Title
Hand Gesture and Mathematics Learning: Lessons From an Avatar

Permalink
https://escholarship.org/uc/item/9rk0m1x3

Journal
Cognitive Science, 41(2)

ISSN
0364-0213

Authors
Cook, Susan Wagner
Friedman, Howard S
Duggan, Katherine A
et al.

Publication Date
2017-03-01

DOI
10.1111/cogs.12344

Peer reviewed

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 41 (2017) 518–535

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12344

Hand Gesture and Mathematics Learning: Lessons From an Avatar

Susan Wagner Cook,[a] Howard S. Friedman,[b] Katherine A. Duggan,[b] Jian Cui,[c] Voicu Popescu[c]

[a]*Department of Psychological and Brain Sciences, Delta Center, University of Iowa*
[b]*Department of Psychology, University of California, Riverside*
[c]*Department of Computer Science, Purdue University*

Received 30 September 2014; received in revised form 7 August 2015; accepted 11 September 2015

Abstract

A beneficial effect of gesture on learning has been demonstrated in multiple domains, including mathematics, science, and foreign language vocabulary. However, because gesture is known to co-vary with other non-verbal behaviors, including eye gaze and prosody along with face, lip, and body movements, it is possible the beneficial effect of gesture is instead attributable to these other behaviors. We used a computer-generated animated pedagogical agent to control both verbal and non-verbal behavior. Children viewed lessons on mathematical equivalence in which an avatar either gestured or did not gesture, while eye gaze, head position, and lip movements remained identical across gesture conditions. Children who observed the gesturing avatar learned more, and they solved problems more quickly. Moreover, those children who learned were more likely to transfer and generalize their knowledge. These findings provide converging evidence that gesture facilitates math learning, and they reveal the potential for using technology to study non-verbal behavior in controlled experiments.

Keywords: Gesture; Learning; Mathematics; Animated pedagogical agent; Nonverbal behavior; Cognitive development; Instruction

1. Introduction

Teacher gesture speaks to learners more effectively than words alone. Including hand gesture in instruction increases learning of a wide variety of concepts, including Piagetian conservation (Church, Ayman-Nolley, & Mahootian, 2004; Ping & Goldin-Meadow,

---

Correspondence should be sent to Susan Wagner Cook, Department of Psychological and Brain Sciences, E11 Seashore Hall, University of Iowa, Iowa City, IA 52242. E-mail: susan-cook@uiowa.edu

2008), symmetry (Valenzeno, Alibali, & Klatzky, 2003), and mathematical equivalence (Singer & Goldin-Meadow, 2005). For example, one study of Piagetian conservation (Church et al., 2004) found that instruction with gesture led to nearly two times as many students demonstrating deep learning relative to instruction without gesture. Meta-analysis confirms that speakers' gestures benefit listeners' comprehension (Hostetter, 2011). However, understanding the causal link between gesture and learning has been complicated by the confounding of gesture with other communicative cues.

To establish causal links between gesture and learning outcomes, we need pure experimental manipulations of gesture. Many studies use controlled instructions delivered by an experimenter who gestures in one condition and does not gesture in a comparison condition (e.g., Cook & Goldin-Meadow, 2006; Rowe, Silverman, & Mullan, 2013; Singer & Goldin-Meadow, 2005). However, instruction delivered by a live experimenter introduces confounding factors. First, it is impossible for a live instructor to be blind to condition. This opens the door for experimenter bias, and observer-expectancy effects, whereby the experimenter may elicit precisely the expected behaviors of interest (Rosenthal, 1976). It is not clear that live actors can appropriately monitor all of their behavior, particularly when considerable attention is devoted to controlling speech and gesture. Controlled live instructions are also unnatural because they are practiced and therefore do not include the hesitations, false starts, and other characteristics of spontaneous communication. Moreover, changes in instruction across time, as experimenters gain experience, may interact with experimenter bias to further distort observed effects.

An alternative to live instructions is to use carefully constructed videos, matching behaviors that are not of interest so that the only feature known to vary across conditions is gesture (e.g., Cook, Duffy, & Fenn, 2013; Valenzeno et al., 2003). Video instruction is stable over time and eliminates experimenter bias in delivery. Researchers can use a single audio track to ensure that the speech is identical across conditions, and the videos can be carefully constructed for similarity in non-gestural non-verbal behavior. However, this approach is labor intensive and requires researchers to learn to produce highly practiced behavior or to lip-sync, which may not effectively preserve the audio-visual correspondence between speech and lip movements. Both live and video approaches may also bias researchers to use instructions that are simple and short so that other variables can be controlled, decreasing external validity since teaching often involves extended interactions. Finally, video presentation differs from live presentation along a number of dimensions, including the reduction of a three-dimensional movement into a (typically smaller) two-dimensional space and a lack of perceived interactivity and social presence. Although there is research to suggest that gestures presented on video can influence learning much like gesture presented "live" (Hostetter, 2011), other research suggests that children sometimes learn more from live presentation (Barr, 2010).

Thus, although much research to date has found that gesture can increase learning, fundamental methodological issues in studying gesture limit the conclusions that can be drawn. An alternative to live and video instructions is to use a computer-generated, embodied pedagogical agent, or avatar, to present instructions with and without gesture. Although this approach necessarily shares many drawbacks of video presentation, avatars have some unique benefits. Avatars allow for greater control of stimuli, including complete control over both the presence and timing of potentially confounding non-verbal behaviors, including face, lip, and body movements, even with long and complex stimuli. Avatars may be perceived as more social and interactive than video presentations, and they can be programmed to respond contingently to learner behavior. Furthermore, physical and other aspects of avatars can be readily manipulated in ways that videos (and live humans) cannot, allowing for testing of hypotheses that is otherwise not feasible. Of course, these benefits come with some additional costs, in that avatars may not be perceived and processed in the same way as more naturalistic stimuli (e.g., Black, Chang, Chang, & Narayanan, 2009). There is also a risk that avatars that are superficially too natural may even enter the “uncanny valley” (resulting in revulsion toward a not-quite human figure) and be perceived negatively by viewers rather than positively (Mori, MacDorman, & Kageki, 2012).

Avatars can support learning, both by directly communicating content to learners (Atkinson, 2002; Barlow, 1997; Baylor & PALS Research Group, 2003; Lester et al., 1997; Lusk & Atkinson, 2007; Moreno & Mayer, 2007) and by creating a positive learning environment (Heidig & Clarebout, 2011; Lester et al., 1997). In one example, Buisine and Martin (2007) used an avatar to compare learning from instruction including pointing gesture that were redundant with speech (with the referent fully specified in speech), pointing gestures that were complementary to speech (with the referent referred to only deictically in speech (e.g., “that”)) and speech without meaningful gesture. Buisine and Martin (2007) found that recall of the presented material was best when speech was fully informative and gesture was redundant with speech. Gesture that was redundant with speech also led to the highest subjective ratings of agent quality.

However, several factors limit the conclusions that can be drawn from prior work with instructional avatars who gesture. Most work has focused exclusively on pointing gesture (e.g., Baylor & Kim, 2008, 2009; Craig, Gholson, & Driscoll, 2002; Dunsworth & Atkinson, 2007; Mayer & DaPra, 2012; Moreno, Reislein, & Ozogul, 2010) without including the variety of gesture types typically seen during human communication. In addition, prior research has failed to control for agents’ eye gaze, body position, and other non-verbal behavior (e.g., Atkinson, 2002; Baylor & Kim, 2009; Craig et al., 2002; Dunsworth & Atkinson, 2007; Lusk & Atkinson, 2007; Mayer & DaPra, 2012; Moreno et al., 2010), and some prior research has confounded instructor speech and instructor gesture by changing the verbal instructions along with the gestures (Buisine & Martin, 2007)). When speech and other non-verbal behaviors are not comparable across conditions, the observed effects cannot be uniquely attributable to gesture. Finally, because gestures have typically been implemented without attention to the form and timing of spontaneously produced gestures, it is not clear whether previous conclusions also apply to naturalistic communication (Atkinson, 2002; Buisine & Martin, 2007; Craig et al., 2002; Dunsworth & Atkinson, 2007; Lusk & Atkinson, 2007; Mayer & DaPra, 2012; Moreno et al., 2010).

Thus, although prior research suggests that avatar gesture is beneficial for learning (Atkinson, 2002; Bergmann & Macedonia, 2013; Buisine & Martin, 2007; Dunsworth & Atkinson, 2007; Lusk & Atkinson, 2007; Moreno et al., 2010), because of concerns about both internal and external validity these findings provide only preliminary evidence for a general beneficial effect of gesture on learning. To provide strong evidence for a functional role of gesture on learning, it is necessary to manipulate gesture while completely controlling speech and all other non-verbal behaviors. In addition, using gestures that are based on gestures used spontaneously by human instructors will make findings more likely to generalize to spontaneous human interaction.

We used an instructional avatar to investigate how gesture influences learning of mathematical equivalence. Understanding of math equivalence predicts children's later math achievement even when controlling for baseline math achievement, IQ, and SES, suggesting that understanding of equivalence is a fundamental concept in mathematics (Devlin, McNeil, Carrazza, Byrd, & McKeever, 2015). Moreover, much is known about how instructors and students gesture during lessons on mathematical equivalence (Goldin-Meadow, Kim, & Singer, 1999; Singer & Goldin-Meadow, 2005). Thus, equivalence offers an excellent and important case study for examining the effects of gesture on learning using an instructional avatar.

We hypothesized that children would benefit from avatar gesture much like they benefit from human gesture, and that the effect of gesture would be seen even when other non-verbal behaviors were fully controlled. Specifically, we expected avatar gesture to facilitate understanding of trained material, as well as promoting near and far transfer (Cook et al., 2013). If we can find beneficial effects of gesture even when other behaviors are fully controlled, this would provide converging evidence to support the claim that gesture underlies the learning seen in other studies.

2. Method

2.1. Participants

Sixty-five children (31M, 34F, M_{age} 9;0 years, range 7;4–10;5) participated in this study. Children were recruited via a database created from birth records. Parents received a mailing and were subsequently contacted by phone to answer any questions and schedule sessions if interested. Children received a paperback book for participating. Data from one child with Asperger's syndrome was excluded from the analysis.

2.2. Materials

We developed an instructional avatar implemented via a computer animation character and used this platform to investigate the effect of avatar gesture on student learning. Twelve avatar stimulus videos were created for this study. Each video depicted an avatar standing in front of a virtual white board while explaining a mathematical equivalence problem (see Fig. 1; supplementary online material). There were six pairs of stimuli. Members of a pair had the same problem depicted on the whiteboard and used the same audio track. The audio track was recorded by a male

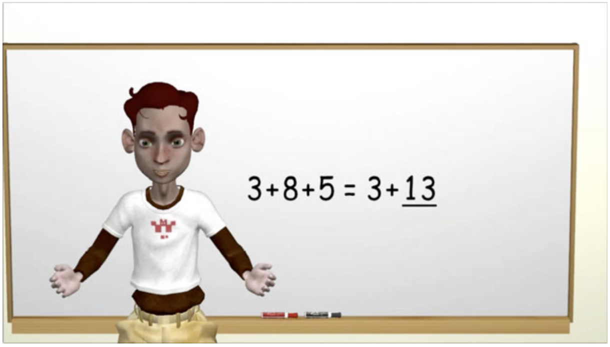

Fig. 1. Example of a still frame depicting a bimanual beat gesture.

speaker sitting in an isolated room in front of a computer and was collected without accompanying gestures. One video in each pair contained additional hand gestures, whereas the other did not contain any gesture, but otherwise non-verbal behavior was identical across the members of a pair and was highly similar across all six instructional videos.

In the videos with hand gestures, the avatar produced two types of gestures—content gestures designed to reinforce the conceptual content of the lesson, and bimanual beat gestures with outward focused movements (described below and depicted in Fig. 1) intended to help increase the charisma and appeal of the avatar (Duggan et al., n.d.). The content gestures were not generated automatically (Cassell et al., 1994) but rather were adapted from instructions used in prior research (Cook et al., 2013) and depicted the notion of equivalence between the two sides of the equation using two-handed balance gestures and sequences of points with the two hands separately indicating the two sides of the equation. These content gestures were developed based on extensive previous experience with children and adults explaining the concept of mathematical equivalence and were consistent with the gestures spontaneously used by children and adults. Consistent with the behavior of real teachers (Alibali & Nathan, 2012), we included a variety of gesture types, and the most frequently used gesture in our instructions was a pointing gesture. All content gestures were timed such that the preparation and onset of the stroke of the gesture slightly preceded the relevant speech, but the gestures were not specifically fine-tuned to match the content or prosody of the accompanying speech like spontaneous gesture (Chiu & Marsella, 2011). The beat gestures did not depict specific content, but rather the outward-focused movements were timed to important content words in the accompanying audio track. Adults often use beat gesture in mathematics instruction with children (Alibali & Nathan, 2012). The beat movements were based on the social psychological literature on gesture and charisma (DePaulo & Friedman, 1998). Other than the presence or absence of gestures, the stimulus videos were identical in the avatar's body position, facial expression, posture, eye gaze, and lip movements. Video stimuli are available as supporting online material (Supporting online material).

2.3. Procedure

All stimuli were presented on a computer using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). Children had unlimited time to solve each problem. An experimenter was available to assist as necessary, and parents observed from an adjoining room through an open window. Children were not typically given any assistance from the experimenter, who provided minimal responses to questions (e.g., responding "just put whatever you think is best" when children requested assistance or stated that they did not know how to solve the problems).

Prior to observing the instruction with the avatar, children solved six matching addends mathematical equivalence problems on the computer. Each of these problems had three single-digit numbers (2–9) on the left side of the equal sign and one number and an answer blank on the right side. The answer blank occurred in either the left or the right position on the right-hand side. The addend on the right side was always the same as one addend from the left side, and it was located in a matching position across the equal sign (e.g., $5 + 9 + 7 = 5 + \_$ or $7 + 8 + 4 = \_ + 4$). Each problem was displayed on the computer monitor, and children typed their answers using the keyboard.

After this pretest, children participated in a video tutorial. Children first viewed an introductory statement from the avatar about the principle of mathematical equivalence. They then viewed six explanations of novel, equal addends mathematical equivalence problems presented by the avatar in a fixed order. The correct answer was present during all explanations given by the avatar, and the avatar provided explanations of correct solutions to the problem.

Children were randomly assigned to an experimental condition. For children in the gesture condition, the introduction and the six videos contained accompanying hand gesture, and for children in the no gesture condition, these videos did not contain any hand gestures. After each avatar explanation, children solved an equal addends mathematical equivalence problem of their own at the computer, with the answer blank located in the same position as the problem immediately previously described by the avatar. Children were not given any feedback on their solutions.

After the lesson, children completed a posttest at the computer. Posttest problems were presented in a fixed order. The first six problems on the posttest were novel equal addends equivalence problems with matching addends, so they were identical in form to the pretest and to the problems explained by the avatar. Next there were four transfer problems, two with equal addends not located in a matching position and two with no equal addends. Finally, there were six conceptual questions, which were true/false questions about equality, adapted from prior research (Matthews, Rittle-Johnson, McEldoon, & Taylor, 2012). Children's responses and reaction times were recorded for analysis.

3. Results

A multilevel logistic regression model was used to account for variability across individual subjects as well as variability in problem difficulty. The log of the odds of correctly solving each individual posttest problem was predicted from an interaction of instructional condition and problem type (trained problems; transfer problems; conceptual problems) with a covariate for performance on the pretest. A random problem intercept was included to account for variation in difficulty across problems and problem types. Random subject intercepts and by subject random effects for type of problem, condition, and their interaction were also included (the full random effects structure as recommended by Barr, Levy, Scheepers, & Tily [2013]).

As in prior research on mathematical equivalence (Cook et al., 2013; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014), many children were successful on the pretest, with 27 children solving more than half of the pretest problems correctly. These children do not have room for improvement and may not be likely to benefit from gesture. We eliminated children scoring greater then 50% on the pretest from subsequent analysis. When considering only those children who were not successful on more than half of pretest problems ($n = 38$),[1] there was an effect of test with the posttest and transfer tests more difficult than the true/false conceptual test ($\beta_{\text{Posttest}} = -2.33$, $z = 2.14$, $p = .0325$; $\beta_{\text{Transfer}} = -2.40$, $z = 1.93$, $p = .053$). The pretest covariate was positive, but not significant ($\beta_{\text{Pre}} = 2.57$, $z = 0.73$). There was a significant effect of condition, with children in the gesture condition performing better than children in the no gesture condition ($\beta_{\text{Gesture}} = 2.75$, $z = 2.92$, $p < .01$, see Fig. 1). There were no significant interactions between condition and test (all ps $> .39$). Averaging across all tests, children in the gesture condition were correct on 86% of problems after training, while children in the no gesture condition were correct on only 73% of problems (Fig. 2).

One possible explanation for this pattern of performance is that children in the no gesture condition did not understand the instructions given during training, because of the lack of informative gesture. Because of the equal addends present in the problem, when the avatar names this number without an accompanying pointing gesture it is potentially ambiguous. However, an examination of performance during the instructional period suggests that children in both conditions were equally successful at understanding the provided instructions. Children in both conditions were equally successful at solving problems during the tutorial (GC 70% vs. NGC 68%). Thus, although children in both groups were equally successful during the lesson, children in the gesture condition were more successful on the posttest of learning given after instruction.

This finding suggests that children in the gesture condition may have acquired more transferable knowledge than those who learned without gesture. Accordingly, we assessed the generality of children's knowledge as an additional perspective on differences in learning across conditions. To do so, performance on the near transfer and conceptual problems was examined only for children who were successful on the trained problems on the posttest (correct on at least 66% of posttest matching addends equivalence prob-

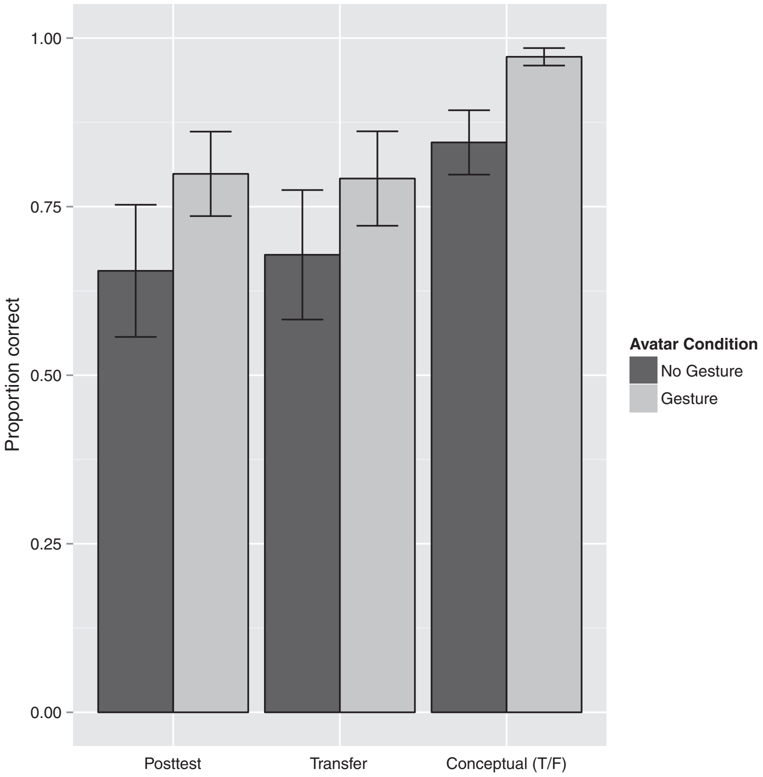

Fig. 2. Proportion of problems solved correctly on each test by children in each instructional condition.

lems, n = 29). The log of the odds of correctly solving each problem was again predicted from instructional condition as the factor of interest interacting with the test (transfer problems; conceptual problems). A random problem intercept was included to account for variation in difficulty across problems of each type. Random subject intercepts and by subject random effects for type of problem, condition, and their interaction were also included. As can be seen in Fig. 3, there was an effect of gesture condition ($\beta_{\text{Gesture}} = 3.86$, $z = 2.59$, $p < .01$), with children in the gesture condition doing better than children in the no gesture condition. The interaction between condition and test was not significant. Thus, even when we only consider those children who were successful on the trained problems after instruction, children who learned with gesture were more successful at transfer and conceptual problems compared with children who learned without gesture. Moreover, the effect of gesture condition on transfer and generalization in this group was reliable even when number of problems correct on the posttest is included as an additional covariate, suggesting that it is how the knowledge was acquired, in addition to how much knowledge was acquired, that is important in supporting performance on the transfer and conceptual problems.

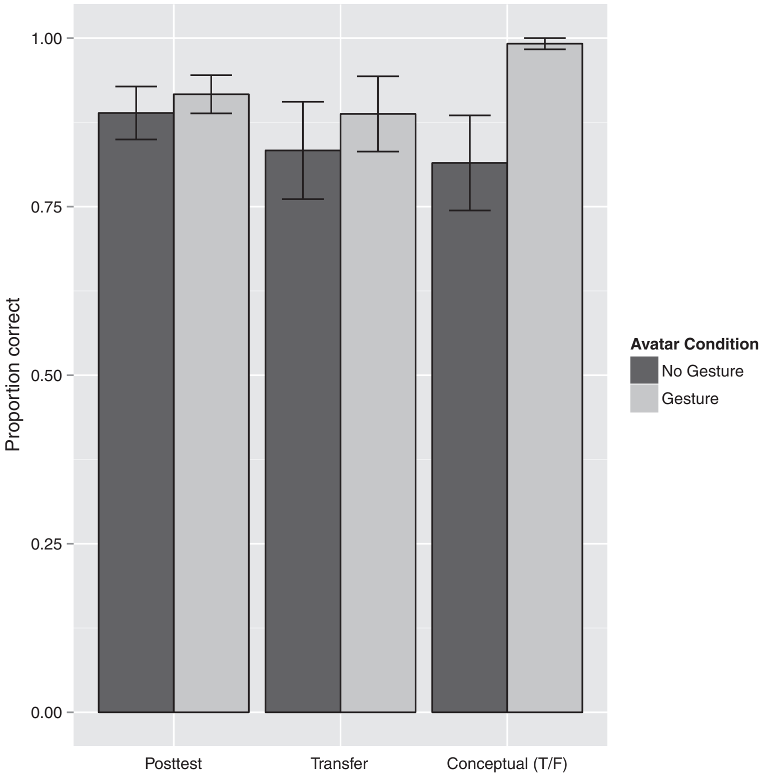

Fig. 3. Proportion of problems solved correctly on each test by children in each instructional condition who were successful on posttest.

In addition to assessing whether or not children solved problems correctly, we also measured the amount of time children spontaneously took to solve the problems. Because solution time is quite different for correct and incorrect solutions given the different mathematical operations required, our analysis of solution time included only those children who were not overwhelmingly correct prior to instruction, and who primarily solved problems correctly after instruction, and only correct solutions. We also excluded problems with time to solve greater than the 95th percentile for that type of problem (trained problems; transfer problems; conceptual problems). Our criteria for success after instruction were defined as scoring greater than or equal to 50% correct on the trained problems after instruction. The log of the amount of time to solve the problem was predicted from the interaction of instructional condition and the type of problem. A random problem intercept was included to account for variation in difficulty across problems of each type. Random subject intercepts, by subject random effects for type of problem, condition, and their interaction, were also included. For these analyses, we report the *t* value as the statistic of interest, as there is not a consensus on the degrees of freedom needed to calculate an exact *p* value. We considered *t* values ≥ 2 as significant findings. As can be seen

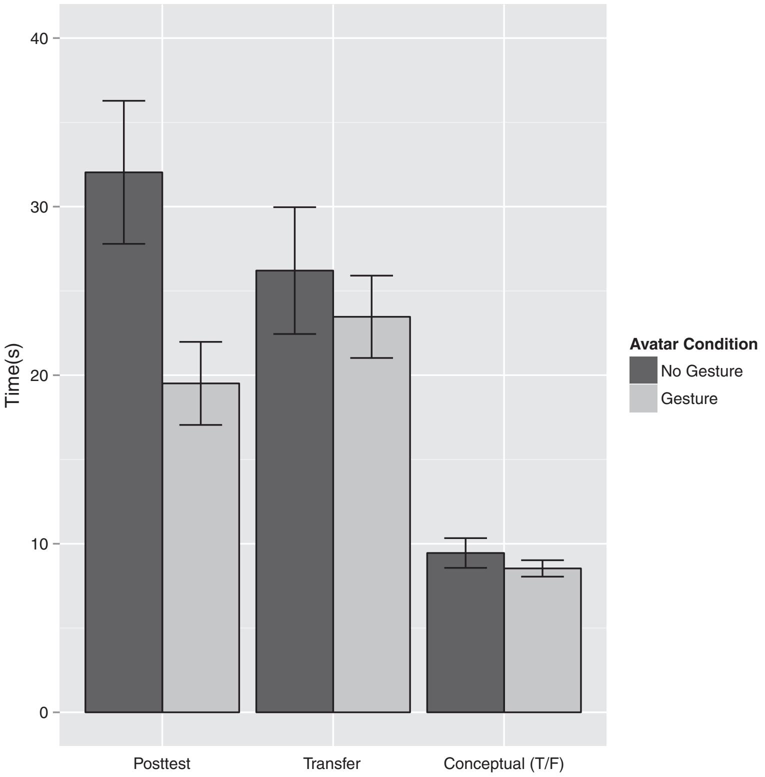

Fig. 4. Time to solve correct equivalence problems for children who learned.

in Fig. 4, there was an interaction of condition and test, with children in the gesture condition solving the posttest trained problems more quickly than children in the no gesture condition ($\beta_{\text{Gesture}} = -0.39$, $t = 2.5$). The posttest trained problems and the transfer problems also took more time than the true/false generalization problems ($\beta_{\text{Posttest}} = 0.78$, $t = 4.0$, $\beta_{\text{Transfer}} = 0.97$, $t = 4.5$; Fig. 4).

4. Discussion

Children who view a mathematics lesson from a computer-generated avatar who gestures learn more than children who learn from an avatar who produces the identical eye gaze and lip and body movements without the accompanying gestures. Children who learn from the avatar who gestures also are more likely to transfer their knowledge compared with children who see the lesson without gesture, even when compared with children who do learn from the instructions without gesture. Moreover, children who learn after instruction with gesture also solve trained problems more quickly. These findings extend prior research on teacher gesture to a highly controlled environment. In conjunction with prior research comparing instruction with and without gestures using live (e.g., Singer & Goldin-Meadow, 2005) and videotaped (Cook et al., 2013) presentation, there is now considerable evidence to suggest that gestures themselves can facilitate learning. There is also new evidence that gesture has specific effects on transfer and solution time.

The influence of gesture on learning that was observed with the avatar was consistent in effect size with those observed in prior studies of gesture and learning (see Hostetter, 2011, for a meta-analysis of the effects of gesture on communication). The gestures used in this study were delivered with a computer animation avatar, but they facilitated learning similarly to gestures produced by human instructors. Previous research suggests that children interact and converse with cartoon avatars much like they interact with human partners, even when they prefer the human partner (Hyde, Kiesler, Hodgins, & Carter, 2014).

We included transfer and conceptual problems designed to test children's conceptual understanding of equality. The effect of gesture was not limited to problems exactly like those seen during instruction, but rather extended to these additional tests of children's knowledge. This is consistent with prior work suggesting that gesture offers a benefit to transfer (Cook et al., 2013) and may promote conceptual knowledge. For example, one study (Cutica & Bucciarelli, 2008) demonstrated that learners were more likely to make deep inferences after instruction that included accompanying gesture, although in this study prosody and intonation were not controlled.

Although there is now considerable evidence that gestures can improve learning (Church et al., 2004; Richland & McDonough, 2010; Valenzeno et al., 2003), there is not a clear understanding of the mechanism by which gesture may enhance learning. Although our findings do not speak directly to mechanism, they do reveal that gesture helps create knowledge that can be deployed quickly, and that transfers and generalizes to conceptual understanding. Explanations of how gesture supports learning will need to account for changes in both the amount that is learned as well as changes in the nature of what is learned.

The instructions used in this study included a variety of gestures, and it is likely that different gestures work via different mechanisms and serve different functions for learners. One possibility is that some gestures help learners better understand the spoken instructions. The representational balance gestures provide a visual representation of equality that may facilitate understanding of the concept (Arcavi, 2003). The sweeping gestures clearly indicate the sides of the problem, which many children fail to encode correctly (McNeil & Alibali, 2004). The pointing gestures provide a visual cue for the referent of speech, likely facilitating identification and processing of the appropriate referents (Louwerse & Bangerter, 2010; Ozcelik, Arslan-Ari, & Cagiltay, 2010).

We think it is unlikely, however, that the only benefit to gesture in our paradigm is due to facilitating, disambiguating, or improving understanding of speech. Children in both conditions performed equally well during the instructional period, suggesting that children in the no gesture condition did not have great difficulty understanding the spoken instructions. There were also effects of gesture on learning even when considering only children who learned from the instruction. These children clearly had effectively unpacked the meaning of the instructor's speech. Thus, it seems that gesture contributes something to children's learning that is beyond simply promoting understanding in the moment.

Another possibility is that some gestures may make the spoken instructions easier to process, reducing demand on cognitive systems and freeing up resources that then support learning (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Wagner, Nusbaum, & Goldin-Meadow, 2004). This possibility is consistent with the finding that children who learn with gesture are subsequently faster to solve problems after training compared with children who learn, but without gesture. However, this account leaves open the question of exactly how gestures ease processing.

Deictic gestures that directly indicate or that are located near their referents may also help learners by providing a direct and visible link between the verbal instructions and the immediate environment (Ballard, Hayhoe, Pook, & Rao, 1997). Children learning mathematics need to link their mental representations to problem representations in the external environment. Gesture and other deictic representations may provide a powerful cue for bridging internal and external representations in support of learning and transfer.

One possibility is that what is unique about gesture is that all of its properties come together in a single representation: It is a deictic, visual-motor embodied representation that is uniquely synched with speech. Alternatively, gesture may simply function as a highly effective visual cue for learners (e.g., de Koning, Tabbers, Rikers, & Paas, 2010; Jamet, Gavota, & Quaireau, 2008; Kalyuga, Chandler, & Sweller, 1999; Koning, Tabbers, Rikers, & Paas, 2009; Lin & Atkinson, 2011). If so, then we should expect other visual cues to support learning like gestures. However, findings on the beneficial effects of visual signaling in multimodal learning have been mixed, with many studies failing to find benefits of visual cues (e.g., Bartholomé & Bromme, 2009; Berthold & Renkl, 2009; Huk, Steinke, & Floto, 2009; Tabbers, Martens, & Merriënboer, 2010). Moreover, direct comparison of hand gesture with other methods of visual signaling suggests that gesture may be unique in its effects on learners (e.g., Moreno et al., 2010). For example, De Koning and Tabbers (2013) compared visual signaling via an arrow and via a picture of a pointing hand and found learning benefits associated with viewing the pointing hand. Similarly, Johnson and colleagues (De Koning & Tabbers, 2013) found that students with low prior knowledge benefitted more from agent gesture than from an arrow.

Gesture may also influence learning indirectly by making instruction more appealing and engaging, increasing learner interest, attention, and motivation and making learning seem less difficult (Baylor & Ryu, 2003; van Mulken, André, & Müller, 1998). Gesturing humans and avatars may be perceived as more natural and more social than non-gesturing human avatars, and they may thereby engage learners in interaction likely to promote deep processing by priming a social interaction schema (Mayer, 2005; Mayer & DaPra, 2012). Indeed, prior research has found that gesture increases viewers' perception of the language competence and likeability of animated agents (Cassell & Thorisson, 1999), and other research has similarly reported that gesture can facilitate perception of the quality of the agent's explanation (Buisine & Martin, 2007). Moreover, viewers respond most positively to agents whose gestures should read gestures are based on those of an individual person, rather than modeled from data from multiple individuals (Bergman, Kopp & Eyssel, 2010).

The computer-generated instructional avatar offers a new tool for uncovering exactly how gesture supports learning because it allows for testing hypotheses with a high degree of experimental control. The avatar allows one to generate experimental stimuli that do not vary along dimensions that are not of interest. As demonstrated here, using the avatar, one can present stimuli with accompanying eye gaze, lip, and body movements that do not vary across gesture conditions. In addition, computer-generated avatars can be programmed to respond contingently to various behaviors of the learner, including learner gesture or eye gaze. Thus, the avatar will likely provide a useful tool for testing mechanistic accounts of learning with gesture, at both an individual and group level (following [Baylor & Rhu, 2003]).

Gesture has been shown to be helpful in supporting learning for a wide variety of content, and in a variety of paradigms, including learning via analogy (Richland & McDonough, 2010), foreign language vocabulary learning (Macedonia, Müller, & Friederici, 2010), and science learning (Singer, Radinsky, & Goldman, 2008), as well as learning from discourse (Cutica & Bucciarelli, 2008). Gestures have also been shown to be especially helpful for atypical learners. For example, children with ADHD (Wang, Bernas, & Eberhard, 2004) and children with language impairment (Kirk, Pine, & Ryder, 2011; Weismer & Hesketh, 1993) seem to benefit from gesture even more than typically developing children. The avatar may offer a platform to bring gesture to a variety of diverse populations. Using an instructor avatar implemented as a computer animation character, we can go from exploring when and why gesture benefits learning to personalizing gesture to facilitate learning and communication in general.

The research reported here addresses an important confound in prior work investigating the function of gesture with respect to learning of mathematical equivalence. Differences in eye gaze, face, lip and body movements, and other non-verbal behavior cannot explain the beneficial effect of gesture on learning or transfer of learning observed here, because these behaviors were identical across experimental conditions. Thus, we can conclude one way that children can learn math is by hand.

Acknowledgments

The authors thank Nicoletta Adamo-Villani for her help with the animation of the avatar, and Martha W. Alibali for helpful comments on a prior version of the manuscript. We thank Andrew Allyn for help with stimulus preparation, subject recruitment, and data collection. This research was supported by NSF Award No. 1216984 to Howard S. Friedman, NSF Award No. 1217137 to Susan Wagner Cook, and NSF Award No. 1217215 to Voicu Popescu from the National Science Foundation. Access to the animation system will be provided via other papers and presentations being developed from this collaborative project.

Note

1. The inclusion criteria did not influence this pattern of findings, and so we opted for a fairly liberal inclusion criteria.

References

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, *21*(2), 247–286. doi:10.1080/10508406.2011.611446

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, *52*(3), 215–241.

Atkinson, R. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, *94*(2), 416–427. doi:10.1037//0022-0663.94.2.416

Ballard, D. H., Hayhoe, M., Pook, P., & Rao, R. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(4), 723–767.

Barlow, S., Converse, S., Kahler, S., Lester, J., & Stone, B. (1997). Animated pedagogical agents and problem solving effectiveness: A large scale empirical evaluation. In B. du Buolay, & R. Mizoguchi (Eds.), *Artificial intelligence in education* (p. 23). Amsterdam: IOS Press.

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review*, *30*(2), 128–154 http://doi.org/10.1016/j.dr.2010.03.001.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), doi:10.1016/j.jml.2012.11.001

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology*, *101*(2), 282–293. http://doi.org/10.1037/a0014312.

Baylor, A. L., & PALS Research Group. (2003). *The impact of three pedagogical agent roles* (pp. 928–929). Presented at the second international joint conference. New York: ACM Press. doi:10.1145/860575.860729

Baylor, A. L., & Kim, S. (2008). The effects of agent nonverbal communication on procedural and attitudinal learning outcomes. *Intelligent Virtual Agents*, doi:10.1007/978-3-540-85483-8_21

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, *25*(2), 450–457. doi:10.1016/j.chb.2008.10.008

Baylor, A. L., & Ryu, J. (2003). The effects of image and animation in enhancing pedagogical agent persona. *Journal of Educational Computing Research*, *28*(4), 373–394. doi:10.2190/V0WQ-NWGN-JB54-FAT4

Bergmann, K., Kopp, S., & Eyssel, F. (2010). Individualized gesturing outperforms average gesturing — Evaluating gesture production in virtual humans. In Albeck et al. (Eds.), *Proceedings of the 10th Conference on Intelligent Virtual Agents* (Vol. 6356, pp. 104–117). Berlin: Springer. doi:10.1007/978-3-642-15892-6_11

Bergmann, K., & Macedonia, M. (2013). A virtual agent as vocabulary trainer: Iconic gestures help to improve learners' memory performance. In R. Aylett et al. (Eds.), *Proceedings of the 10th Conference on Intelligent Virtual Agents* (Vol. 8108, pp. 139–148). Berlin: Springer. doi:10.1007/978-3-642-40415-3_12

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70. doi:10.1037/a0013247

Black, M., Chang, J., Chang, J., & Narayanan, S. (2009). Comparison of child-human and child-computer interactions based on manual annotations. In *Proceedings of the Workshop on Child*, Computer and Interaction. Cambridge, USA.

Buisine, S., & Martin, J.-C. (2007). The effects of speech–gesture cooperation in animated agents' behavior in multimedia presentations. *Interacting with Computers*, *19*, 1–10. doi:10.1016/j.intcom.2007.04.002

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., & Stone, M., et al. (1994). *ANIMATED CONVERSATION: Rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents* (pp. 413–420). Presented at the the 21st annual conference, New York: ACM Press. doi:10.1145/192161.192272

Cassell, J., & Thorisson, K. R. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, *13*(4–5), 519–538. doi:10.1080/088395199117360

Chiu, C. C., & Marsella, S. (2011). How to train your avatar: A data driven approach to gesture generation. In *Intelligent Virtual Agents*, (pp. 127–140) Berlin: Springer.

Church, R., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, *7*(4), 303–319.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, *84*(6), 1863–1871. doi:10.1111/cdev.12097

Cook, S., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds. *Journal of Cognition and Development*, *7*(2), 211–232. DOI: 10.1207/s15327647jcd0702_4.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*(2), 428. doi:10.1037/0022-0663.94.2.428

Cutica, I., & Bucciarelli, M. (2008). The deep versus the shallow: Effects of co-speech gestures in learning from discourse. *Cognitive Science*, *32*(5), 921–935. doi:10.1080/03640210802222039

De Koning, B. B., & Tabbers, H. K. (2013). Gestures in instructional animations: A helping hand to understanding non-human movements? *Applied Cognitive Psychology*, *27*(5), 683–689. doi:10.1002/acp.2937

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. II, pp. 3–40). Boston: McGraw Hill.

Devlin, B., McNeil, N., Carrazza, C., Byrd, C., & McKeever, M. O. (2015). Early understanding of math equivalence predicts future math achievement. Presented at the Society for Research in Child Development, Philidelphia, PA.

Dunsworth, Q., & Atkinson, R. K. (2007). Fostering multimedia learning of science: Exploring the role of an animated agent's image. *Computers & Education*, *49*(3), 677–690. doi:10.1016/j.compedu.2005.11.010

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher"s hands tell the student"s mind about math. *Journal of Educational Psychology*, *91*(4), 720–730.

Goldin-Meadow, S., Nusbaum, H., Kelly, S., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*, 516–522.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, *6*(1), 27–54. doi:10.1016/j.edurev.2010.07.004

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis *Psychological Bulletin*, *137*(2), 297–315. doi:10.1037/a0022128

Huk, T., Steinke, M., & Floto, C. (2009). The educational value of visual cues and 3D-representational format in a computer animation under restricted and realistic conditions. *Instructional Science*, *38*(5), 455–469. doi:10.1007/s11251-009-9116-7

Hyde, J., Kiesler, S., Hodgins, J. K., & Carter, E. J. (2014). *Conversing with children* (pp. 1787–1796). Presented at the the 32nd annual ACM conference. New York: ACM Press. doi:10.1145/2556288.2557280

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction*, *18*(2), 135–145.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*(4), 351–371.

Kirk, E., Pine, K. J., & Ryder, N. (2011). I hear what you say but I see what you mean: The role of gestures in children's pragmatic comprehension. *Language and Cognitive Processes*, *26*(2), 149–170. doi:10.1080/01690961003752348

Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, *21* (2), 113–140. doi:10.1007/s10648-009-9098-7

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, *20*(2), 111–122.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). *The persona effect: Affective impact of animated pedagogical agents* (pp. 359–366). Presented at the the SIGCHI conference. New York: ACM Press. doi:10.1145/258549.258797

Lin, L., & Atkinson, R. K. (2011). Using animations and visual cueing to support learning of scientific concepts and processes. *Computers & Education*, *53*(3), 650–658.

Louwerse, M. M., & Bangerter, A. (2010). Effects of ambiguous gestures and language on the time course of reference resolution. *Cognitive Science*, *34*(8), 1517–1529. doi:10.1111/j.1551-6709.2010.01135.x

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology*, *21*(6), 747–764. doi:10.1002/acp.1347

Macedonia, M., Müller, K., & Friederici, A. D. (2010). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, *32*(6), 982–998. doi:10.1002/hbm.21084

Matthews, P., Rittle-Johnson, B., McEldoon, K., & Taylor, R. (2012). Measure for measure: What combining diverse measures reveals about children's understanding of the equal sign as an indicator of mathematical equality. *Journal for Research in Mathematics Education*, *43*(3), 316–350. doi:10.5951/jresematheduc.43.3.0316

Mayer, R. E. (2005). Principles of multimedia learning based on social cues: Personalization, voice, and image principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 201–212). New York: Cambridge University Press.

Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology-Applied*, *18*(3), 239–252. doi:10.1037/a0028616

McNeil, N., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science*, *28*(3), 451–466.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*(3), 309–326.

Moreno, R., Reislein, M., & Ozogul, G. (2010). Using virtual peers to guide visual attention during learning. *Journal of Media Psychology: Theories, Methods, and Applications*, 22(2), 52–60. doi:10.1027/1864-1105/a000008

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *Robotics & Automation*, *19*(2), 98–100. doi:10.2307/25704807?ref=no-x-route:e46664d78673b8826934eedef4fc9d67

van Mulken, S., André, E., & Müller, J. (1998). *The persona effect: How substantial is it?* (pp. 53–66). London: Springer. doi:10.1007/978-1-4471-3605-7_4

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science: A Journal of the American Psychological Society/APS*, *25*(4), 903–910. doi:10.1177/0956797613518351

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior*, *26*(1), 110–117. doi:10.1016/j.chb.2009.09.001

Ping, R. M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, *44*(5), 1277–1287.

Richland, L. E., & McDonough, I. M. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology*, *35*(1), 28–43. doi:10.1016/j.cedpsych.2009.09.001

Rosenthal, R. (1976). *Experimenter effects in behavioral research*. New York: Halsted Press.

Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, *38*(2), 109–117. doi:10.1016/j.cedpsych.2012.12.001

Singer, M., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, *16*(2), 85–89.

Singer, M., Radinsky, J., & Goldman, S. R. (2008). The role of gesture in meaning construction. *Discourse Processes*, *45*(4–5), 365–386. doi:10.1080/01638530802145601

Tabbers, H. K., Martens, R. L., & Merriënboer, J. J. G. (2010). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, *74*(1), 71–81. doi:10.1348/000709904322848824

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers" gestures facilitate students" learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*(2), 187–204. doi:10.1016/S0361-476X(02)00007-3

Wagner, S., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language*, *50*(4), 395–407.

Wang, X., Bernas, R., & Eberhard, P. (2004). Engaging ADHD students in tasks with hand gestures: A pedagogical possibility for teachers. *Educational Studies*, *30*(3), 217–229. doi:10.1080/0305569042000224189

Weismer, S. E., & Hesketh, L. J. (1993). The influence of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech Language and Hearing Research*, *36*(5), 1013–1025. doi:10.1044/jshr.3605.1013

Supporting Information

Additional Supporting Information may be found in the Supporting information tab for this article:

Video S1. Examples of the avatar video instruction for the gesture (GE1.mp4)

Video S2. Examples of the avatar video instruction for no gesture conditions (NG1.mp4)

Appendix

Speech and gesture transcript of the instructional lesson

Speech	Gesture
Intro (only before 1st problem)	
Hi everyone.	
Today we are going to [learn about the equal sign].	2H outward-focused beat gesture
[The equal sign] is a symbol that tells us about the two things on either side of it.	2H outward-focused beat gesture
Whatever is on [one side] of the equal sign needs to be the same amount	LH outward beat
as whatever is on [the other side] of the equal sign.	RH outward beat
Let's see how this works.	
Problem script (repeated for six problems—example problem 3 + 8 + 5 = _ + 5):	
Remember,	
[the equal sign] means that	RH Point to equal sign
the total amount [on the left side]	LH balance gesture
must be the same [as the right side]	RH balance gesture
This will help us [figure out] what goes	RH Simple Beat
inside the [blank].	RH Point to blank
[One side] needs to equal	RH Sweep right side
[the other side].	LH Sweep left side
You know you have [the right answer]	2H outward-focused beat gesture
when the two sides are the same amount.	RH Simple Beat
[Lets] figure out	
how to do this.	
[Three]	LH pt to 3 on left side
plus [eight]	LH pt to 8
equals eleven,	
[Eleven]	LH Simple Beat
plus [five]	LH pt to 5
equals sixteen,	
and [what number]	RH pt to blank
plus [three]	RH pt to 3 on right side
equals [sixteen]?	RH Simple Beat
[Three]	RH Pt to 3 on right side
plus [thirteen]	RH Pt to 13
[equals] sixteen. [Beat]	RH Simple Beat
If you look at [both sides]	RH Simple Beat
[they equal the same amount]	2H Balance gesture
which is [sixteen]	LH balance gesture
and [sixteen].	RH balance gesture
So one side equals the other side.	
OK, [now it is your turn].	2H outward-focused beat gesture