Psychological Science

From Action to Abstraction: Using the Hands to Learn Math
Miriam A. Novack, Eliza L. Congdon, Naureen Hemani-Lopez and Susan Goldin-Meadow

The online version of this article can be found at:

Published by:

$SAGE

On behalf of:

aps
ASSOCIATION FOR
PSYCHOLOGICAL SCIENCE

Association for Psychological Science

Additional services and information for *Psychological Science* can be found at:

Email Alerts: http://pss.sagepub.com/cgi/alerts

Subscriptions: http://pss.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> OnlineFirst Version of Record - Feb 6, 2014

What is This?

From Action to Abstraction: Using the Hands to Learn Math

Miriam A. Novack, Eliza L. Congdon, Naureen Hemani-Lopez, and Susan Goldin-Meadow
The University of Chicago

Abstract
Previous research has shown that children benefit from gesturing during math instruction. We asked whether gesturing promotes learning because it is itself a physical action, or because it uses physical action to represent abstract ideas. To address this question, we taught third-grade children a strategy for solving mathematical-equivalence problems that was instantiated in one of three ways: (a) in a physical action children performed on objects, (b) in a concrete gesture miming that action, or (c) in an abstract gesture. All three types of hand movements helped children learn how to solve the problems on which they were trained. However, only gesture led to success on problems that required generalizing the knowledge gained. The results provide the first evidence that gesture promotes transfer of knowledge better than direct action on objects and suggest that the beneficial effects gesture has on learning may reside in the features that differentiate it from action.

There is ample experimental evidence that performing actions on the external world affects one's internal representations (e.g., James, 2010; Kontra, Goldin-Meadow, & Beilock, 2012; Sommerville & Woodward, 2010; Wilson, 2002). Recently, there has been growing interest in a special kind of action—the actions people produce when talking, often called gestures (McNeill, 1992). Like other forms of action, gestures involve movements of the body, but whereas other forms of action have a direct effect on the world, gestures do not (e.g., twisting a jar lid results in an open jar; gesturing a twisting movement does not). Yet gestures, like other forms of action, can have profound effects on thinking and learning (Goldin-Meadow, 2003). These observations raise questions about how the fundamental differences between gesture and action[1] affect learning. In the study reported here, we compared the consequences of learning a concept by directly acting on objects as opposed to gesturing about those objects. Our goal was to better understand how these two types of movements affect learning.

One critical feature of action-based learning is that it often involves manipulating or interacting with physical objects. For example, a child using base-10 blocks to learn how to carry addends in an addition problem literally carries the blocks from one column to another. Traditional theories of cognitive development suggest that children can often succeed in solving problems with physical objects before they succeed with symbolic representations (Bruner, 1966; Piaget, 1953). External symbols can be acted upon, which allows motor processes to be integrated with abstract ideas (Lakoff & Nunez, 2000). Yet despite the theoretical benefits of action, findings regarding the effectiveness of encouraging learners to act on concrete objects have been mixed (see McNeil & Uttal, 2009; Mix, 2010; Sarama & Clements, 2009). Children can become preoccupied with irrelevant details of perceptually rich symbols (McNeil, Uttal, Jarvin, & Sternberg, 2009) or see their actions as relevant only to the objects on which they were trained rather than to a

Corresponding Author:
Miriam A. Novack, Department of Psychology, 5848 S. University Ave., University of Chicago, Chicago, IL 60637
E-mail: mnovack1@uchicago.edu

more general concept (Uttal, Scudder, & DeLoache, 1997). If physical objects focus a child's attention on irrelevant aspects of a procedure rather than on the concept underlying the procedure, then the child may not be able to generalize what he or she has learned to a new context (Mix, 2010). Action may thus be helpful in teaching children to solve a particular problem, but may fare less well in teaching them to extend that knowledge to new problems.

Gesture, like action, is an act of the body, but unlike action, it does not involve directly manipulating objects. Rather, gestures are representational hand movements that vary in how veridically they represent actions. Gestures can represent many aspects of the action they reference (e.g., pantomiming twisting a jar lid), or they can selectively represent only the movement of the action (e.g., tracing the circular path of the lid). Chu and Kita (2008) examined the different types of gestures people spontaneously produced while explaining their strategies in performing a mental rotation task. Adults initially used mimetic gestures, miming the act of rotating the object, but over the course of the experimental session, they used more and more abstract gestures, tracing the path of the object as it rotated. Chu and Kita hypothesized that this shift reflected an internalization of the action strategy, which had become less tied to the concrete rotation action and thus more abstract.

Goldin-Meadow, Cook, and Mitchell (2009) used an abstract gesture to help 9- to 10-year-old children solve mathematical-equivalence problems such as $4 + 3 + 6 =$ __ $+ 6$. Children were taught to produce a V-point gesture to the first two numbers (the 4 and the 3 in this instance) and then to point at the blank. These movements, which were modeled after the spontaneous gestures of children who know how to solve these problems correctly (Perry, Church, & Goldin-Meadow, 1988), were designed to help the children see that the problems can be solved by grouping and adding the two numbers on the left side of the equation that do not appear on the right side and then putting the sum in the blank. Children who were asked to produce these hand movements during a math lesson were able to extract the grouping strategy despite the fact that they were never explicitly told what the movements represented. Surprisingly, even children who were trained to make a V-point to the wrong addends (3 and 6 in this example) learned grouping, although less well than children trained to point to the correct numbers. Gesture's power as a teaching tool may thus reside not only in its ability to focus children's attention on a specific aspect of a problem, but also in its ability to convey substantive ideas about the relational structure of the problem.

These findings make it clear that children can glean novel insights from an abstract gesture. But it is not yet known whether abstract gesture can support generalization beyond the particular problem on which it was taught, nor whether abstract gesture is more (or less) effective as a teaching tool than concrete action. Our study was designed to address these questions.

We used the procedure from Goldin-Meadow et al. (2009) to investigate whether varying the concreteness of the movements children were trained to produce during a math lesson would influence their ability to generalize the knowledge they had gained to new problem forms. Each child was assigned to one of three training conditions: In the action condition, children were taught to pick up the magnetic number tiles placed over the first two numbers in the problem (magnetic tiles had been placed on all of the numbers), to group the tiles in one hand, and then to hold the hand over the blank. In the concrete-gesture condition, children were taught to mime the action of picking up, grouping, and moving the number tiles, but without ever actually touching the tiles. In the abstract-gesture condition, children were taught to produce a V-point gesture to the first two number tiles and then to point at the blank.

If action-based instruction is the best way to scaffold learning an abstract concept, then we would expect rates of learning on trained problems to be higher in the action condition than in either of the gesture conditions. But if the procedure instantiated in action is too tied to the particulars of the problem, then we would expect children in the action condition to fail to generalize what they had learned. In contrast, if abstract gesture focuses attention on the aspects of the procedure particularly relevant to solving the problem (and away from aspects that are irrelevant), we would expect abstract gesture to be effective not only in teaching children how to solve the problems on which they were trained, but also in helping them generalize this knowledge to new problem forms. We were agnostic about the effects of concrete gesture. Concrete gesture replicates many of the features of action and therefore may behave like action, perhaps tying knowledge to the particular problem on which it was learned. However, because concrete gesture does not involve manipulating objects, it might free knowledge from a particular context so that it can be more easily generalized to new problem forms.

Method

Participants

One hundred forty-two third-grade children were tested in their elementary schools in the Chicagoland area; third-grade children were studied because children of this age typically fail to solve mathematical-equivalence addition problems correctly, and are therefore an ideal

population within which to explore the impact of different types of instruction on learning. To ensure that the participants in all conditions were equally naive with respect to the mathematical-equivalence problems, we excluded all children who solved any of the six problems on the pretest correctly ($n = 44$).[2] In addition, 1 child was excluded because of a language barrier, and 7 were excluded because of experimenter error. The remaining 90 children (60 females, 30 males) ranged in age from 8.46 to 10.48 years ($M = 9.23$ years, $SD = 0.38$) and were racially and ethnically diverse. All participants were tested in the schools they attended, which were recruited through phone calls and e-mails to principals. Children's parents had submitted written consent prior to the children's participation in the study. The children were randomly assigned to one of three conditions: action ($n = 30$), concrete gesture ($n = 31$), or abstract gesture ($n = 29$). Children received a pencil, stickers, and a certificate for participating, and the teachers of participating classrooms received a $40 gift certificate to a local learning store.

Materials

Math problems were written in black marker on a white, portable, dry-erase magnetic board. During training, black magnetic number tiles were placed over the numbers in the problem. There were two forms of math problems on the pretest and posttest. In the ABC problems, the last number on the left side of the equation (c) was repeated on the right side (e.g., $a + b + c = \_\_ + c$); in the PQR problems, the first number on the left side (p) was repeated on the right side (e.g., $p + q + r = p + \_\_$).

Design and procedure

The children were tested individually in an unused room at their school. The procedure was divided into five parts: pretest, preinstruction, instruction, posttest, and generalization test. Experimenter A administered the pretest, posttest, and generalization test. Experimenter B administered the preinstruction and instruction. Experimenter A was not present during instruction and was therefore blind to the training condition.

Pretest. All children first completed a written pretest containing six problems, three in the ABC form and three in the PQR form. After a child solved the problems, Experimenter A wrote the problems on the dry-erase board one at a time and asked the child to explain how he or she got each answer.

Training. The training procedure consisted of preinstruction and instruction. The purpose of the preinstruction (three problems in ABC form) was to teach the

children the words and hand motions they would be asked to repeat during instruction. In all conditions, the experimenter wrote the problems on the board one by one (e.g., $2 + 9 + 4 = \_\_ + 4$), covering all the numbers with corresponding black magnetic number tiles. All the children were taught to look at a problem and say, "I want to make one side equal to the other side," the equalizer strategy (Perry et al., 1988). They were also taught hand movements, which differed by training condition (see Fig. 1). On each of the three preinstruction problems, the experimenter demonstrated the words and hand movements and then asked the child, "Can you say those words and do those hand movements for me?" Neither the experimenter nor the child actually solved any of the preinstruction problems.

During instruction, the experimenter and child alternated in solving 12 ABC problems on the dry-erase board; we refer to ABC problems as *trained* problems. When it was the experimenter's turn, she wrote a problem on the board, covered the numbers with magnetic tiles, and wrote the correct answer in the blank. She then explained how to solve the problem using the equalizer strategy, but did not move her hands at all. For example, for the problem $6 + 3 + 8 = \_\_ + 8$, the experimenter put 9 in the blank and said, "I want to make one side equal to the other side. You see, 6 plus 3 plus 8 is 17, and 9 plus 8 is 17, so one side is equal to the other side." The experimenter then wrote a new problem on the board and covered all the numbers with corresponding magnetic tiles. The child was asked to produce the words and hand movements he or she was taught in preinstruction and was then asked to solve the problem. The child was told whether the answer was right or wrong, but was not given the correct answer if the answer was incorrect. The child's solution was then erased, and the child was asked to repeat the words and hand movements taught during preinstruction.

Posttest. After instruction, the children were given a second six-question paper-and-pencil test to measure immediate improvement. This test had three ABC and three PQR problems and thus was comparable to the pretest.[3] As in the pretest, the children first completed the problems and then explained their solutions at the dry-erase board. Because the PQR problems were seen only in the pretest and posttest and not during training, we refer to them as *near-transfer* problems.

Generalization test. Following the posttest, each child was given a paper-and-pencil generalization test. The test contained two problems that could not be solved using the grouping strategy but required a deeper understanding of mathematical equivalence (e.g., in the problem $2 + 5 + 3 = \_\_ + 6$, none of the addends on the left side of the

Action

"I want to make one side [while simultaneously picking up the first two number tiles, the 2 and 9 in this example] equal to the other side [while simultaneously placing the number tiles on the blank space]."
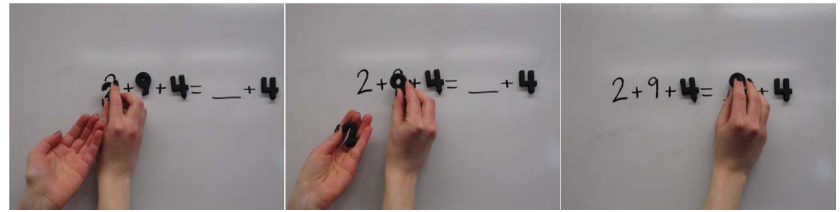
Concrete Gesture

"I want to make one side [while simultaneously moving their hands as if picking up the first two number tiles, but not physically touching the pieces] equal to the other side [while simultaneously pretending to place the number tiles on the blank space without physically moving the tiles]."
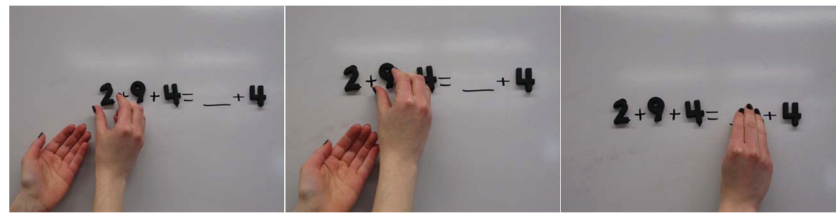
Abstract Gesture

"I want to make one side [while simultaneously producing a V-point gesture under the first two numbers] equal to the other side [while simultaneously pointing with their index finger to the blank space]."
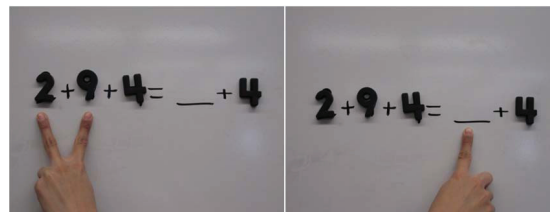
Fig. 1. Illustrations of the hand movements taught during the preinstruction phase of the training procedure to children in the action, concrete-gesture, and abstract-gesture conditions.

equation are repeated on the right). We refer to these problems as *far-transfer* problems.[4]

Coding

The entire session was videotaped. Gesture and speech during children's pretest and posttest explanations were transcribed and coded according to a previously established coding system (Perry et al., 1988). Three independent coders determined whether or not the equalizer strategy was produced in speech in explanations of posttest trials. Each posttest trial was coded by two of the three coders, and each coder coded two thirds of the total trials. Reliability was between 94% and 97% for all three pairs of coders (kappas ranged from .88 to .93).

Results

The histograms in Figure 2 show the distribution of children (in the three conditions combined) according to the number of problems they solved after instruction, separately for the trained, near-transfer, and far-transfer problems (recall that these children did not solve any problems successfully before instruction). Because the distribution for each of the three problem types was nonnormal and bimodal (i.e., children tended to either get all the problems wrong or get all the problems right), we gave children credit for solving the problems of a given type if they produced a correct answer, or an answer that was off by 1 but accompanied by a correct explanation, on all exemplars of that problem type. We classified children who met this criterion for trained problems as learners. We then considered whether these learners extended their understanding to near- and far-transfer problems. Learners were considered to have exhibited near transfer if they produced a correct answer, or an answer that was off by 1 but accompanied by a correct explanation, on all exemplars of near-transfer problems. Learners were considered to have exhibited near and far transfer if they met this same criterion for all near- and far-transfer problems.
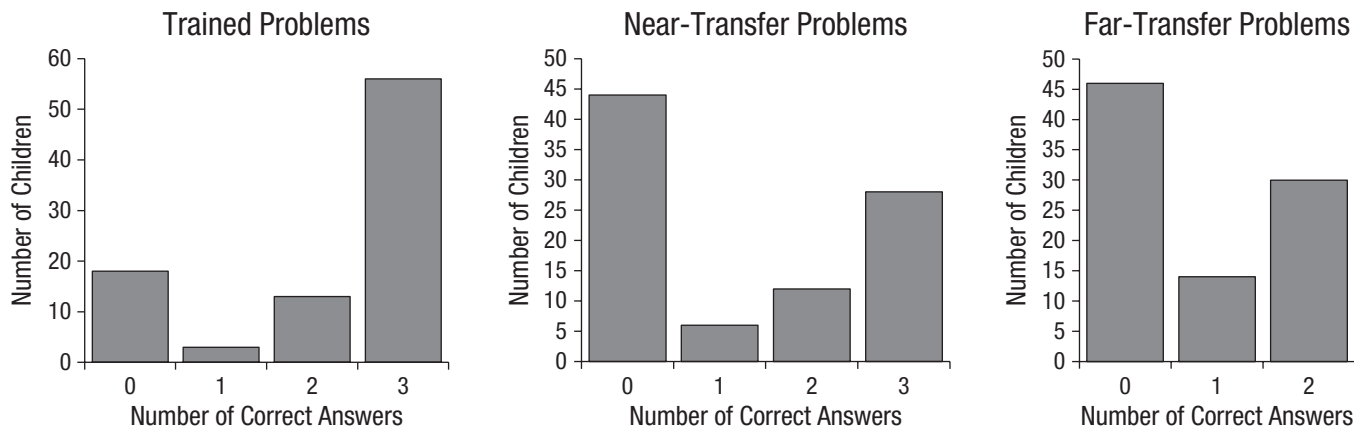
Trained Problems

Near-Transfer Problems

Far-Transfer Problems

Fig. 2. Distribution of children (across all conditions) according to the number of problems solved correctly after instruction. Results are presented separately for trained, near-transfer, and far-transfer problems.

Children in all three conditions were equally likely to be classified as learners (see Fig. 3a). In a fixed-effects binomial logit model, we used condition (action, concrete gesture, abstract gesture) to predict success on the trained problems (i.e., being classified as a learner; the abstract-gesture condition was used as the baseline, as it was likely to have the lowest rates of learning). There was no effect of condition on success on trained problems (action condition: $\beta = -0.22$, $z = -0.41$, $p = .68$; concrete-gesture condition: $\beta = 0.71$, $z = 1.27$, $p = .20$); children did not display significantly different odds of succeeding on the trained problems as a function of condition. Releveling the model with the action condition as the baseline revealed no difference between the action and abstract-gesture conditions ($\beta = 0.22$, $z = 0.41$, $p = .68$), but a marginal difference between the action and concrete-gesture conditions ($\beta = 0.92$, $z = 1.68$, $p = .09$).

We next focused on success on the near- and far-transfer problems, and did so only for learners, children who had successfully solved all the trained problems ($n = 16$ for the action condition, $n = 23$ for the concrete-gesture condition, $n = 17$ for the abstract-gesture condition). Focusing first on the near-transfer problems, we used a fixed-effects binomial logit model to predict learners' probability of getting all near-transfer problems correct (see Fig. 3b). We found a main effect of training condition: Learners in the action condition had significantly lower odds of mastering near-transfer problems than did learners in the abstract-gesture condition ($\beta = -1.58$, $z = -1.97$, $p = .049$), but learners in the concrete-gesture condition did not differ from learners in the abstract-gesture condition ($\beta = -0.031$, $z = -0.05$, $p = .962$). Releveling the model to make the action condition the baseline, we found that learners in both the concrete-gesture ($\beta = 1.55$, $z = 2.03$, $p = .04$) and the abstract-gesture ($\beta = 1.58$, $z = 1.97$, $p = .049$) conditions performed significantly

better than learners in the action condition. Overall, learners in the action condition were less likely to generalize the knowledge they had gained on the trained problems to the near-transfer problems than were learners in either gesture condition, but there was no difference between learners in the two gesture conditions.

Finally, we examined performance on the far-transfer problems (see Fig. 3c). We again used a fixed-effects binomial logit model, this time predicting the log-odds of being a learner who exhibited both near and far transfer. We again found a main effect of training condition: Children in the action condition had significantly lower odds of success on all three problem types than did children in the abstract-gesture condition ($\beta = -2.06$, $z = -2.30$, $p = .02$), and children in the concrete-training condition had marginally lower odds of success than did children in the abstract-gesture condition ($\beta = -1.15$, $z = -1.71$, $p = .09$). When we releveled the model to make the action condition the baseline, we found that the concrete-gesture condition was not statistically different from the action condition ($\beta = 0.90$, $z = 1.01$, $p = .31$), but the abstract-gesture condition was significantly better than the action condition ($\beta = 2.06$, $z = 2.30$, $p = .02$). The pattern of results displayed in Figure 3 suggests that abstract-gesture training was the most effective in encouraging learners to generalize the knowledge they had gained during instruction, action training was the least effective, and concrete-gesture training was somewhere in between.

To further explore the impact of gesturing on children's understanding of the principle underlying mathematical equivalence, we examined children's spoken explanations on the posttest. Children in all three conditions were trained during preinstruction to parrot the words of the equalizer strategy (a strategy based on the principle that the two sides of an equation must be equivalent). We examined whether the hand movements
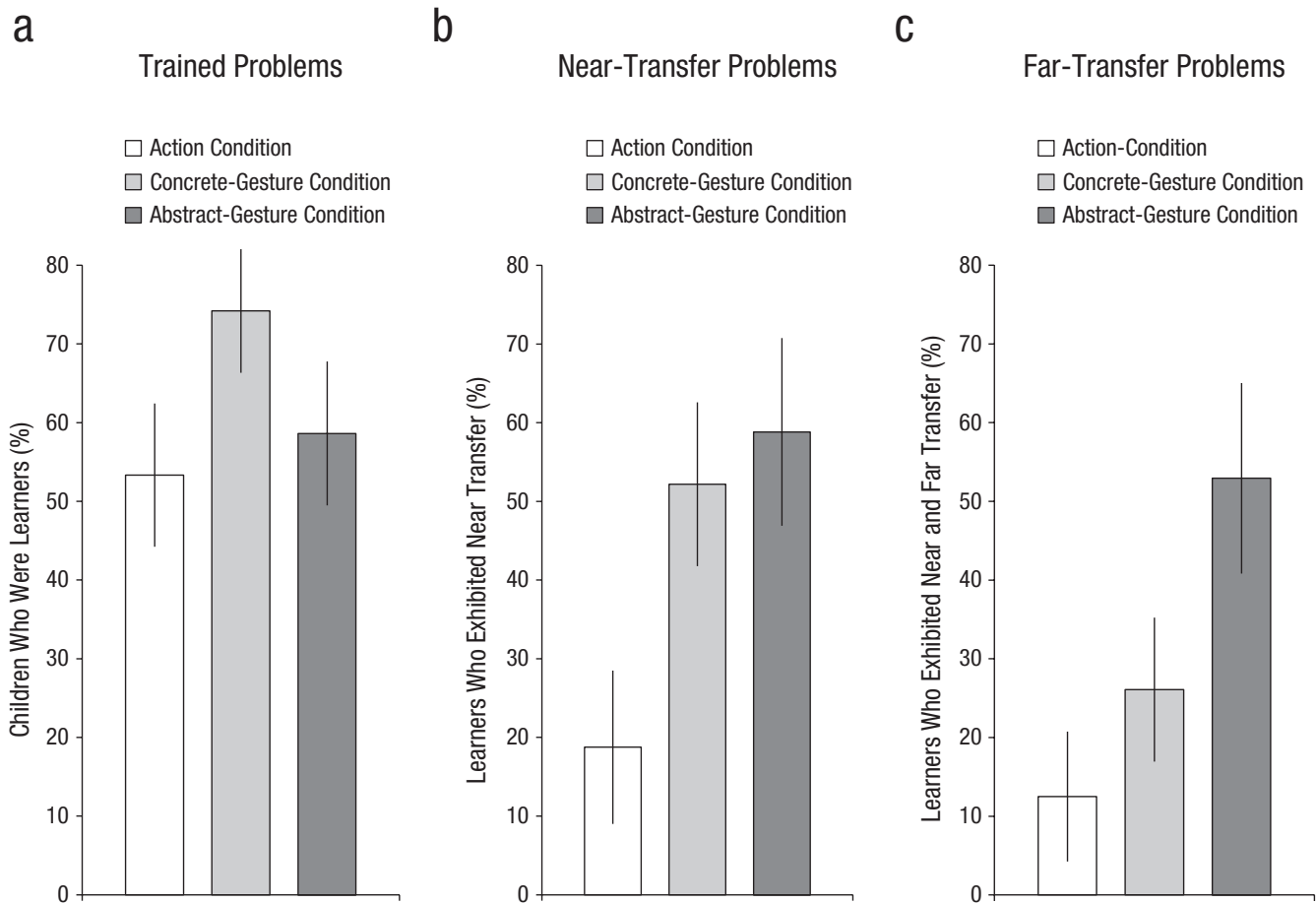
a

Trained Problems

☐ Action Condition
▨ Concrete-Gesture Condition
▧ Abstract-Gesture Condition

b

Near-Transfer Problems

☐ Action Condition
▨ Concrete-Gesture Condition
▧ Abstract-Gesture Condition

c

Far-Transfer Problems

☐ Action-Condition
▨ Concrete-Gesture Condition
▧ Abstract-Gesture Condition

Fig. 3. Children's success on the three problem types by condition: (a) the percentage of children who were learners (i.e., given credit for solving all the trained problems), (b) the percentage of learners who were given credit for solving all the near-transfer problems, and (c) the percentage of learners who were given credit for solving all the near- and far-transfer problems. See the text for an explanation of the criteria for being given credit for solving the problems of a given type. Error bars represent the standard errors of the percentages.

the children produced along with the words they parroted during instruction (i.e., their training condition) influenced how likely they were to glean meaning from those words. To determine whether the children gleaned meaning from the parroted equalizer strategy, we correlated the number of times a child's posttest explanations referred to the equalizer strategy with that child's success on the posttest problems: 0 = no success on any problems; 1 = learners (i.e., success on only trained problems); 2 = learners who exhibited near transfer (i.e., success on trained and near-transfer problems); and 3 = learners who exhibited near and far transfer (i.e., success on trained, near-transfer, and far-transfer problems). We found a significant positive relation between the two measures only for children who had repeated the equalizer strategy while producing a concrete ($r = .76$, $p < .001$) or abstract ($r = .64$, $p < .001$) gesture during instruction, and not for children who had repeated the words

while acting on the numbers ($r = .25$, $p = .17$). These results suggest that the children gleaned meaning from the words they were trained to say if they were required to produce gestures (but not actions) along with those words.

Discussion

Previous research has shown that action and gesture support learning across a variety of ages and contexts. Here we exploited the similarities and differences between action and gesture to compare their effects on learning. We found that acting gave children a relatively shallow understanding of a novel math concept, whereas gesturing led to deeper and more flexible learning. Furthermore, the form of the gesture mattered: Abstract gesture facilitated generalization, whereas concrete gesture, which mimicked the hand movements of action, brought learners to an

intermediate stage of conceptual understanding. Learning in this task was thus correlated with the concreteness of the hand movements children produced during training. The more concrete movements seemed to tie the knowledge gained to the training context, which suggests that the beneficial effects gesture has on learning may stem not only from gesture's base in action, but also from its ability to abstract away from action.

Our findings are consistent with the literature on abstraction. For example, the *concreteness-fading* theory proposes that learning is best supported by first introducing concrete representations and then moving on to more symbolic or abstract representations (Goldstone & Son, 2005; McNeil & Fyfe, 2012). This theory raises the possibility that the movement-based learning in our mathematics paradigm could have been even more successful had we used a within-subjects training design in which we provided learners with increasingly abstract representations of the grouping strategy. Would children have learned even better had they begun with action and then moved to gesture in a second training session? A fading technique of this sort might be particularly effective with children who are struggling with the concept being taught.

Action did not promote generalization on our math task, but manipulating concrete symbols may not be ideal for learning in a domain as symbolic and abstract as mathematics. The utility of teaching ideas through action on objects may depend on whether an answer can be "read off" of the manipulative (see Samara & Clements, 2009). Consider, for example, a mental rotation task in which doing the action *gives* the answer (i.e., once the object is rotated, one can see whether or not it is the same object as the comparison object). Using action may be more effective than using gesture to teach a task like mental rotation. Recent evidence from a study of mental transformation in kindergarten children provides some support for this hypothesis; training children to rotate the objects led to greater success immediately after instruction than did training children to gesture the rotation (although children who were trained to gesture did catch up at a 1-week follow-up; Levine, Goldin-Meadow, Carlson, & Hemani-Lopez, 2014). The age of the learner might also have an impact on the relative effectiveness of action and gesture. Not only do young children appear to internalize ideas through action experience (Vygotsky, 1978), but they also find gesture to be more difficult to interpret than action (Novack, Goldin-Meadow, & Woodward, 2014). Thus, although we found that gestures were better than action in promoting generalization, additional research is needed to determine the pervasiveness of this effect across domains and ages.

Why does gesture promote the kind of learning that leads to generalization? Our findings suggest that gesture may play a role in helping learners process the words it accompanies less superficially. In this study, we trained children to produce the equalizer strategy in speech, and children in all three conditions followed our instructions, producing the words on every problem they attempted to solve. However, repeating these words led to deep learning (i.e., generalization to near- and far-transfer problems) only when they were produced along with gesture, either concrete or abstract. Children in the action condition repeated the words of the equalizer strategy while continuing to answer problems incorrectly, which suggests that they did not really understand what they were saying. Our findings thus lead to the intriguing hypothesis that saying words while gesturing may help a learner integrate and internalize those words, whereas saying words while acting may not (see also Cook, Mitchell, & Goldin-Meadow, 2008).

In sum, our findings provide the first evidence that gesture not only supports learning a task at hand but, more important, leads to generalization beyond the task. Children appear to learn underlying principles from their actions only insofar as those actions can be interpreted symbolically. Although gesture can be thought of as simulated action (Hostetter & Alibali, 2008), the features of gesture that differentiate it from action may be precisely what makes it useful for generalization. On a continuum from action to abstraction, gesture is more abstract than action but still less abstract than verbal language. Perhaps it is being situated in this comfortable middle ground, with one foot in the concrete and one foot in the abstract, that makes gesture such a powerful tool for learning.

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.

Notes

1. As noted, *action* is a superordinate term that includes gesture; for ease of presentation, from this point on, we use the term *action* to refer only to actions that have a direct and physical effect on the world, that is, to refer to actions that are not gestures

2. We also coded children's speech and gesture in explaining how they solved the pretest problems and classified children as *mismatchers* if they produced at least three explanations in which the strategy conveyed in gesture was different from the strategy conveyed in speech in the same explanation (see Perry et al., 1988, who found that mismatchers are particularly ready to learn mathematical equivalence). We found no differences in the number of mismatchers across conditions, $F(1, 2) = 0.73$, $p = .485$, and no difference in the number of mismatching explanations across conditions, $F(1, 2) = 1.533$, $p = .22$. These results confirm the absence of preexisting differences among the conditions. In addition, including mismatcher status and number of mismatches as predictors in the models reported in the Results section did not alter the findings, and neither measure interacted with any of the reported measures.

3. We used two versions of the mathematical-equivalence test; if a child was given version A at the pretest, that child was given version B at the posttest, and vice versa.

4. The generalization test also included two problems that could be solved by generalizing grouping to multiplication (e.g., $2 \times 4 \times 3 = \_\_ \times 3$) and two problems that could be solved by grouping two nonadjacent addends (e.g., $7 + 2 + 3 = 2 + \_\_$). We did not use the scores on these problems to assess far transfer, as both of these problem types can be solved by using some version of grouping.

References

Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, MA: Harvard University Press.

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, *137*, 706–723. doi:10.1037/a0013157

Cook, S., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*, 1047–1058. doi:10.1016/j.cognition.2007.04.010

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

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20*, 267–272. doi:10.1111/j.1467-9280.2009.02297.x

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences*, *14*, 69–110. doi:10.1207/s15327809jls1401_4

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gesture as simulated action. *Psychonomic Bulletin & Review*, *15*, 495–514. doi:10.3758/PBR.15.3.495

James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science*, *13*, 279–288. doi:10.1111/j.1467-7687.2009.00883.x

Kontra, C. E., Goldin-Meadow, S., & Beilock, S. L. (2012). Embodied learning across the life span. *Topics in Cognitive Science*, *4*, 731–739. doi:10.1111/j.1756-8765.2012.01221.x

Lakoff, G., & Nunez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

Levine, L. C., Goldin-Meadow, S., Carlson, M., & Hemani-Lopez, N. (2014). *Training mental rotation skill in young children: Action trumps gesture in the shorter- but not the longer-term*. Manuscript submitted for publication.

McNeil, N. M., & Fyfe, E. R. (2012). "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction*, *22*, 440–448. doi:10.1016/j.learninstruc.2012.05.001

McNeil, N. M., & Uttal, D. H. (2009). Rethinking the use of concrete materials in learning: Perspectives from development and education. *Child Development Perspectives*, *3*, 137–139. doi:10.1111/j.1750-8606.2009.00093.x

McNeil, N. M., Uttal, D. H., Jarvin, L., & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. *Learning and Instruction*, *19*, 171–184. doi:10.1016/j.learninstruc.2008.03.005

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Mix, K. S. (2010). Spatial tools for mathematical thought. In K. S. Mix, L. B. Smith, & M. Gasser (Eds.), *Space and language* (pp. 41–66). New York, NY: Oxford University Press.

Novack, M. A., Goldin-Meadow, S., & Woodward, A. L. (2014). *Learning from gesture: How early does it happen?* Manuscript submitted for publication.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*, 359–400. doi:10.1016/0885-2014(88)90021-4

Piaget, J. (1953, November). How children form mathematical concepts. *Scientific American*, *189*, 74–79.

Sarama, J., & Clements, D. (2009). Concrete computer manipulatives in mathematics education. *Child Development Perspectives*, *3*, 145–150. doi:10.1111/j.1750-8606.2009.00095.x

Sommerville, J. A., & Woodward, A. L. (2010). The link between action production and action processing in infancy. In F. Grammont, D. Legrand, & P. Livet (Eds.), *Naturalizing intention in action* (pp. 67–90). Cambridge, MA: MIT Press.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology*, *18*, 37–54. doi:10.1016/S0193-3973(97)90013-7

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes* (pp. 79–81). Cambridge, MA: Harvard University Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*, 625–636.