Consolidation and Transfer of Learning After Observing Hand Gesture

Susan Wagner Cook
University of Iowa

Ryan G. Duffy and Kimberly M. Fenn
Michigan State University

Children who observe gesture while learning mathematics perform better than children who do not, when tested immediately after training. How does observing gesture influence learning over time? Children (n = 184, ages 7–10) were instructed with a videotaped lesson on mathematical equivalence and tested immediately after training and 24 hr later. The lesson either included speech and gesture or only speech. Children who saw gesture performed better overall and performance improved after 24 hr. Children who only heard speech did not improve after the delay. The gesture group also showed stronger transfer to different problem types. These findings suggest that gesture enhances learning of abstract concepts and affects how learning is consolidated over time.

The spontaneous use of hand gesture is a pervasive feature of human communication. Even young children produce hand gestures when telling stories (Colletta, Pellenq, & Guidetti, 2010), solving problems (Church, 1999; Garber, Alibali, & Goldin-Meadow, 1998), and engaging in conversation (Boyatzis & Satyaprasad, 1994). Gestures are not merely epiphenomenal to speech but can aid in the learning of new material. Children learn more when they gesture while learning mathematical tasks (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Cook & Goldin-Meadow, 2006; Cook, Mitchell, & Goldin-Meadow, 2008; Feyereisen, 2006). Importantly, children can learn mathematics from producing gesture even when the gesture is not spontaneous and the child gestures only because he or she is instructed to do so (Broaders et al., 2007; Goldin-Meadow, Cook, & Mitchell, 2009). Thus, there is evidence for a causal role of producing gesture in changing children’s mathematical knowledge.

It is perhaps not surprising that producing gesture improves learning, given that producing actions can improve declarative memory. Memory for action phrases is better when the actions are performed than when the same words are simply studied (cf. Cohen, 1981; Cohen & Stewart, 1982; Saltz & Donnenwerth-Nolan, 1981). However, gesture production does more than simply improve memory; when individuals gesture during learning, their gestures may help them actively uncover knowledge (Broaders et al., 2007). Students are more likely to endorse a valid problem-solving strategy if they have previously demonstrated the strategy with their hands (Garber et al., 1998), and adults are more likely to transfer knowledge if they have previously gestured in a manner consistent with transfer (Beilock & Goldin-Meadow, 2010).

Gesture production may also directly affect the conceptual knowledge that children construct. This can be seen in two ways: first, third, and fourth-grade students who mimic their instructor’s hand gestures when solving mathematics problems are more likely to obtain correct solutions than children who participate in the same instruction and do not mimic, suggesting that the production of gesture can contribute to improved understanding (Cook & Goldin-Meadow, 2006). Second, third, and fourth-grade students who are explicitly told to mimic their instructor’s gestures are often able to verbalize the mathematical reasoning behind the gestures, even when that reasoning is never verbalized by the instructor (Goldin-Meadow et al., 2009). Thus, children seem to extract conceptual knowledge from gestures that they produce.

Merely observing gesture can also aid in learning. Valenzeno, Alibali, and Klatsky (2003) found that preschool children who saw an instructor produce meaningful gestures while learning about symmetry...
learned more than children who did not see the instructor gesture at all. Similarly, first-grade children who saw gesture during a videotaped lesson on conservation learned more than children who did not see gesture (Church, Ayman-Nolley, & Mahootian, 2004). Related studies have shown that observing gesture facilitates learning in a variety of different paradigms, including the learning of mathematical equivalence (e.g., \(2 + 4 + 6 = 2 + \_\_\); Perry, Berch, & Singleton, 1995; Richland & McDonough, 2010; Singer & Goldin-Meadow, 2005).

We have also shown that a combination of gesture observation and gesture production can contribute to the stability of learning over time in third- and fourth-grade students (cf. Cook, Yip, & Goldin-Meadow, 2010; Cook et al., 2008), but it is not known whether mere observation of gesture has a similar effect. In one study, children who both saw and produced gesture while learning mathematical equivalence showed maintenance of learning at a follow-up test given 4 weeks later. For these children, performance at the follow-up test was similar to performance on an immediate posttest. In contrast, when participants did not gesture during training, there was a decline in performance at the follow-up test. This is particularly interesting because the group that did not gesture performed as well as the group that gestured on the immediate posttest (Cook et al., 2008). Therefore, producing gesture helps maintain learning that otherwise would have been lost over time, suggesting that gesture does something more than simply contribute to understanding in the moment.

Our primary goal was to examine how gesture observation affects initial learning and subsequent consolidation of learning. Much of the previous research on gesture observation has tested recall in the same session in which material was learned (e.g., Church, Garber, & Rogalski, 2007; Ping & Goldin-Meadow, 2008; Riseborough, 1981; Straube, Green, Weise, Chatterjee, & Kircher, 2009; Woodall & Folger, 1985). A few related studies have used a much longer time interval, testing performance in mathematical equivalence several weeks after learning (e.g., Cook et al., 2008; Cook et al., 2010) but these studies all investigate gesture production, not gesture observation. No study has investigated the effect of observing gesture on learning over time. Furthermore, the effect of producing or observing gesture on learning over a shorter, 24-hr time period remains unresolved. We chose to test learning over a 24-hr period because prior work suggests that memory consolidation occurs in the first 24 hr after learning (see Diekelmann & Born, 2010; Margoliash & Fenn, 2008, McGaugh, 2000; for reviews). Therefore, if gesture influences consolidation processes, we should see an effect after 24 hr.

In addition, although numerous studies have shown that gesture can improve mathematical learning, much of this work has involved one-on-one interaction between an instructor and student (cf. Cook & Goldin-Meadow, 2006; Goldin-Meadow et al., 2009). This design allows for conditions to be balanced across classrooms and teachers but does not address the question of whether gesture also affects mathematical learning in a classroom setting. Previous work has shown that students learn more when they receive individualized, rather than group-level, instruction (Bloom, 1984). However, gesture might help children overcome the difficulties of group-level instruction and show even greater effects in a less controlled classroom environment, where students may experience greater distraction and difficulty maintaining attention. Indeed, gesture did have a positive benefit on learning on a classroom-wide level in a study of conservation (Church et al., 2004). Thus, the second goal of the present research was to investigate whether seeing gesture would affect mathematical equivalence learning on a classroom-wide level.

We provided a lesson in mathematical equivalence to classrooms of elementary school children. Children who are learning the fundamentals of mathematics have difficulty acquiring conceptual knowledge of the equal sign. These children often generalize procedural interpretations of the equal sign, interpreting it as a cue to combine terms rather than as a symbol expressing a relation between quantities (Kieran, 1981; McNeil & Alibali, 2005). Importantly, these generalizations can be avoided by simply changing what learners see. In one study, students who saw equations during training that involved arithmetic (e.g., \(15 + 13 = 28\)) performed significantly worse on a posttest in equivalence (\(4 + 2 + 6 = \_\_ + 6\)) than students who saw equations that did not involve arithmetic (e.g., \(28 = 28\)). Because students in each group heard the exact same spoken instructions, this suggests that children’s conceptualization of equality may be based in part on their visual experience during learning (McNeil, 2008).

In this study, children were instructed in mathematical equivalence with speech and gesture or with speech alone. They were tested immediately after training and after a 24-hr delay. In addition, we assessed transfer to other types of mathematical equivalence problems at the delayed test. We
hypothesized that providing learners with a representation in gesture would highlight the relational nature of the equal sign and facilitate acquisition and transfer of knowledge of equality, both immediately and after a 24-hr delay.

**Method**

**Participants**

One hundred and eighty-four children (88 male) were recruited for participation in this study. Children were either in the second \((n = 98)\), third \((n = 80)\), or fourth \((n = 6)\) grade (the fourth-grade students were recruited from a split third- and fourth-grade classroom). Children from a wide age range were included because previous research has demonstrated that understanding the equal sign is extremely difficult for American children across multiple grade levels (McNeil, 2007). We recruited a diverse sample from 22 classrooms in seven public elementary schools in central Michigan. Average ethnic composition of the schools was as follows: 60.8% European American (range = 36%–95%), 20.8% African American (range = 3%–37%), 16.8% Hispanic (range = <1%–23%), 2.4% Asian American (range = 0%–6%), and 1.16% Native American (range = 0%–2%). Across the schools, the average proportion of students who qualified for free or reduced-price lunch was 55.4% (range = 8%–81%).

Sixty-eight additional students participated in the study but were excluded from all analyses because they either failed to complete all parts of the study \((n = 18)\), visibly copied off of another student \((n = 2)\), or got at least one problem correct on the pretest \((n = 48)\). Of those who obtained at least one problem correct on the pretest, 10 students were in the second grade (9% of second-grade students) and 38 were in the third grade (32% of third-grade students). No fourth-grade students solved an equation correct on the pretest. Consent forms were distributed to parents a week prior to data collection. Students who did not return a consent form were still permitted to engage in the activity, but their work was discarded at the conclusion of the session. In addition, each child provided either written (third- and fourth-grade students) or verbal (second-grade students) assent prior to participating in the experiment.

**Materials**

**Tests.** Five tests were used in this experiment: three tests of mathematical equivalence with respect to addition and two transfer tests (see Appendices A and B for test materials). The tests of mathematical equivalence contained six equivalence problems with an equivalent addend on either side of the expression. Each problem featured three single-digit addends on the left side of an equal sign. On the right side of the equal sign, one of the addends from the left side appeared again and was adjoined with a blank space. Students had to determine what number should be placed in the blank to ensure that the sum of the values on the left was equivalent to the sum of the values on the right. For example, the problem \(7 + 2 + 9 = 7 + \_\) is properly solved by placing the number 11 in the blank. Each equation on every test was novel and was not repeated elsewhere in the experiment.

The transfer tests were designed to measure students’ ability to generalize learning to new types of equivalence problems. Because of the wide age range of children participating in this study, two separate transfer tests were constructed: one for second-grade students and one for third- and fourth-grade students. Both transfer tests contained four problems similar in structure to those found on the other tests and seen during instruction. The transfer test for second-grade students contained equivalence addition problems that used nonequivalent addends. Thus, none of the numbers from the left side of the problem appeared on the right side of the problem (e.g., \(4 + 5 + 7 = 3 + \_\)). These problems allow children to use the same strategy used during instruction, but are more difficult because they do not allow for the same calculation procedures to be used. Equivalence problems with nonequivalent addends have been used as transfer problems in previous work and children who correctly solve equivalence problems with repeated addends do not always transfer their understanding to problems without repeated addends (Alibali, 1999; Rittle-Johnson & Alibali, 1999).

The transfer test for third- and fourth-grade students used multiplication in place of addition and used equivalent multiplicands (e.g., \(6 \times 2 \times 3 = 6 \times \_\)). These problems also allow children to use the same strategy used during instruction but require children to implement this strategy with a different mathematical operation. Thus, both transfer tests required students to generalize, either to a new problem structure or to a new problem operation.

**Training videos.** Twelve videos were created to illustrate how to solve equivalence addition problems. The average duration of the videos was 15.5 s (SD = 1.38). In each video, an equivalence problem with equivalent addends was displayed on a
screen, and an instructor stood alongside it and explained how to solve a single equivalence problem. In half of the videos, the instructor gestured while she explained the problems and in the other half, she did not gesture.

The instructor explained how to solve each problem by first verbalizing the equalizer strategy. This technique stresses that both sides of an equation must yield the same numerical value and that each side can be calculated separately (Goldin-Meadow, Kim, & Singer, 1999). Thus, to explain the problem $8 + 6 + 2 = \_ + 2$, the instructor said, “I want to make one side equal to the other side. Eight plus six plus two is sixteen, and fourteen plus two is sixteen. So, one side is equal to the other side.” In the gesture videos, the instructor also employed two hand gestures while speaking, using different hands to refer to the two sides of the equation. Whenever she said the words “one side,” she swept her left hand back and forth beneath the left half of the equation and when referring to “the other side,” she swept her right hand back and forth beneath the right half of the equation. The speech-alone videos featured the same instructor explaining the same problems, using exactly the same verbal scripts and intonation. However, in these videos, the instructor’s hands rested naturally at her side, and no gestures were produced. Thus, each problem used during training had both a gesture and speech-alone video.

Pairs of stimulus videos were created for this study. The gesture and speech-alone videos were matched for intonation and rhythm of speech so that the only difference between the two videos was the instructor’s use of hand gesture. The stimuli were produced by an experimenter (SWC) who has extensive prior experience conducting live experiments with these exact instructions. During the videos, the instructor stood to the left side of the problem with her body oriented toward the problem. She was not directly facing the problem; the observers were able to see the right side of her face during the entire video. The instructor used standardized eye gaze, turning to and looking at each portion of the equation as she referred to it. The instructor also used standardized intonation. To ensure that the audio tracks were comparable across conditions, we first recorded multiple takes of the instructor explaining 12 candidate problems with gesture and in speech alone. We then selected matched videos of each of the 12 candidate explanations and played the audio from pairs of videos (gesture and speech alone) concurrently and selected the 6 stimulus pairs that we judged to have the least amount of perceivable variation across tracks. Example stimuli and audio amplitude plots can be viewed online (see online supporting information Figure S1).

Procedure

All participants from a given class took part in the experiment simultaneously in their regular classroom environment. Each classroom was randomly assigned to either a speech and gesture or speech-alone condition. The only difference between the training in the two conditions was the videos used during training. Children in the gesture condition watched the six videos including gesture and children in the speech-alone condition watched the six videos that did not contain gesture. Whenever possible, care was taken to run at least one gesture and one speech-alone classroom in each school to counterbalance potential effects of location.

Each student was first given a pretest and had unlimited time to complete it. Children were then told that they would watch some videos to learn how to solve similar problems. During training, the children first viewed an individual video and then solved a single equivalence problem (Appendix A). This sequence (video, then problem) was repeated six times for a total of six training videos and six problems. Videos were projected on a screen at the front of the classroom and auxiliary computer speakers were used to ensure that all participants could hear the instructor’s speech. The audio was loud enough to be heard clearly from the back of the classroom. The order of the videos and training problems was fixed; all participants saw the videos and completed the problems in the same order.

Immediately following training, the children were asked to complete the first posttest and were given as much time as they needed. After this, the students were told that the experimenters would return the next day, but were not informed that they would be tested again. Twenty-four hours later, the experimenters returned to the classroom and administered the second posttest and the transfer test, again placing no time constraints on the participants. Because the problems on all of the tests were unique, students were unable to simply memorize answers from earlier portions of the experiment and needed to rely on their ability to adaptively employ the strategy presented during training.

Results

We used a multilevel logistic regression model to account for variability across individual subjects as
well as variability in problem difficulty. We predicted the log of the odds of correctly solving each problem on each of the tests (analyzed individually), with instructional condition as the factor of interest, and grade, subject, and problem as random effects. The problem effect accounted for the different transfer test given to second-grade students. We did not include class or school as random effects in the analysis because models with both class and grade, school and grade, or school and class did not always converge in our preliminary analyses, and the effect of grade was much more robust than the effect of class or school. The results reported here do not change substantially if class is considered a random effect instead of grade. The results also do not change substantially if socioeconomic status (as measured by the percentage of students at each school who qualify for free or reduced-price lunch) is included in the model as a control parameter to account for variability across schools.

There was a main effect of gesture condition on each test, with children in the speech and gesture group performing better than the speech-alone group on Posttest I ($\beta = 3.12$, $z = 3.24$, $p < .01$), Posttest II ($\beta = 3.13$, $z = 3.97$, $p < .0001$), and the transfer test ($\beta = 2.24$, $z = 2.66$, $p < .01$; see Figure 1). We also analyzed performance on all tests in a single analysis. There was a main effect of gesture ($\beta = 1.42$, $z = 2.27$, $p = .018$) and a significant interaction between condition and test ($\beta = 0.61$, $z = 2.08$, $p = .038$). Performance in the gesture group reliably improved from Posttest I to Posttest II, whereas the speech-alone group did not show improvement. Children in the gesture group also tended to perform better than the speech-alone group on the transfer test, relative to each group’s posttest performance ($\beta = 0.50$, $z = 1.47$, $p = .14$), but this effect failed to reach significance.

In a more direct test of the effect of gesture over time, we analyzed performance on Day 2 tests (Posttest II and transfer test), controlling for performance on Posttest I, allowing us to take into account the differential performance immediately after instruction on Day 1. The effect of gesture on Day 2 performance was reliable even when Day 1 performance was included as a covariate, with children in the gesture group performing better than children in the speech-alone group ($\beta = 1.05$, $z = 2.27$, $p = .023$). In this analysis, the interaction between condition and test was not significant, suggesting that gesture during instruction facilitated both retention and transfer after 24 hr, above and beyond what can be accounted for based on the amount of initial learning. Thus, the differential effects observed on Posttest II and on the transfer test were not simply due to children in the speech-alone group doing worse immediately after training. Instead, gesture affects initial learning, and additionally affects how learning is consolidated over time.

**Discussion**

In this study, we found a robust effect of observing gesture on both initial learning and maintenance of learning. After observing gesture, learning increased over the following 24 hr, and this increase was independent of the amount of initial learning. In addition, we extended previous findings on the efficacy of gesture as an instructional tool even when instruction is not individualized. When learning to solve abstract problems that require algebraic reasoning, children can benefit from gesture presented on video at the classroom level. Finally, we found that gesture also facilitates transfer of learning 24 hr after initial acquisition.

How does observing gesture facilitate learning? Because the audio tracks were matched across our experimental conditions, we can rule out the explanation that the beneficial effect of gesture on learning is driven by changes in prosody that may be associated with gesture production. This is always a potential confound in studies using live presentation of hand gesture. However, our videos were carefully produced and selected to ensure that prosody did not differ between conditions.

One possibility is that gesture may simply engage student attention in a distracting environment. Gesture is a dynamic visual representation
and is likely to be effective at capturing attention. Indeed, instruction with gesture has been shown to lead learners to turn their head away from the video less than comparable instruction without gesture (Valenzano et al., 2003). However, we believe that increased attention is unlikely to explain all of the effects we observed. If attention underlies learning, then students who showed similar amounts of learning should have also attended similarly during instruction. Therefore, we would expect that students who performed similarly on the first posttest should also perform similarly on the second posttest. However, we found an effect of gesture on the second posttest even after controlling for performance on Day 1, suggesting that gesture has a unique effect on learning, above and beyond what would be expected by increased attention.

Instead, we propose that gesture is changing something about the knowledge that children acquire from the instruction and this leads to improved performance across time. One possibility is that gesture may clarify or provide conceptual information that is not readily apparent in the accompanying speech. One important aspect of mathematical equivalence is understanding that an equation actually consists of two sides that need to be related to one another, rather than a simple string of numbers that should be operated over (McNeil & Alibali, 2004). Children typically do not appropriately encode equivalence problems and appropriate encoding is associated with success in solving equivalence problems (McNeil & Alibali, 2004). In our instructions, gesture may help to clarify the meaning of the word \textit{side}, leading children in the gesture condition to be more likely to encode the problem correctly and form correct conceptual representations. Indeed, there is evidence that gesture can improve understanding of the message in speech. Both children and adults use gestural cues to disambiguate statements with multiple possible interpretations (Kelly, 2001; Kelly, Barr, Church, & Lynch, 1999) and gesture can also help to clarify ambiguous language. Children are more likely to correctly identify poorly articulated words if the speaker points to the objects that are being labeled (Thompson & Massaro, 1994).

Although we believe that gesture may change the way that students understand the accompanying speech, it is likely that this effect is not unique to gesture. Alternative manipulations of either the spoken instructions or the training materials may also help children form appropriate conceptual knowledge and produce results similar to the pattern of performance seen in the gesture group. For example, it is possible that if the instructor explained more fully what is meant by a \textit{side}, or the equations were written with a considerable amount of blank space between the two sides of each equation, students might develop greater understanding of the concept of side. Future work will be necessary to determine whether gesture is a particularly effective or unique tool for influencing student understanding.

Gesture may also enhance learning via mechanisms that are more difficult to engage with other auxiliary aids. For instance, gesture may promote lasting learning by creating motor representations for mathematics problem solving. It is well established that motor representations are particularly effective at creating lasting knowledge. Producing actions can increase memory for action words (cf. Cohen, 1981; Cohen & Stewart, 1982; Saltz & Donnenwerth-Nolan, 1981) and motor simulation may similarly affect memory for accompanying words and concepts. Although mathematics is typically thought of as an abstract domain, mathematical concepts may be supported by more concrete visual and motor representations (Cook, 2011; Lakoff & Núñez, 2000). Gesture may facilitate learning by providing a procedural, motor representation that is complementary to the basic declarative understanding of the problem conveyed in speech and that is particularly likely to support lasting learning.

On this account, observing gesture would need to facilitate motor encoding in observers. Work in a variety of paradigms suggests that observed actions recruit similar neural networks as those used to produce the same action (see Rizzolatti & Craighero, 2004, for a review) and that this motor simulation can facilitate understanding of observed actions (Knoblich & Sebanz, 2006; Prinz, 1997; Reed & Farah, 1995). Moreover, the motor system is engaged more when gesture accompanies speech than when gesture does not accompany speech (Skipper, Goldin-Meadow, Nusbaum, & Small, 2009; Skipper, van Wassenhove, Nusbaum, & Small, 2007). This suggests that listeners may simulate motor actions when observing gesture and that gesture observation may implicate motor regions during communication. This motor representation may help to create more enduring memory, potentially by providing a more proceduralized representation of the information.

The motor representation may also provide an especially robust cue to problem structure. In our instruction, the same gestures were repeated across multiple problems. However, because the numbers in the problems changed, the problems themselves and the spoken explanations of these problems...
varied. Gesture may provide learners with a cue to the underlying problem structure that is especially portable and can be directly applied to other similar problems (Alibali & Nathan, 2007; Richland, Zur, & Holyoak, 2007). Gestures may be unique in this property—gestures are both abstracted from and dependent on their context for meaning, which allows them to simultaneously express abstract relations while linking these relations to specific entities available in the environment. In contrast, the specificity of spoken reference may make it difficult to simultaneously encode both general and specific information.

We have shown that seeing gesture does not simply facilitate performance directly after training. Instead, additional effects of gesture emerge with time; performance in the gesture group increases across a 24-hr interval. One explanation of this finding is that gesture may interact with offline memory processing during sleep, or consolidation, in the intervening period. Consolidation can strengthen memory and even enhance performance above levels seen immediately after training, similar to the pattern observed here (see Diekelmann & Born, 2010; Margoliash & Fenn, 2008; McGaugh, 2000, for reviews). Here, we show that memory with gesture is enhanced after a 24-hr interval that includes a sleep phase so it is possible that these effects may be the result of sleep-dependent consolidation. Indeed, sleep-dependent consolidation has been found in learning in the motor system (cf. Brawn, Fenn, Nusbaum, & Margoliash, 2010), which we propose is recruited when observing gesture.

In conclusion, hand gesture not only facilitates understanding in the moment, but also affects how knowledge is represented over time. Gestures can improve learning even when presented to classrooms of children via videotaped instruction. Thus, one way to facilitate mathematics learning in children is to provide learners with a helping hand.

References


### Appendix A: Test and Training Equations With Equivalent Addends

<table>
<thead>
<tr>
<th>Pretest</th>
<th>Immediate posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 + 6 + 4 = ___ + 4</td>
<td>6 + 3 + 4 = ___ + 4</td>
</tr>
<tr>
<td>5 + 3 + 6 = 5 + ___</td>
<td>7 + 2 + 4 = 7 + ___</td>
</tr>
<tr>
<td>6 + 8 + 7 = ___ + 7</td>
<td>5 + 6 + 8 = ___ + 8</td>
</tr>
<tr>
<td>9 + 8 + 7 = 9 + ___</td>
<td>5 + 7 + 8 = 5 + ___</td>
</tr>
<tr>
<td>2 + 9 + 6 = ___ + 6</td>
<td>8 + 3 + 5 = ___ + 5</td>
</tr>
<tr>
<td>5 + 9 + 8 = 5 + ___</td>
<td>8 + 7 + 6 = 9 + ___</td>
</tr>
</tbody>
</table>

### Training

<table>
<thead>
<tr>
<th>24-hr posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 + 9 + 8 = 4 + ___</td>
</tr>
<tr>
<td>7 + 5 + 8 = ___ + 8</td>
</tr>
<tr>
<td>3 + 8 + 7 = 3 + ___</td>
</tr>
<tr>
<td>3 + 2 + 7 = ___ + 7</td>
</tr>
<tr>
<td>9 + 8 + 4 = 9 + ___</td>
</tr>
<tr>
<td>6 + 8 + 5 = ___ + 5</td>
</tr>
</tbody>
</table>

### Appendix B: Transfer Test Items

#### Transfer test for second-grade students

| 6 + 4 + 2 = ___ + 3 |
| 4 + 5 + 7 = 2 + ___ |
| 5 + 2 + 3 = ___ + 4 |
| 6 + 2 + 3 = 5 + ___ |

#### Transfer test for third- and fourth-grade students

| 1 × 4 × 2 = ___ × 2 |
| 4 × 5 × 4 = 4 × ___ |
| 5 × 2 × 3 = ___ × 3 |
| 6 × 2 × 3 = 6 × ___ |

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s website:

**Figure S1.** Audio Amplitude Plots From the Gesture (top) and No Gesture (bottom) Stimulus Videos.