studies 1 and 2 as a function of test (pretest vs. posttest) and group (action vs. observation). The red dashed lines indicate chance performance, which was lower in Study 2 than in Study 1 because of the addition in Study 2 of a third response option, "same" (indicating that the torque of the two systems was equal). Error bars indicate ±1 *SEM*.

(ANOVA), $F(1, 42) = 0.01$, $p > .250$ (see Fig. 2, left panel). However, an ANOVA controlling for pretest accuracy revealed that group did have a significant effect on post-test performance, $F(1, 41) = 5.21$, $p = .028$, $\eta_p^2 = .113$. Students in the action group showed a significantly nonzero (~10%) gain in accuracy from pretest to posttest, $t(21) = 3.07$, $p = .006$. Those in the observation group did not, $t(21) = -0.01$, $p > .250$ (see Fig. 2, left panel).[1]

Experiencing torque, felt as a resistive force, allowed students to investigate properties of angular momentum when aspects of the wheel system—the wheels' spin direction, spin speed, size, and tilt—were varied. For example, they experienced the fact that faster spin speed means larger angular momentum and, in turn, a greater resistive force. In the double-wheel system, the individual angular momentum vectors for the two wheels were combined, and the directions of the two vectors had to be taken into account. In Study 2, we extended our initial findings by exploring whether action experience specifically aids understanding of the vector nature of angular momentum.

Study 2

Our second study ($N = 36$; 16 males) replicated and extended Study 1. (We scheduled 40 subjects, but one member of the last pair did not show up at the lab, and 2 subjects were dropped because of computer error.) In order to investigate whether subjects learned about the vector nature of angular momentum, we included test trials in which wheels of the same size spun at the same speed but in opposite directions and adjusted training accordingly. In this special case, the individual angular momentum vectors for the two wheels cancel, and tilting the axle of the system results in zero resistive force. We categorized test trials according to whether students could consider only the magnitude of angular momentum to determine which avatar experienced more force (magnitude-dependent trials) or had to take into account the directional vector properties of one or both of the systems to respond correctly (vector-dependent trials, in which the wheels spun in opposite directions). We reasoned that the action group in particular should show learning on the vector-dependent trials (when the vector directions opposed one another), as they would feel this striking contrast when they handled the wheels.

Test trials in Study 2 again showed two avatars simultaneously, but Template manipulated one of two different systems of wheels. This allowed us to increase the number of trials to 126 (plus 6 practice trials). The apparatus Template manipulated remained constant in blocks of 15 or 16 trials (a red fixation cross cued the end of each block). The wheels' size, relative direction of spin (same or opposite), spin speed, and tilt rate (rather than tilt direction, as in Study 1) were manipulated. In this version of the test, both wheels in a given pair had the same size and spin speed. We also included no-spin trials in which Woody tilted an apparatus with wheels that were not spinning. These manipulated factors were crossed as fully as possible.

After each test trial, subjects answered the question, "Is Woody or the Template experiencing more force?" This was the same question used in Study 1, but with slightly different wording intended to ensure subjects' understanding. We also elaborated on the meaning of the question in the task instructions. The design of this test meant that in some trials, the two avatars both experienced zero force. Therefore, a third response option—"same"—was added and explained in the task instructions.

The training trials were updated from those in Study 1 to reflect the factors tested (e.g., differences in tilt rate), and subjects in the action group tilted the wheels seven times per trial rather than five. To control the tilt-rate manipulation during training, we asked subjects in the action group to tilt the wheels in time with a metronome.

As in Study 1, we found no group difference in pretest accuracy, $F(1, 34) = 0.37$, $p > .250$, but the action group

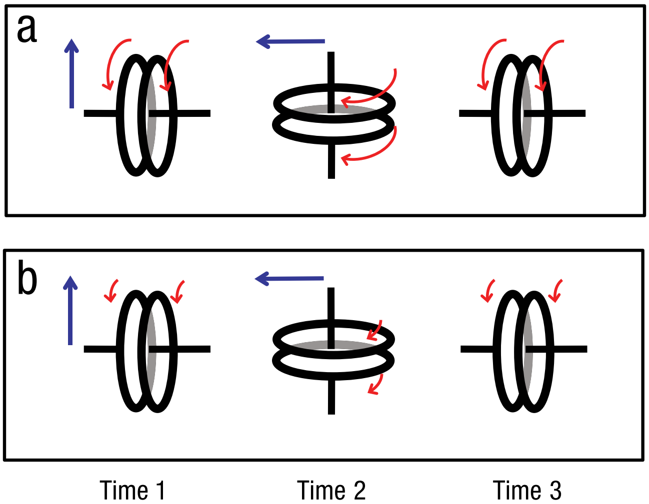

Fig. 3. Example problem from the angular momentum quiz completed outside the scanner in Study 3. In each problem, subjects were shown two sequences and asked to indicate which would result in more resistive force. The red arrows indicate the speed and direction of each wheel's spin. The blue arrows indicate the tilt rate and tilt direction for each system. In this example, the only difference between A and B is the speed at which the wheels are spinning; the wheels in A are faster, and the resistive force is therefore greater in A.

showed an advantage of approximately 10% in accuracy at posttest (controlling for pretest performance), $F(1, 33) = 5.09$, $p = .031$, $\eta_p^2 = .134$ (Fig. 2, right panel). This advantage was driven by the vector-dependent trials (those in which the wheels spun in opposite directions, resulting in cancellation). The action group significantly outperformed the observation group at posttest (controlling for pretest performance) on vector-dependent trials, $F(1, 33) = 5.67$, $p = .023$, $\eta_p^2 = .147$, but not on magnitude-dependent trials, $F(1, 33) = 0.35$, $p > .250$.[2] Thus, action experience enhances students' ability to account for the vector nature of angular momentum.

Study 3

Although Studies 1 and 2 demonstrated that action experience enhances students' understanding of angular momentum, it remained unclear how this enhancement occurs. In our third study, we used functional magnetic resonance imaging (fMRI) to investigate the neural correlates of learning. Thirty-five right-handed subjects (16 males) were scanned; we selected this sample size to closely match the sample size in Study 2 and other similar fMRI studies (Beilock et al., 2008). However, 2 subjects were dropped: One was unable to complete the experiment because of discomfort in the scanner, and the anatomical scan of another subject was lost as a result of a technical error.

Subjects first took a pretest and then completed action or observation training before being scanned while they completed the posttest. The training and tests were the same as those used in Study 2. The scanning allowed us to explore possible group differences in activation of sensory and motor brain regions resulting from training. Subjects completed four runs of fast event-related trials in the scanner (3-T Philips Achieva scanner, eight-channel Philips SENSE head-coil). Each run was self-paced, lasted for approximately 5 to 6 min, and consisted of 30 to 34 trials interspersed with fixation (128 trials in total). Each trial was triggered by the onset of the volume acquisition that followed the completion of the previous trial. Optseq software (http://surfer.nmr.mgh.harvard.edu/optseq/) was used to optimize the design. A T2*-weighted echo-planar imaging sequence was used to acquire functional images covering the whole brain (36 axial slices) with a repetition time of 2,000 ms. After the functional acquisition, a 7-min T1-weighted anatomical scan occurred.

Within 24 hr after scanning, all subjects were asked to complete an online follow-up quiz similar in content to what they might experience in a college-level introductory physics class. Before taking the quiz, the students watched a brief video in which a physics instructor explained the quiz and gave a thorough description of how to interpret the stimuli. As illustrated in Figure 3, two sequences were presented on each trial, and subjects were asked to compare the torque in the two sequences. Specifically, they were instructed: "Choose A if a person performing sequence A would experience more resistive force; Choose B if a person performing sequence B would experience more resistive force; Choose SAME if people performing sequences A & B would experience equal resistive force." Subjects completed 50 trials, across which the wheels' size, spin speed, spin direction, and tilt rate varied. The systems in some of the quiz problems closely resembled the bicycle-wheel system subjects encountered during training (as illustrated in Fig. 3), but other problems tested subjects on systems that looked different (e.g., a single sphere or globe rotating about an axle). Nineteen of the 33 subjects who completed the scanning session also completed this quiz. Subjects varied in how soon after scanning they completed the quiz, but all completed it within 24 hr. The timing did not differ between the two training groups and did not correlate with performance.

There were no group differences in accuracy on either the pretest or the posttest; there was only a main effect of test, $F(1, 31) = 31.72$, $p < .001$, such that both groups improved from the pretest to the posttest. There were no group differences in reaction time on either test.

Next, we looked at how the training groups differed in their quiz performance. The action group answered

74.5% (SE = 8%) of the questions correctly, and the observation group answered 52.2% (SE = 6.9%) of the questions correctly; the difference was significant, $F(1, 17) = 4.47$, $p = .049$, $\eta_p^2 = .208$.

For the neural data from the posttest, all preprocessing steps and whole-brain data analyses were conducted using BrainVoyager QX (Version 2.3.1; Brain Innovation, Maastricht, The Netherlands). Functional images were first slice-time-corrected and then motion-corrected using sinc interpolation; they were then spatially smoothed with a 4-mm full-width/half-maximum Gaussian kernel. Each functional run was manually aligned to the subject's 3-D anatomical image, and both were then transformed into Talairach space. We began analyzing the neural data by performing a group contrast at the whole-brain level to look for regions where the activation difference between problem trials and fixation (as assessed via the blood-oxygen-level-dependent, or BOLD, signal) differed between the action group and the observation group. Starting with a voxel-wise threshold of $p < .001$ at the whole-brain level, we ran a Monte Carlo simulation in BrainVoyager to set a minimum cluster threshold of 360 anatomical voxels (or ten 3 × 3 × 4 functional voxels) and applied this threshold to define the regions listed in Table 1. Regions defined by the group contrast were similar whether we included all 33 subjects who completed the scanning portion of the experiment or the subset of 19 who completed both scanning and the follow-up quiz (see Table 1). This contrast revealed greater activation for the action group compared with the observation group in several *predicted* regions, that is, regions known to be involved in action planning and production and to be sensitive to previous action experience (Beilock et al., 2008; Calvo-Merino et al., 2005; Cross, Hamilton & Grafton, 2006): dPMc, primary motor/somatosensory cortex (M1/S1), superior parietal lobe (SPL), supplementary motor area (SMA), and the cerebellum (see Fig. 4a).

Finally, we asked whether activation within the predicted brain regions found to differ as a function of training group predicted quiz accuracy (n = 19; see Table 1). Obtaining this result would support our hypothesis that the extent of the sensorimotor brain system's involvement is related to angular momentum understanding. Across the training groups, activation in the left M1/S1 region of interest functionally defined in the group contrast described earlier significantly predicted quiz performance ($r = .579$, $p = .009$; Fig. 4b). Using a bias-corrected bootstrapping procedure (1,000 resamples) with SPSS code (Preacher & Hayes, 2004), we found that activation in left M1/S1 mediated (accounted for) the influence of training on quiz performance (ab = 0.15; see Fig. 5); the indirect effect ($c - c'$) of training on quiz accuracy through M1/S1 activity was nonzero (95% CI = [0.0082, 0.3846]).

Activation in SMA/anterior cingulate cortex was also predictive of quiz scores across the training groups. However, the bootstrapping procedure (1,000 resamples) indicated that the indirect effect of this activation on quiz performance was not significantly different from zero.

The M1/S1 activation observed during the posttest likely had a bilateral component (which was detected at a more liberal voxel-wise threshold). Several other areas of significance also had a bilateral component detectable at a more liberal threshold, but there are reasons why the pattern of activation was stronger in the left hemisphere. First, all scanned subjects were right-handed. Second, the physical setup of the training was such that the actor's left hand anchored the set of wheels, while the right hand felt the brunt of the torque (the laser pointer in the axle and the vertical target line on the wall were on the actor's right-hand side). Though the actor's precise movements varied, this setup may have made the motor representation of changes in torque more salient in the right hand, arm, and shoulder, and consequently resulted in a left-lateralized neural signal. This pattern is typical, as noted in a recent meta-analysis by Caspers, Zilles, Laird, and Eickhoff (2010).

Our findings in Study 3 provide a causal model for how physical experience enhances understanding of angular momentum. Action experience (relative to observation) leads to increased activation of sensorimotor systems important for representing dynamic physical concepts. This activation, in turn, enhances understanding of torque and angular momentum (as assessed via our quiz). We have not yet ruled out the possibility that a more distributed pattern of activation plays a causal role (though this distributed pattern would likely involve M1/S1). Further research utilizing transcranial magnetic stimulation could bolster the specific causal model shown in Figure 5.

Study 4

In a final randomized field experiment in a college-level physics class, we asked whether the benefits of action experience could be seen on a quiz and homework completed several days after engaging with the bicycle-wheel system. Subjects in Study 4 were drawn from an introductory physics course at DePaul University. Ninety-seven students were enrolled in the course; 94 participated in the lab on the day of our experiment. For the analyses presented here, we excluded those students who had previous experience manipulating a similar system of bicycle wheels, those who failed to stick to their assigned roles during the lab activities, and those who did not attend the angular momentum lecture presented the same week as the lab. Fifty-seven students who met these criteria completed the quiz; these students were included in the analyses of quiz performance. Fifty-nine students who met these criteria completed the homework assignments; these students were included in analyses of homework performance.

Table 1. Results From Study 3: Clusters of Activation Resulting From Group Contrasts at the Whole-Brain Level and Correlations Between Activation and Quiz Accuracy

Region	Peak Talairach coordinates: all scanned subjects (N = 33)			Peak Talairach coordinates: subjects who completed the quiz (n = 19)			Correlation between activation and quiz accuracy (n = 19)
	x	*y*	*z*	*x*	*y*	*z*	
	Action group > observation group						
Predicted regions							
L dorsal premotor cortex				–19	–8	54	.355
R ventral premotor cortex				29	–5	42	.294
L primary motor/somatosensory cortex	–28	–41	63	–34	–23	57	.579*
R primary motor cortex				56	–11	39	.359
L superior parietal/secondary	–19	–65	30	–19	–65	30	.295
somatosensory cortex				–22	–65	54	.222
	–31	–50	54	–31	–50	54	.304
Cerebellum	20	–59	–27				
	5	–80	–18				
	–1	–56	–9	2	–53	–9	.323
	–4	–65	–39	–1	–65	–43	.277
	–25	–59	–27				
Supplementary motor area/anterior	–7	10	36	–7	4	33	.528*
cingulate cortex	26	34	12				
Additional regions							
R inferior frontal gyrus				41	7	27	
R cuneus				20	–71	18	
R superior temporal sulcus	47	–29	15				
R superior frontal gyrus	23	55	33				
Red nucleus	5	–20	–3	5	–23	–9	
L lingual gyrus	–13	–53	3	–28	–65	0	
	–37	–50	–3				
L insula	–34	–5	12	–37	–5	12	
	–31	–23	6				
L fusiform gyrus				–40	–38	–9	
L inferior temporal gyrus	–55	–50	–12				
	Observation group > action group						
L inferior parietal lobule	–52	–44	39				
R insula	47	–41	21				

Note: All clusters listed in the table were significant, $p < .001$ (cluster corrected). R = right; L = left.
*$p < .05$.

Working in 4-person lab groups (2 actors and 2 observers per group), students were randomly assigned to action or observation roles.[3] Each group was given one set of wheels. In an activity closely modeled after the training paradigm used in Studies 1 through 3, the students conducted experiments in which they sequentially compared how this system behaved as one relevant factor was changed at a time (e.g., they compared a fast tilt rate with an extremely slow one, keeping wheel size, spin speed, and spin direction constant). Each group worked through the set of four experiments with the wheels twice. First, 1 actor and 1 observer performed the experiments (Act 1). During this time, the students who were not currently playing the actor and observer roles were in charge of reading out the specifics for each trial and setting up the physical system appropriately (e.g., spinning the wheels at the correct speed). Following the experiments, all 4 students completed a worksheet designed to connect the experience (action or observation) with the bicycle wheels to the equations for angular momentum and torque (see Fig. 6b). The students were allowed to work together in their group, but were asked not to explicitly

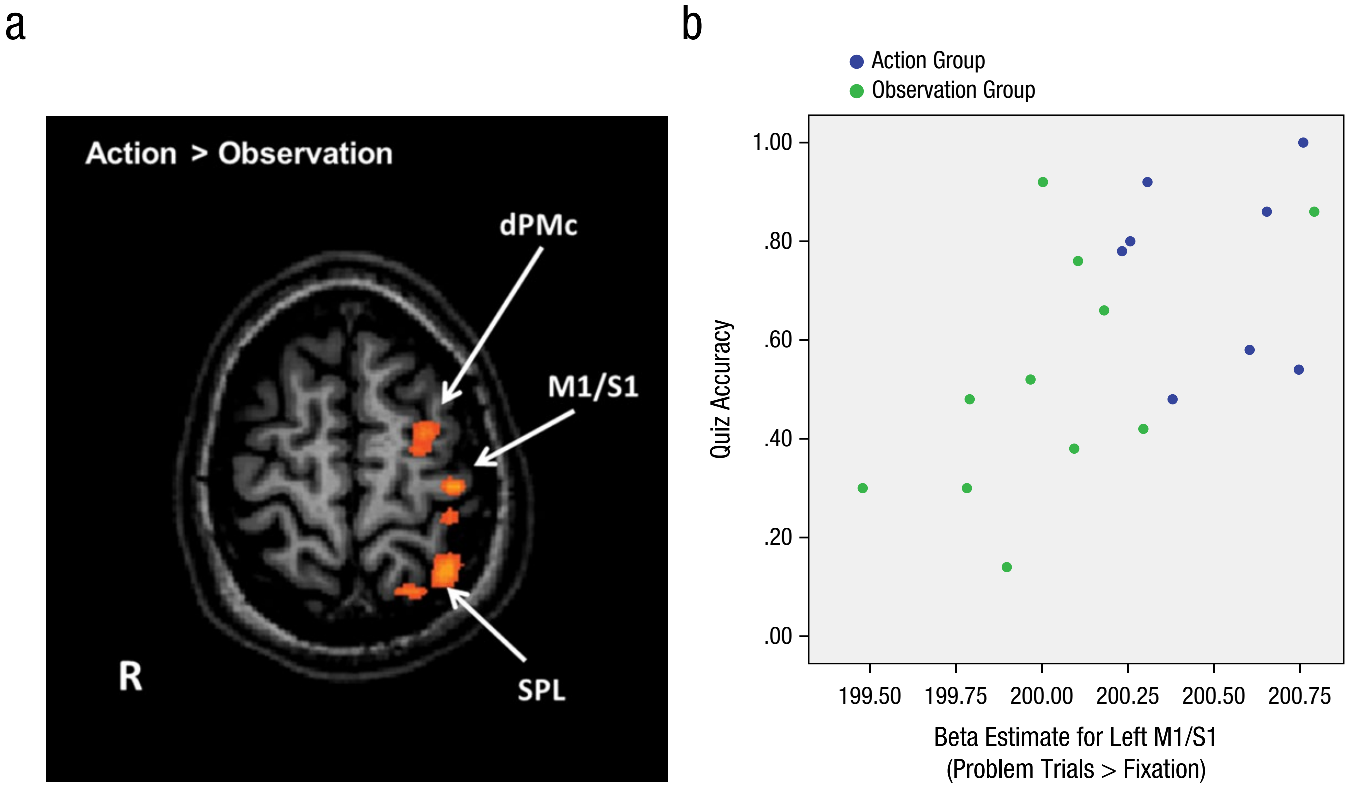

Fig. 4. Imaging results from the posttest in Study 3. The brain image (a), an axial slice at $z = 57$, shows a subset of neural regions in which the activation difference between problem trials and fixation was greater in the action group than in the observation group, for the 19 subjects who completed both the scanning and the online quiz. dPMc = dorsal premotor cortex; M1/S1 = primary motor/somatosensory cortex; SPL = superior parietal lobule; R = right. The scatter plot (b) shows the relation between accuracy on the online quiz and β weight for activation within the left M1/S1 region from the group contrast (problem trials > fixation); each plotted point represents a subject.

compare their experiences. Next, the students who had been the facilitators during Act 1 took on the actor and observer roles and performed the four experiments (Act 2). The students then completed the worksheets a second time. During the lab, students in the observation group were not allowed to handle the wheels, but group members were not prevented from working together in answering questions. An instructor guided this 2-hr exercise as part of normal classroom activities.

Each trial was contrasted with the one before it, with one factor changed at a time so as to emphasize that factor. Experiment 1 consisted of six trials testing the influence of spin speed on the angular momentum of the system (see Fig. 6a). Experiments 2 through 4 tested the influence of wheel size (i.e., moment of inertia), spin direction (i.e., vector cancellation), and tilt rate, respectively. The students were told that they would be graded only on the completeness of the worksheets, not on the consistency of their answers across time (i.e., answers should reflect current thinking). We found no effect of role order (i.e., serving as the actor or observer during Act 1 vs. Act 2) on any measure of accuracy in the classroom.

Several days after the lab, and after hearing a lecture on the material, students took a quiz on torque and angular momentum. The quiz was typical of the course, including multiple-choice, short-answer, and quantitative problems (for examples, see Fig. 6c). Placing our investigation within the context of a physics course allowed us to increase the variety and complexity of questions the students encountered. There were seven questions on the quiz, but some had multiple parts. In total, 17 question parts (problems) were scored. For the purposes of our data analysis, all problems were weighted equally in the scoring, and each problem was scored independently. This meant, for example, that if a student made an error in one problem and then carried an incorrect value into calculations for the next problem, he or she could still be scored as answering the second problem correctly if the calculations were carried out correctly. The scoring was strictly performance based. Students had to demonstrate both conceptual and computational competence to receive credit for the calculation-based problems and had to give sufficient justification for their answers to receive credit for the

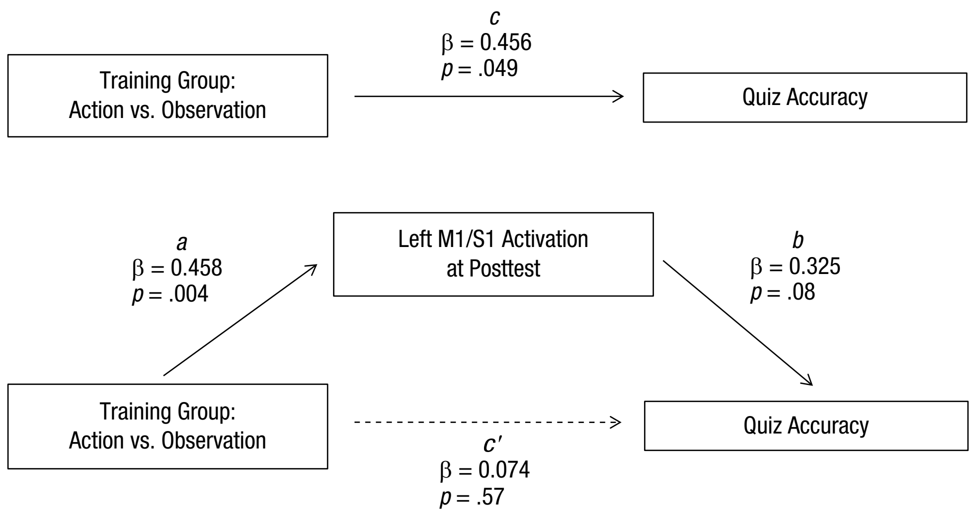

Fig. 5. Results from Study 3: effect of training group (action vs. observation) on quiz accuracy as mediated by activation in left primary motor/somatosensory cortex (M1/S1) during the posttest.

short-answer problems. Each multiple-choice problem had only one correct solution. Two experimenters scored the full data set, and interrater reliability was high, Cohen's κ = .96 across all problems. A third experimenter acted as a tiebreaker for scoring solutions to quantitative and short-answer problems. We also coded the rich data set from the quiz with a less strict coding scheme; the same pattern of results emerged.

Overall, the action group earned quiz grades that were about 7% higher than those of the observation group (76% vs. 69%), despite the fact that grades on other quizzes taken during the same school term were matched between the two groups. Figure 7 summarizes group performance for each problem on the quiz; students in the action group outperformed those in the observation group on 14 of the 17 problems (more often than would be predicted by chance, according to a nonparametric sign test on the mean group difference for each problem; p = .013).

In Study 2, we found that action experience was especially important for reasoning on vector-dependent trials. To further explore the impact of action experience on performance in the classroom, we compared students' accuracy on two quantitative problems that were well matched overall, but differed in vector dependence. Problems 11 and 14 both required students to calculate the angular momentum of a physical system involving spinning objects (Fig. 6c). The system in Problem 11 contained only one wheel (magnitude-dependent problem), whereas the system in Problem 14 contained two disks spinning in opposite directions (vector-dependent problem). The two groups performed similarly on the magnitude-dependent problem (action group: M = 94%, SEM = 4.1, 95% CI = [86%, 102%]; observation group: μ = 87%, SEM = 7.2, 95% CI = [73%, 101%]; each group's mean fell within the other group's 95% CI. However, the action group performed better than the observation group on the vector-dependent problem (action group: M = 68%, SEM = 8.1, 95% CI = [52%, 84%]; observation group: μ = 43%, SEM = 10.6, 95% CI = [22%, 64%]; each group's mean fell outside the other group's 95% CI. Thus, as in Study 2, the benefit of physically experiencing the properties of angular momentum was stronger when a question required application of vectors than when it did not.

We also wanted to compare students' performance when the angular momentum changed over time from an initial to a final state (dual-state problems) with their performance when the angular momentum did not change (single-state problems). The 17 problems were thus broken down into four types: magnitude-dependent single-state problems, magnitude-dependent dual-state problems, vector-dependent single-state problems, and vector-dependent dual-state problems. Problems of the last type were most complex and required students to consider both the direction of angular momentum vectors within the system and the changing orientation of vectors over time. An ANOVA with independent variables of temporal nature (single- vs. dual-state problems), vector nature (magnitude- vs. vector-dependent problems), and training group (action vs. observation) revealed that the action group's performance advantage on the quiz depended on the type of problem; that is, there was a three-way interaction of temporal nature, vector nature, and training group, $F(1, 55)$ = 4.11, p = .047. Results were in line with our prediction that the action group would have an advantage on problems that required the application of a

Act 1–Experiment 1: How is the spin speed of the wheels related to the total angular momentum of the bicycle wheels?

Use the set of big wheels for all of Experiment 1.

Note: After you complete each trial below, check off the box next to the trial to indicate its completion.

□1 (reference trial) First, neither wheel will spin. Enforcer, make sure the wheels are not spinning. Actor, you will tilt the axle at a quick pace (which you will keep consistent for the next several trials).

□2 Next, one wheel will spin at a slow speed and the other will not spin. Enforcer, spin just the left wheel (upward from your perspective) at a slow speed. Actor, tilt the axle at the same quick pace as before.

□3 Next, one wheel will spin at a fast speed and the other will not spin. Enforcer, spin just the left wheel upward at a fast speed. Actor, tilt the axle at the same quick pace as before.

□4 (reference trial) Neither wheel will spin. Enforcer, make sure the wheels are not spinning. Actor, tilt the axle at the same quick pace as before.

□5 (same as #2) One wheel will spin at a slow speed and the other will not spin. Enforcer, spin just the left wheel upward at a slow speed. Actor, tilt the axle at the same quick pace as before.

□6 (same as #3) One wheel will spin at a fast speed and the other will not spin. Enforcer, spin just the left wheel upward at a fast speed. Actor, tilt the axle at the same quick pace as before.

PAUSE here. Once all of the groups have reached this point, the instructor will give a brief demonstration to reiterate the appropriate tilt and spin speeds. The instructor will then ask everyone to move on to Experiment 2 together.

b

3. a) What physics quantity was varied in Experiment 2, when we changed the size of the bicycle wheels? (circle one)

I Δt magnitude of $\vec{\omega}$ direction of $\vec{\omega}$

b) Based on the definition of (equation for) angular momentum, how did varying the size affect the total angular momentum of the bicycle wheels? Briefly justify your answer.

c

11. A single bicycle wheel is spinning about its axle at an angular speed of 8.7 rad/s. The wheel has a moment of inertia of 1.6 kg·m^2.

What is the magnitude of the angular momentum of the wheel?

14. Two disks are free to rotate about the same axle. The moment of inertia of disk 1 is 4.2 kg·m^2, and the moment of inertia of the disk 2 is 1.4 kg·m^2. Disk 1 spins counterclockwise with an angular speed of 2.7 rad/s, and disk 2 spins clockwise with an angular speed of 8.1 rad/s.

What is the magnitude of the angular momentum of the two-disk system?

Fig. 6. Materials used in the physics classroom for Study 4: (a) excerpt from the lab exercise, (b) example question from the worksheets completed after the experiments, and (c) example problems from the classroom quiz. Problem 11 was categorized as magnitude dependent, and Problem 14 was categorized as vector dependent.

multiple vectors: The action group outperformed the observation group by about 12 percentage points on vector-dependent dual-state questions (Problems 12, 15, 16, and 17; action group: 27% correct, observation group: 15% correct). Note that the interaction was such that the action group had an advantage for three of the four

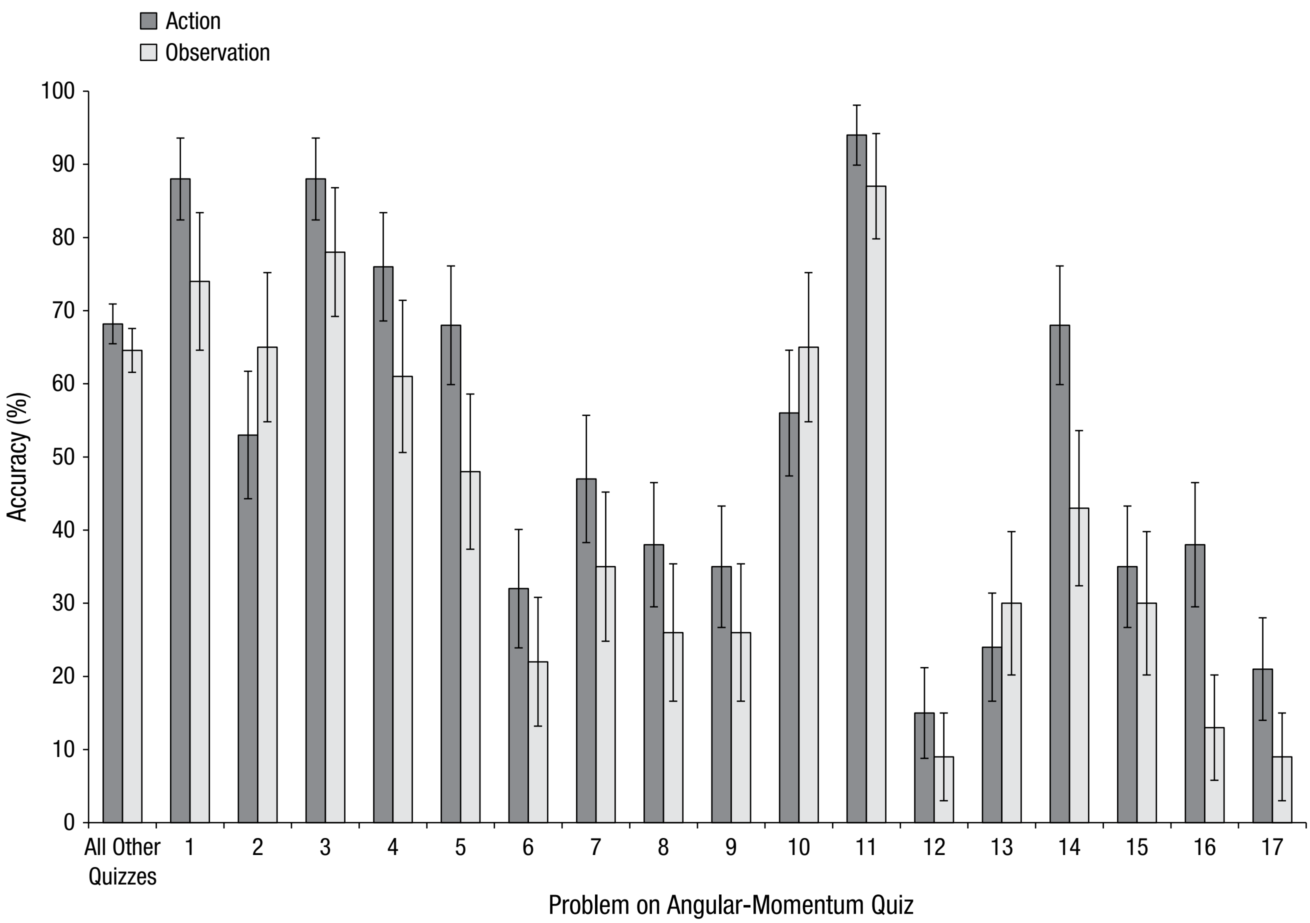

Fig. 7. Results from Study 4: the action group's and observation group's performance on each problem on the angular momentum quiz and on all other quizzes taken during the same course. Error bars represent ±1 *SEM*.

categories (magnitude-dependent single-state and vector-dependent single-state problems, in addition to vector-dependent dual-state problems).

The quiz was a true snapshot of students' understanding after the lab and lecture, as it was their first individual encounter with problems on angular momentum. In addition, students received homework problems at the end of the week after the quiz, which was a change from practice as usual (i.e., usually students received homework assignments before the quiz on a unit). The homework had 12 relevant problems. (Three problems were excluded from our analyses because they did not relate to the angular momentum topic or they were designed with several parts that built on one another and students received feedback as they worked through the steps.) As for the quiz, all problems were weighted equally in the scoring. The online homework system allowed students to attempt a problem multiple times so that they could learn as they worked on the assignment, but only first answers were included in the analyses reported here. Each student received different values for the variables in the problems, so the correct answers were different for different students, but students were allowed to work together and use their notes. The homework assignment was due about 1 week after the training experience and a few days after the quiz. Because the students likely learned from taking the quiz (McDaniel, Roediger, & McDermott, 2007), the homework data should be viewed with the caveat that they do not necessarily represent a "pure" measure of how students' knowledge varied as a function of whether they played the role of actor or observer.

Nonetheless, when we categorized the 12 problems of interest according to their temporal and vector nature, as we did for the quiz problems, we found a pattern of results similar to that observed on the quiz: The benefit of physical experience was most evident for complex problems that required consideration of multiple vectors. The action group outperformed the observation group by about 12 percentage points on the vector-dependent dual-state homework problems (action group: 49% correct; observation group: 37% correct). The two groups did not differ significantly in their average performance on other homework assignments for the course (action group: 86%; observation group: 84%). When we excluded 5 students who completed the homework but not the quiz (n = 54), the same pattern of results for the homework problems emerged.

In summary, in an authentic classroom environment, students who felt the consequences of the vector nature of angular momentum outperformed students who observed the same phenomena.

Discussion

Recent technological advances have prompted the implementation of online classrooms and virtual laboratories in science education. As a result, it is necessary to understand when direct experience with the physical world is beneficial for learning and how to integrate the tenets of embodied cognition into virtual learning environments (de Jong, Linn, & Zacharia, 2013; Han & Black, 2011; Lindgren & Johnson-Glenberg, 2013; Pouw, van Gog, & Paas, 2014). We have shown that brief, meaningful physical experience with science content enhances learning by activating sensorimotor brain systems used to execute similar actions in the past. These findings specify a mechanism underlying the value of physical experience in science education, and lead the way for classroom practices in which experience with the physical world is an integral part of learning. Our data show the value of a match between the sensory and motor input experienced during learning and the information needed at test, and they suggest that science concepts involving kinetics may be particularly well suited for learning via physical experience. When physical experience is closely tied to the to-be-learned content, subsequent activation of sensory and motor systems can effectively support students' reasoning. This kind of experience may be most influential in the initial stages of learning, when students are resolving misconceptions (Zacharia, Loizou, & Papaevripidou, 2012), and in areas of science in which kinetics come into play (e.g., physics, engineering, and chemistry).

Author Contributions

C. Kontra co-designed and performed all the experiments, analyzed the data, and cowrote the manuscript. D. J. Lyons performed Experiment 4 and analyzed data from that experiment. S. M. Fischer and S. L. Beilock co-designed all the experiments, supervised the project, and cowrote the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was supported by the National Science Foundation through CORE DUE-1348614, FIRE DRL-1042955, and the Spatial Intelligence and Learning Center (SBE-0541957 and SBE-1041707), as well as by the Spencer Foundation.

Open Practices

All data and materials have been made publicly available via the Harvard Dataverse Network and can be accessed at http://thedata.harvard.edu/dvn/dv/physics. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/view/ and http://pss.sagepub.com/content/25/1/3.full.

Notes

1. A repeated measures ANOVA on accuracy for Question 1 revealed a main effect of test (pretest vs. posttest), $F(1, 42) = 4.25$, $p = .045$, qualified by a significant two-way interaction of test and group, $F(1, 42) = 4.33$, $p = .044$. A repeated measures ANOVA on accuracy for Question 2 revealed a nonsignificant interaction and no main effects. There were no significant effects on reaction time for either question.
2. A repeated measures ANOVA on accuracy revealed a main effect of test, $F(1, 34) = 9.38$, $p = .004$, qualified by a significant two-way interaction of test and group, $F(1, 34) = 5.42$, $p = .026$. The only significant effect on reaction time was a main effect of test, $F(1, 34) = 7.50$, $p = .010$; both groups responded more quickly on the posttest compared with the pretest.
3. During the lab the week after our experiment concluded, all students were given action experience (a chance to manipulate the bicycle wheels).

References

Alberts, B. (2013). Prioritizing science education. *Science*, *340*, 249.

Allen, I. E., & Seaman, J. (2010). *Class differences: Online education in the United States, 2010*. Babson Park, MA: Babson Survey Research Group.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*, 84–91.

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences, USA*, *105*, 13269–13273.

Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, *15*, 1243–1249.

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage*, *50*, 1148–1167.

Cross, E. S., Hamilton, A. F. de C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage*, *31*, 1257–1267.

Cummings, K. (2008, October). *The Rensselaer studio model for learning and teaching: What have we learned?* Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC.

de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, *340*, 305–308.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, USA*, *111*, 8410–8415.

Glenberg, A. M. (1997). What memory is for [Target article and commentaries]. *Behavioral & Brain Sciences*, *20*, 1–55.

Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. *Computers & Education*, *57*, 2281–2290.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher*, *42*, 445–452.

McDaniel, M. A., Roediger, H. L., III, & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, *14*, 200–206.

Meltzer, D. E., & Thornton, R. K. (2012). Resource Letter ALIP-1: Active-learning instruction in physics. *American Journal of Physics*, *80*, 478–496.

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *351*, 1002–1005.

Pouw, W. T. J. L., van Gog, T., & Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educational Psychology Review*, *26*, 51–72.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717–731.

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: Office of the President.

Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.

Zacharia, Z. C., Loizou, E., & Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Childhood Research Quarterly*, *27*, 447–457.

Zwann, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1–11.