



Learning math by hand: The neural effects of gesture-based instruction in 8-year-old children

Elizabeth M. Wakefield¹ · Eliza L. Congdon² · Miriam A. Novack³ · Susan Goldin-Meadow⁴ · Karin H. James⁵

Published online: 20 May 2019
© The Psychonomic Society, Inc. 2019

Abstract

Producing gesture can be a powerful tool for facilitating learning. This effect has been replicated across a variety of academic domains, including algebra, chemistry, geometry, and word learning. Yet the mechanisms underlying the effect are poorly understood. Here we address this gap using functional magnetic resonance imaging (fMRI). We examine the neural correlates underlying how children solve mathematical equivalence problems learned with the help of either a speech + gesture strategy, or a speech-alone strategy. Children who learned through a speech + gesture were more likely to recruit motor regions when subsequently solving problems during a scan than children who learned through speech alone. This suggests that gesture promotes learning, at least in part, because it is a type of action. In an exploratory analysis, we also found that children who learned through speech + gesture showed subthreshold activation in regions outside the typical action-learning network, corroborating behavioral findings suggesting that the mechanisms supporting learning through gesture and action are not identical. This study is one of the first to explore the neural mechanisms of learning through gesture.

Keywords Gesture · Learning · Mathematics · Neuroimaging · Neural mechanisms

Introduction

When people talk, they spontaneously move their hands – they gesture. Importantly, these movements are not mere hand waving. Gesture can be used to express and manipulate ideas, and asking learners to produce gesture facilitates learning in many domains – mathematics (e.g., Congdon et al., 2017; Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Cook, & Mitchell, 2009; Perry, Church, & Goldin-Meadow,

1988; Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018), quantity concepts (e.g., Church & Goldin-Meadow, 1986; Pine, Lufkin, & Messer, 2004), language concepts and vocabulary (LeBarton, Goldin-Meadow, & Raudenbush, 2014; Macedonia, Muller, & Friederici, 2011; Wakefield, Hall, James, & Goldin-Meadow, 2018; Wakefield & James, 2015), and even moral reasoning (Beaudoin-Ryan & Goldin-Meadow, 2014). Although gesture is recognized as a powerful learning tool, the mechanisms underlying its impact are not well understood.

We begin to address this gap by identifying the neural correlates associated with learning via gesture. We taught children to produce a strategy for solving mathematical equivalence problems either through speech *and* gesture or through speech alone. After this active training period, we used functional magnetic resonance imaging (fMRI) to compare the neural patterns recruited by each group of children as they solved additional problems during a scan, without speech or gesture.

Missing addend mathematical equivalence problems (e.g., $3+4+5= _ +5$) are an important precursor to algebra, yet children struggle with these problems (e.g., Falkner & Karen, 1999; Ginsburg, 1989; Saenz-Ludow & Walgamuth, 1998). In previous behavioral studies, children have been taught

✉ Elizabeth M. Wakefield
ewakefield1@luc.edu

¹ Department of Psychology, Loyola University Chicago, Chicago, IL, USA

² Department of Psychology, Bucknell University, Lewisburg, PA, USA

³ Department of Psychology, Northwestern University, Evanston, IL, USA

⁴ The University of Chicago, Chicago, IL, USA

⁵ Department Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

problem-solving strategies in speech to help them better understand these problems, and research has shown an added benefit when children are asked to produce gesture along with the speech strategy during a lesson (Cook et al., 2008; Goldin-Meadow et al., 2009). For example, Cook et al. (2008) taught children to say, “I want to make one side equal to the other side,” during a math lesson. Some children were taught only the words; others were taught the words plus gestures conveying the same strategy (e.g., underlining the left side of the equation with the hands, and then underlining the right side). Children who produced the words and gestures during instruction retained what they had learned better than children who produced only the words (Cook et al., 2008; see also Goldin-Meadow et al., 2009).

There are many characteristics of gesture that could contribute to its ability to help children learn new concepts, including the fact that it can direct a learner’s attention (e.g., Wakefield, Novack, et al., 2018) or help to spatially represent an idea (Cooperrider, Gentner, & Goldin-Meadow, 2016). Here we focus on the fact that gesture is a type of action and thus engages the motor system when produced. There is reason to believe that the neural patterns underlying learning through gesture are similar to those underlying learning through action simply because gesture is a type of action. Gesturing, like acting on objects, involves creating a motor plan (Hostetter & Alibali, 2008). But, unlike acting on objects, which has a direct effect on the external world, gesturing has its effect by representing information (e.g., producing a twisting gesture near a jar lid does not, in and of itself, result in getting the lid off the jar; in contrast, producing a twisting action on the lid does).

Focusing on a gesture’s identity as an action has particular relevance for developmental populations, as active learning has long been a central theme of early development (e.g., Piaget, 1952). Moreover, recent developmental studies have found support for the role action plays in learning domains as varied as social-cognitive processing (Gerson & Woodward, 2014; Sommerville, Woodward, & Needham, 2005), reading comprehension (Glenberg, Brown, & Levin, 2007), word learning (Smith, 2005), and mental rotation (Wiedenbauer & Jansen-Osmann, 2008). Action is particularly relevant for mathematics education given the widespread use of manipulatives (physical objects that instantiate mathematical concepts) in elementary school classrooms (see Mix, 2010, for review).

Perhaps because of these widespread effects, the mechanisms driving action-based changes in the brain have been well studied. Findings from the action-learning neuroimaging literature indicate that actively learned information recruits sensory and motor areas more heavily than passively learned information in both children (James, 2010; James & Bose, 2011; James & Swain, 2011; Kersey & James, 2013) and adults (e.g., Butler & James, 2013; Butler, James, & James,

2011; James & Atwood, 2009; James & Gauthier, 2006; James & Maouene, 2009; Longcamp, Anton, Roth, & Velay, 2003). For example, when actively exploring a novel object, both children and adults coordinate sensory-motor processes and, in so doing, change the functional network involved in processing the object. This motor experience leads to augmented sensory-motor representations that can be drawn on later when related stimuli are presented (James, 2010; Longcamp et al., 2003; Prinz, 1997). Importantly, this motor activation has been seen in participants who are lying still in the scanner and shown objects that they had previously learned to act on (e.g., Butler & James, 2013; James & Bose, 2011; James & Swain, 2011; for similar findings in letter writing action, see James, 2010; James & Atwood, 2009). Learning through action thus has lasting effects that continue to influence processing, even when learners are no longer producing any actions at all.

By contrast, the mechanisms underlying a gesture’s effects on learning are poorly understood. In the current study, we use neuroimaging to examine these mechanisms in a math-learning paradigm. Because the facilitative effects of gesture on learning mathematical equivalence are well documented, this paradigm provides an excellent test-bed to explore the neural mechanisms involved in learning with gesture. We limit our neuroimaging investigation to those children who have *successfully learned* mathematical equivalence through either speech and gesture or through speech alone. Any differences that we find in patterns of neural activation between the two groups must then reflect processing differences, rather than differences in levels of learning; in other words, differences in *how* the task was learned, rather than differences in *how well* the task was learned. If the learning effects of gesture stem from the fact that gesture is a type of action, then we should see significantly greater recruitment of frontal-parietal motor regions, regions known to be activated after learning through action (e.g., James & Swain, 2011), in children who learned through speech and gesture than in children who learned through speech alone.

Method

Participants

Twenty 7- to 9-year-old children ($M=8.0$ years, eight males) were included in the final analysis. Children were recruited from a larger behavioral study testing how children learn math through gesture. Children were eligible for this study if they failed to correctly solve any of the mathematical equivalence problems on the pretest prior to training *and* correctly solved at least half of the problems on the post-test after training. All of the children who met these criteria from the larger behavioral study were invited to participate (see *Procedure* for

details). All participants were right-handed, native English speakers with no reported history of neurological or psychiatric disorders, and had normal visual acuity. Informed consent was obtained from participants' parents or guardians, and assent was obtained from children. An additional nine children participated but were not included in the final analysis (two refused to complete the experimental session, seven moved excessively during the fMRI session).

Behavioral procedure

One week prior to the fMRI procedure, children participated individually in a one-on-one instructional session modeled after previous studies of math learning using speech and gesture strategies (Cook et al., 2008; Goldin-Meadow et al., 2009). At pretest, children completed six missing-addend mathematical equivalence problems. Three problems were of the form $a+b+c= _+c$ and three problems were of the form $p+q+r=p+_$. Children who answered none of the six problems correctly were then assigned to one of two training conditions. Children in the speech-alone condition learned the phrase, "I want to make one side equal to the other side" (the equalizer strategy). Children in the speech + gesture condition learned the same phrase along with an equalizer gesture: underlining the left side of the problem with the left hand (while saying "one side") and underlining the right side of the problem with the right hand (while saying "other side").

Both groups were then given the same math lesson. The experimenter first taught the child how to solve a mathematical equivalence problem using the equalizer strategy in speech (without gesture), and then asked the child to solve another problem. Children were told to produce the words or words + gestures they had learned before and after attempting to solve the problem. The experimenter repeated this procedure until she had demonstrated six problems and the child had attempted to solve six problems, all of which were of the form $a+b+c= _+c$.

After the lesson, children took a post-test comparable to the pretest. Children who answered at least three of six problems correctly were invited to participate in a neuroimaging session. We limited our fMRI study to children who had reached this criterion because we were interested not in whether children *could* learn, but in *how* they had learned, and how that learning would be reflected neurologically. Ten children were recruited from each training condition and were matched on age (speech alone: $M_{\text{age}} = 7.98$ years; speech + gesture: $M_{\text{age}} = 8.05$ years, $t(18) = .75$, *ns*). A between-samples t-test confirmed that there was no significant difference in the number of correct answers children gave between groups (speech alone: $M = 4.4$, $SD = 1.4$; speech + gesture: $M = 4.1$, $SD = 1.4$; $t(18) = 0.49$, *ns*).

Neuroimaging procedure

fMRI test stimuli

Mathematical equivalence problems of the form $a+b+c= _+c$ were presented during a scan.¹ Children were asked to solve a problem and then press a GO button displayed on the lower right-hand side of the screen (Fig. 1). The problem then disappeared for 1 s and reappeared with a number in the blank. Children were then asked to compare their answer to the answer the computer produced in order to assess its correctness. They were told to press an \times button if the answer did not match the one they generated and a \checkmark button if the answer matched the one they generated.

fMRI session

The imaging session took place a week after the behavioral session. Children completed a follow-up test comparable to the behavioral session post-test. They then participated in a shortened version of the training session, during which they solved three problems, in order to remind them of the strategies they had learned the previous week. As in the original session, children produced either words or words+gestures depending on their training condition. Next, they were introduced to the imaging environment via a MRI simulator in which they listened to the sounds emitted by both functional and structural scans (for more on this methodology, see James & Maouene, 2009). Finally, they were taught what to do during the actual fMRI scan, including how to make responses by pressing one of two large buttons with either their right or left hand. The buttons were attached to a Plexiglass table that was slightly angled and designed to allow children to rest their hands while supine in the scanner. If comfortable after this exposure and instructional session, children then participated in a 25-min fMRI scanning session. Children were monitored during the scan by a trained research assistant who was in the scanning room with them to ensure that no children produced gestures during the problem-solving runs.

During the scan, children completed five runs (one structural, four functional). We focus here on data acquired while children completed two problem-solving runs, which were event-related and self-paced (these runs were always the first and final function run; functional runs 2 and 3 were conducted as part of a larger study). During problem-solving runs, children answered a series of missing addend mathematical equivalence problems. The number of problems solved varied by child; children solved problems at their own pace until the run

¹ Children also solved problems of a different form ($p+q+r=p+_$) to test transfer ability, as problems of this form were not explicitly trained. Because transfer was not the focus of this paper, these problems are excluded from our analyses.

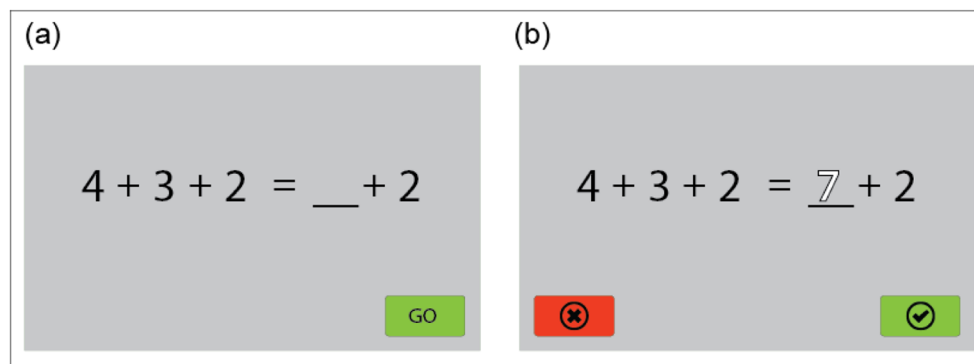


Fig. 1 An example of a problem presented to children during the problem-solving runs in the scanner. Children first saw display (a) and pressed GO when they had mentally solved the problem; they then saw

display (b) and pressed ✓ if their solution matched the number in the display or × if their solution did not match

automatically ended after 120 volumes had been collected (4 min). Children were taught a “Problem Game,” before the scan. For each missing addend problem, children mentally solved the problem, and then pressed the “Go” button with their right hand (Fig. 1a). The problem disappeared for 1 s and then reappeared with an answer in the blank (Fig. 1b). Answers displayed by the computer were either correct or incorrect (incorrect answers were off by ± 2 or 3).² Children were told that sometimes the computer would be “right” and sometimes it would be “wrong,” and their job was to tell the computer when it had done a bad job solving the problem. Thus, when an answer appeared in the blank, they were to indicate whether the computer had correctly or incorrectly answered the problem via a button press (× for incorrect answers; ✓ for correct answers). Evidence that children were compliant with our instructions to solve the problem in their heads is twofold: First, children were highly accurate – we would not expect this to be the case if they were not taking time to solve problems. Second, children spent an average of 17.59 s ($SD = 9.45$ s) solving problems. If children were not trying to solve the problems, we would expect much shorter solution times.

After children responded, there was a rest period of 1, 3, or 5 s (randomized) before another problem appeared. We extracted functional data during the time when children were solving each math problem, irrespective of the child’s subsequent behavioral response. This portion of the trial took place between the appearance of each problem and the point at which children pressed the “Go” button.

We decided to use a self-paced, as opposed to a fixed, design for the problem-solving runs because children vary widely in the amount of time they take to solve mathematical

equivalence problems, and because previous work (Krinzinger et al., 2011) had suggested an advantage for the self-paced design over the fixed design for this type of task. In a fixed design, a child would be given a set amount of time, say 20 s, to solve a problem, but the child might take only 10 s to solve the problem. The remaining 10 s would still be counted as “problem solving,” but as the child was not actually solving the problem, the information collected during this period would add noise to the data. By using a self-paced design, we eliminated the potential for off-task activation, making data collection less noisy – children were guaranteed to be *on-task* during all of the epochs extracted for analysis. Krinzinger et al. (2011) have indeed found that data collected using a self-paced design are more reliable and reproducible than data collected using a fixed design. They also established that, on an arithmetic task similar to ours, there was sufficient power to conduct their analyses when children spent between 36 and 144 s solving the problems. Children spent a comparable, or greater, amount of time solving the problems in our study ($M=159$ s, $SD=44.6$), thus suggesting that we have sufficient power to conduct our analyses.

fMRI parameters

Imaging occurred in a 3-Tesla Siemens Magnetom Trio whole-body MRI system and a phased-array 12-channel head coil in the Imaging Research Facility located in the Indiana University Psychological and Brain Sciences department. The field of view was 220 mm with an in-plane resolution of 64×64 pixels and 33 slices per volume (3.4 mm thick, no gap). Images were acquired using an echo-planar technique (echo time, $TE=30$ ms, time of repetition, $TR=2,000$ ms; flip angle= 70°) for blood oxygen level-dependent (BOLD) imaging. High-resolution T1-weighted anatomical volumes were acquired using a Turbo-flash 3-D sequence: $TI=900$ ms, $TE=2.67$ ms, $TR=1,500$ ms, flip angle= 9° , with 120 sagittal slices of 1.5-mm thickness, a field of view of 192×192 mm, and an isometric voxel size of 1.5 mm^3 .

² We were careful not to include incorrect answers that could have been generated by the types of strategies that children who do not how to solve these problems frequently use. We found in pilot work that presenting children with answers generated by incorrect strategies can lead children who have recently learned how to solve the problems to revert back to previous misconceptions. We did not want to destabilize our fragile learners and therefore chose not to display these types of incorrect answers.

fMRI data analysis procedures

Brain Voyager™ QX version 2.8 (Brain Innovation, Maastricht, Netherlands) was used for all analyses. During preprocessing, images underwent 3D motion correction and linear trend removal (additional steps for motion correction are detailed in the next paragraph), and a Gaussian spatial blurring (FWHM 6 mm) was applied. Individual functional volumes were co-registered with anatomical volumes using an intensity-matching, rigid-body transformation algorithm. Individual anatomical volumes for children were normalized to the Talairach reference frame using the eight-parameter affine transformation with parameters selected in relation to anatomical landmarks, identified manually (Talairach & Tournoux, 1988). Applying the same transformation to the co-registered functional volumes placed the functional data in a common brain space, allowing for comparisons across participants.

Statistical analyses were performed with a random-effects general linear model. Predictors were included for each task condition, generated from the timing of the experimental protocol, and then convolved with a two-gamma hemodynamic response function.

Controlling for excessive head motion is a special concern when imaging children (e.g., Poldrack, Pare-Blagoev, & Grant, 2002; Yerys et al., 2009; Yuan et al., 2009). We accounted for head motion in four ways: (1) Children with motion parameter estimates of movement that exceeded 6 mm in any axis were excluded from analysis. (2) During preprocessing, we used a motion-correction (re-alignment) algorithm to estimate and correct for motion. (3) Motion parameter estimates from the preprocessing stage were included in the GLM model as regressors of no interest (Evans, Todd, Taylor, & Strother, 2010; Poldrack et al., 2002). (4) From the motion estimates, severe motion “spikes” > 5 mm were identified and those time points were re-assigned from the original experimental condition in the protocol to a “motion spike” condition. Motion spike time points were added to the GLM model as dichotomous (and unconvolved) regressors of no interest. Using a combination of continuous time series motion parameter predictors and dichotomous “motion spike” predictors accounts for more error variance than using continuous time series motion parameters alone. Finally, in addition to these efforts to account for motion during the scan, we also measured the mean amount of head displacement during the scan, and found no difference between groups (speech alone: $M=0.79$ mm, $SD=0.49$, speech+ gesture: $M=0.84$ mm, $SD=0.55$, $t(18)=0.21$, *ns*).

Two whole-brain analyses were performed. In the first, we considered data from participants across training conditions. Individual data from all 20 participants were averaged together to create group statistical parametric maps. Correction for multiple comparisons was performed at the cluster level, using

the BrainVoyager™ Cluster-Level Statistical Threshold Estimator. We estimated the cluster-size threshold necessary to produce an effective alpha <.05 through a Monte Carlo simulation (1,000 iterations), given a voxel-wise p -value <0.005. This cluster-threshold (35 mm^3) was applied during the whole-brain analysis. In the second analysis, we looked directly at effects of training condition on children’s processing of mathematical equivalence problems. To this end, Group (speech alone vs. speech + gesture) was entered as a between-group variable into the ANCOVA module in BrainVoyager™, as a predictor of activation during problem solving (the portion of trials that took place between the appearance of each problem and the point at which children pressed the “Go” button). We estimated the cluster-size threshold necessary to produce an effective alpha <.05 through a Monte Carlo simulation (1,000 iterations), given a voxel-wise p -value <0.005 in each group separately. We then chose the more conservative (i.e., larger) cluster threshold to apply during the between-group analysis (29 mm^3).

Results

Behavioral performance and group matching

Before turning to our neuroimaging results, we need to establish that our two experimental groups are well matched. To this end, we compared the groups on: (1) general information about the participants (age at test, gender); (2) their performance on mathematical equivalence problems (e.g., $3+6+5= _ +5$), which were the focus of our training (number of problems solved during Problem Runs; average amount of time taken to solve problems during the Problem Runs; accuracy on problems at post-test, follow-up, and during Problem Runs); and (3) their general arithmetic skills (number of simple addition problems [$3+4= _$] solved in a 4-min period; accuracy on simple addition problems; number of arithmetic errors (e.g., stating that 3 plus 4 is 9) made across all behavioral time points; instances of finger-counting across all behavioral time points, a spontaneous index of dependence on hand movements). Our comparisons were made using both traditional hypothesis testing and Bayesian inference testing. The two types of tests came to the same conclusion – there was no evidence of significant differences between conditions in the behavioral results during the fMRI scan (see Table 1).

Mathematical equivalence problem processing

To ensure that results from our self-paced design were similar to previous studies on mathematical calculation processing in children, we conducted a whole-brain contrast, comparing activation to Problem Solving versus Rest across all individuals. Figure 2 shows a network of regions recruited for

Table 1 Comparison of the two experimental groups on behavioral measures

	Speech alone, M (SD)	Speech and gesture, M (SD)	T	p	Scaled JZA Bayes Factor
General Information					
Age at test	7.98 (0.44)	8.05 (0.58)	0.42	.75	3.01
Gender	6 females 4 males	6 females 4 males			
Performance on Mathematical Equivalence (3+4+5=___+5) Problem-Solving					
Time solving problems during scan (sec)	17.6 (9.7)	18.3 (9.6)	0.63	.87	2.76
Number of problems solved during Problem Solving Run	10.6 (3.1)	9.4 (3.3)	0.85	.41	2.42
Accuracy (P(c) based on d' analysis) during Problem Solving Run	0.96 (0.10)	0.95 (0.15)	0.30	.77	3.12
Number of problems solved correctly at behavioral posttest (out of 6)	4.4 (1.4)	4.1 (1.4)	0.49	.75	2.94
Number of problems solved correctly at behavioral follow-up (out of 6)	4.2 (2.3)	4.3 (2.0)	0.10	.91	3.22
Number of simple addition problems (3+4=___) solved in 4 min	19.0 (5.2)	20.4 (5.9)	0.56	.58	2.85
Accuracy on addition problems (P(c) based on d' analysis)	1.0 (0.0)	0.99 (0.004)	1.25	.08	1.75
Number of arithmetic errors made across all behavioral time points (pretest, training, post-test, follow-up test)	1.6 (2.0)	1.7 (1.8)	0.12	.91	3.22
Instances of finger-counting across all behavioral time points (pretest, training, post-test, follow-up test)	2.7 (3.8)	2.4 (2.5)	0.20	.84	3.19
General Math Ability					

mathematical problem solving that aligns with findings from previous studies with children (e.g., Davis et al., 2009; Kaufmann, Wood, Rubinsten, & Henik, 2011; Rivera, Reiss, Eckert, & Menon, 2005): Children displayed extensive activation in the superior parietal lobe bilaterally, which extended into the inferior parietal gyrus in the left hemisphere, encompassing the horizontal intraparietal sulcus (HIPS) (see Table 2 for locations of peaks and cluster size). These findings establish that children in our study displayed activation in a network known to be recruited by children during mathematical calculation.

Effect of training condition on mathematical equivalence problem processing

We then split the data by training condition to compare activation patterns between the two groups during the Problem Solving runs. Children who learned through speech + gesture showed significantly greater activation in a number of regions, compared to children who learned through speech alone (see Fig. 3; Table 2). This network included bilateral post-central gyrus, anterior to the general superior parietal activation shown by all children, extending into the left pre-central gyrus and the left inferior gyrus, as well as activation in the right middle temporal gyrus, right insula, and right supramarginal gyrus. These areas are similar (although not identical) to the network activated when individuals process information related to action-based learning: The region activated in the left inferior parietal cortex shows activation when individuals recall actions performed on objects (Johnson-Frey, 2004); the bilateral postcentral gyrus activation is slightly medial to activation found when children process stimuli related to actions performed on objects (James & Swain, 2011); and the precentral gyrus has also been implicated in this process (James & Swain, 2011). There were no regions that showed greater activation in children who learned to solve problems through speech alone.

From behavioral work, we know that action and gesture can affect learning in different ways (Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014). Given that the regions in our analyses reflected a general motor-learning network, we conducted an exploratory analysis of sub-threshold activation ($p < .01$). Along with the network already identified, activation was seen in the right angular gyrus (AG, x(53) y(-35) z(27)). Potential implications for this activation are addressed in the *Discussion*, although this activation must be interpreted with caution given the danger of false-positives in neuroimaging work (e.g., Eklund, Nichols, & Knutsson, 2016).

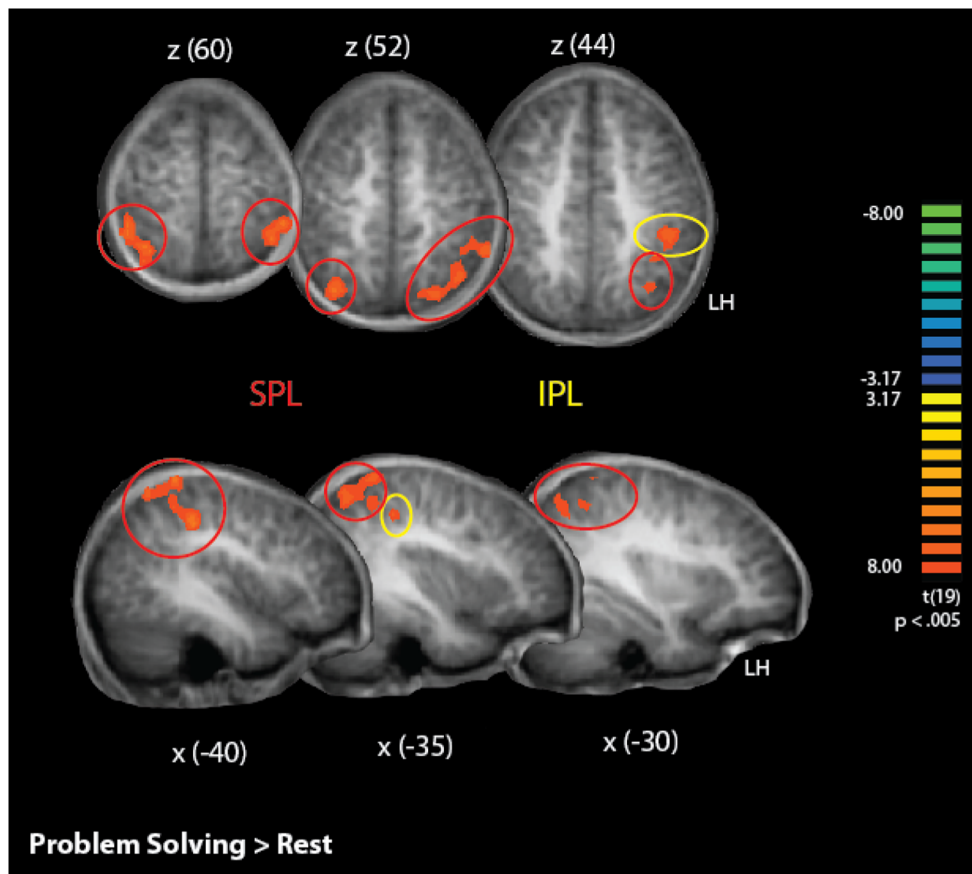


Fig. 2 Regions that showed greater activation during Problem Solving than during Rest (Fixation) for all children

Discussion

In this study, we directly investigated the neural effects of learning through self-produced gesture. Our results highlight properties of gesture that underlie its ability to support learning and insight; specifically, that gesture may help children learn because it engages the motor system in the learning process. As expected based on previous work, we found that all participants in our study activated regions associated with

arithmetic processing in children (e.g., Davis et al., 2009; Kaufmann et al., 2011; Krinzinger et al., 2011; Rivera et al., 2005). This analysis, intended as a manipulation check, suggests that our participants were actively solving the math problems during the scan. In this discussion, we focus on our second finding – differences in neural activation patterns between the two training conditions during problem solving. Children who had previously been trained to solve problems through speech and gesture showed greater activation in a

Table 2 Locations of peaks and cluster size

Contrast	Location	Cluster size	Peak coordinate	Peak t-value
Problem Solving > Rest	Left Superior Parietal Lobule	4,207	x(-39) y(-47) z(58)	5.4
	Right Superior Parietal lobule	2,212	x(41) y(-47) z(57)	4.8
	Left Inferior Parietal Lobule	863	x(-43) y(-35) z(38)	4.8
Problem Solving: Speech + Gesture > Speech Alone	Left Inferior Parietal Lobule	1,515	x(-40) y(-37) z(21)	3.6
	Left Postcentral/Precentral Gyrus	995	x(-28) y(-20) z(54)	3.7
	Right Postcentral Gyrus	1,116	x(23) y(-20) z(54)	3.6
	Left Precentral Gyrus	1,398	x(-18) y(-14) z(60)	3.8
	Right Anterior Middle Temporal Gyrus	922	x(44) y(13) z(-25)	3.8
	Right Posterior Middle Temporal Gyrus/Middle Occipital Gyrus	1,286	x(51) y(-62) z(18)	4.0
	Right Insula	584	x(43) y(-17) z(21)	3.6

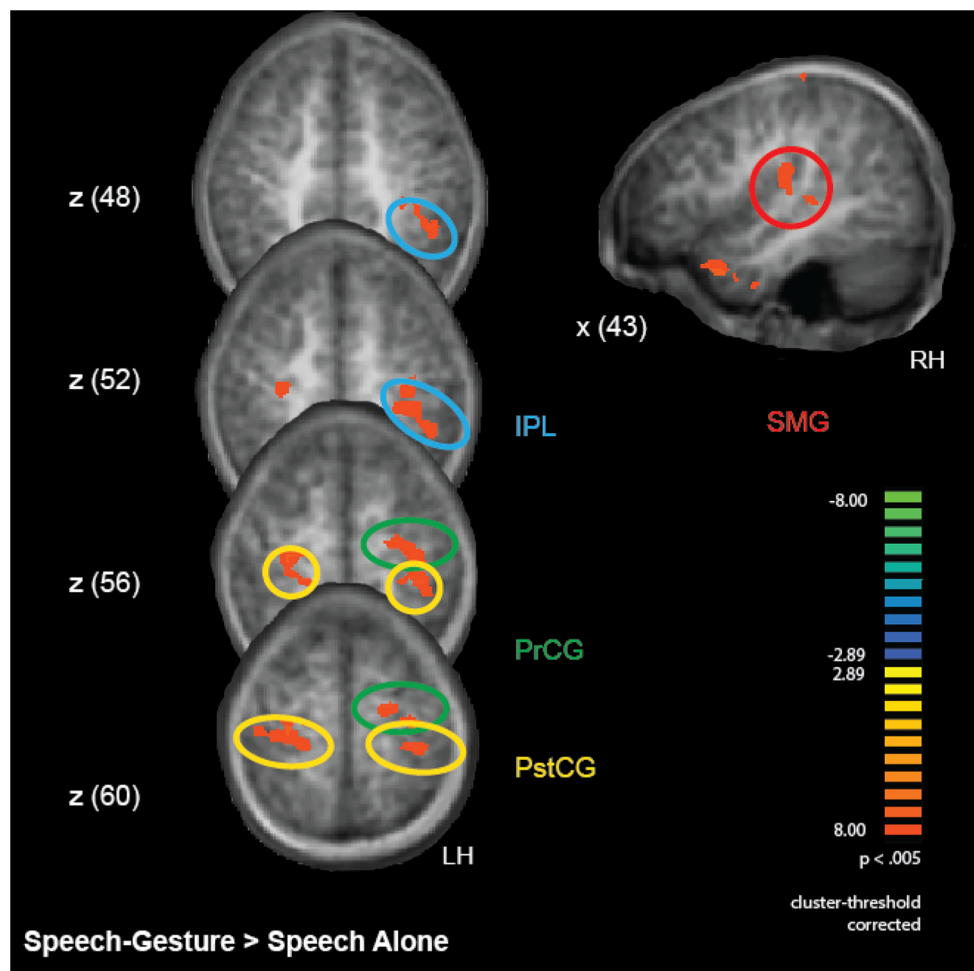


Fig. 3 Regions that showed greater activation during Problem Solving in children who learned through Speech and Gesture than in children who learned through Speech Alone

number of neural regions than children who had been trained to solve problems through speech alone.

Gesture as a type of action

It is well established that there are links between the actions we produce and our perceptual systems. After learning about an object through acting on that object, children show a neural trace of this learning experience when they subsequently perceive the object (James & Swain, 2011). For example, children who learn a novel action word (e.g., “yocking”) while producing an action (e.g., twisting the top of a strange looking object) recruit parietal regions associated with reaching and grasping when they later hear that word. They also recruit frontal premotor regions when shown the object that they acted on during learning. Research with adults shows a comparable effect – areas associated with reaching and grasping are recruited when adults view tools that they had previously actively manipulated (e.g., Johnson-Frey, 2004). Importantly, these effects of action experience occur even when action is not concurrently produced. A neural trace of action experience

is established, such that perceiving visual or auditory information *related* to a verb or object that has been associated with movement is enough to activate motor processing regions. Because gesture is a type of action, we hypothesized that learning through gesture would leave a similar neural trace to that of transitive action. In other words, we predicted that children who learned through gesture would recruit regions generally associated with motor movements, either frontal premotor regions or parietal regions, implicated in reaching and grasping (Culham & Valyear, 2006; Prado et al., 2005), even though children were *not* producing gestures in the scanner.

Our results suggest that gesture’s facilitative effects on learning may be grounded in the same kinds of mechanisms that underlie the facilitative effects found in other types of action. Children who learned through gesture showed significantly greater activation in a large region of parietal cortex bilaterally, extending into frontal motor regions, than children who learned without gesture. Thus, children who learned the math task while gesturing seemed to incorporate that gesture into their lasting understanding of how to solve problems, via

associating a motor representation with the form of the problem and/or with the spoken strategy. Just as children recruit sensorimotor regions when viewing objects that they had earlier learned about through self-produced actions, children may treat the visual form of the math equivalence problem as an object, and recruit sensorimotor regions when viewing and solving problems that they had earlier learned about through self-produced gestures. As another possibility, if children are recalling the spoken strategy they learned earlier when solving problems in the scanner, it may be that a motor association is linked to this “action-laden” speech for children who learned through self-produced gesture, similar to the patterns of activation seen when individuals hear verbs associated with motor actions (e.g., James & Swain, 2011; Pulvermüller, 2005). We cannot distinguish between these mechanisms (and they are not mutually exclusive). However, previous work has shown that gestures can facilitate learning even in the absence of a spoken strategy (Cook et al., 2008), which may make it more likely that the form of the problem (rather than the spoken strategy) is what children are linking to their motor representation.

This finding begins to shed light on *why* gesture is beneficial to learners. Gesture has been hypothesized to arise from the same neural system as the actions we use to move about in the world and to manipulate objects (Hostetter & Alibali, 2008). However, until now, the evidence for this claim has been circumstantial. The fact that we find a similar pattern of activation when children process information learned through gesture as when they process information learned through other types of action suggests that the beneficial effects of gesture on learning seen in previous studies (e.g., Cook et al., 2008; Goldin-Meadow et al., 2009) are likely to be driven, at least in part, by recruitment of the motor system during learning. Future work is needed to determine whether this pattern of activation also occurs when children have been taught through gesture experience, but have *not* successfully learned to solve mathematical equivalence problems. By including only children who had successfully learned to solve mathematical equivalence problems, we are able to make conclusions about *how* children learned to solve the task, not *whether* they learned to solve the task. However, our data leave open the possibility that the motor system may be differently engaged, depending on the extent to which a concept is understood.

Gesture beyond action

Our results are consistent with the idea that gesture supports learning because it is a type of action. However, we know from behavioral work that gesture and action are not identical in their effects on cognitive processes (e.g., Goldin-Meadow & Beilock, 2010; Kelly, Healy, Ozyurek, & Holler, 2014; Novack et al., 2014; Trofater, Kontra, Beilock, & Goldin-

Meadow, 2015). It is therefore not surprising that we found, in an exploratory analysis, sub-threshold activation in the angular gyrus, a region that is *not* specific to the network recruited when individuals process information learned through actions on objects. Pyke, Betts, Fincham, and Anderson (2014) recently showed that the AG is involved in mapping math problems to visuo-spatial referents. The fact that the AG was more heavily recruited by children in the speech + gesture group than by children in the speech-alone group might therefore reflect gesture’s capacity to provide a spatial framework for mathematical equivalence problems. But the AG has also been identified with verbal retrieval of math facts (Dehaene, Piazza, Pinel, & Cohen, 2003). Thus, an alternative possibility is that the enhanced activation of the AG in the speech + gesture group relative to the speech alone group reflects gesture’s ability to maximize learning that takes place through speech (e.g., Congdon et al., 2016). Future work is needed that directly compares the neural correlates underlying learning from gesture with learning from action in order to better understand the similarities, and the differences, between these processes. Paying particular attention to how well children generalize their knowledge after training, and how this generalization is instantiated in the brain, will be important, as previous behavioral studies suggests that gesture has the capacity to support more flexible learning than action on objects (Novack et al., 2014; Wakefield, Hall, et al., 2018).

Finally, we recognize that our relatively small sample size is a general limitation of the study, and that it contributes to the exploratory nature of this analysis. We therefore suggest that these secondary findings be taken cautiously and primarily as guiding hypotheses for future work.

Conclusions

Ours is the first study to use fMRI to investigate the impact that gesture production has on learning. The findings suggest that gesture may benefit learners, at least in part, because it involves the motor system in the learning process. Importantly, the neural differences between children who learned via speech and gesture, compared to children who learned via speech alone, cannot be attributed to behavioral differences between groups, as all of the children in the study learned how to solve the math problems during training. Moreover, the motor signatures found for children in the speech and gesture condition occurred in runs during which children were *not* producing gestures, providing evidence that, as in learning through action on objects, learning through gesture leads to a lasting and embedded neural trace of motor system involvement, which is activated when children later solve the math problems they had learned with gesture.

Acknowledgements The authors wish to acknowledge all the children and their parents who participated in this study; the research assistants who helped with the study, Arianna Gutierrez, Portia Goodin, Natalie Gutkin, Alyssa Kersey, Emily Kubota, and Debby Zemlock, our MR technicians, Arianna Gutierrez, Colleen McCracken, and Sean Berry, and MR physicist Dr. Hu Cheng. This research was supported by an IU Imaging Research Facility Pilot Grant, awarded to K. James, E. Wakefield, and S. Goldin-Meadow; the IU Faculty Research Support Program, awarded to K. James; the National Science Foundation (DRL-1561405) awarded to S. Goldin-Meadow and E. Wakefield; the National Institute of Health, T32 (HD07475); the Institute of Education Sciences (R305 B090025) to S. Raudenbush in support of Novack and Congdon; and the National Institute of Child Health and Development: Institutional Training Grant (HD07475-15) to L. Smith in support of Wakefield.

Data availability None of the data or materials for the experiments reported here are available, and none of the experiments were preregistered.

References

- Beaudoin-Ryan, L., & Goldin-Meadow, S. (2014). Teaching moral reasoning through gesture. *Developmental Science*, *17*, 984–990. <https://doi.org/10.1111/desc.12180>
- Butler, A. J., & James, K. H. (2013). Active learning of novel sound-producing objects: Motor reactivation and enhancement of visuo-motor connectivity. *Journal of Cognitive Neuroscience*, *25*, 203–218. https://doi.org/10.1162/jocn_a_00284
- Butler, A. J., James, T. W., & James, K. H. (2011). Enhanced multisensory integration and motor reactivation after active motor learning of audiovisual associations. *Journal of Cognitive Neuroscience*, *23*, 3515–3528. https://doi.org/10.1162/jocn_a_00015
- Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index on transitional knowledge. *Cognition*, *23*, 43–71. [https://doi.org/10.1016/0010-0277\(86\)90053-3](https://doi.org/10.1016/0010-0277(86)90053-3)
- Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O’Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction*, *50*, 65–74. <https://doi.org/10.1016/j.learninstruc.2017.03.005>
- Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O’Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction*, *50*, 65–74. <https://doi.org/10.1016/j.learninstruc.2017.03.005>
- Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*, 1047–1058. <https://doi.org/10.1016/j.cognition.2007.04.010>
- Cooperrider, K., Gentner, D., & Goldin-Meadow, S. (2016). Spatial analogies pervade complex relational reasoning: Evidence from spontaneous gestures. *Cognitive Research: Principles and Implications*, *1*, 1–28. <https://doi.org/10.1186/s41235-016-0024-5>
- Culham, J. C., & Valyear, K. F. (2006). Human parietal cortex in action. *Current Opinion in Neurobiology*, *16*, 205–212. <https://doi.org/10.1016/j.conb.2006.03.005>
- Davis, N., Cannistraci, C. J., Rogers, B. P., Gatenby, J. C., Fuchs, L. S., Anderson, A. W., & Gore, J. C. (2009). The neural correlates of calculation ability in children: An fMRI study. *Magnetic Resonance Imaging*, *27*, 1187–1197. <https://doi.org/10.1016/j.mri.2009.05.010>
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506. <https://doi.org/10.1080/02643290244000239>
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*, 7900–7905. <https://doi.org/10.1073/pnas.1602413113>
- Evans, J. W., Todd, R. M., Taylor, M. J., & Strother, S. C. (2010). Group specific optimisation of fMRI processing steps for child and adult data. *Neuroimage*, *50*, 479–490. <https://doi.org/10.1016/j.neuroimage.2009.11.039>
- Falkner, L., & Karen, P. (1999). Children’s understanding of equality foundation for algebra. *Teaching Children Mathematics*, *6*, 232–237.
- Gerson, S. A., & Woodward, A. L. (2014). Learning from their own actions: The unique effect of producing actions on infants’ action understanding. *Child Development*, *85*, 264–277. <https://doi.org/10.1111/cdev.12115>
- Ginsburg, H. P. (1989). *Children’s arithmetic- How they learn it and how you teach it*. Austin, Texas: Pro Ed.
- Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology*, *32*, 389–399. <https://doi.org/10.1016/j.cedpsych.2006.03.001>
- Goldin-Meadow, S., & Beilock, S. L. (2010). Action’s influence on thought: The case of gesture. *Perspectives on Psychological Science*, *5*, 664–674. <https://doi.org/10.1177/1745691610388764>
- Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. (2009). Gestures gives children new ideas about math. *Psychological Science*, *20*, 267–271. <https://doi.org/10.1111/j.1467-9280.2009.02297.x>
- Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*, 495–514. <https://doi.org/10.3758/ptr.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. <https://doi.org/10.1111/j.1467-7687.2009.00883.x>
- James, K. H., & Atwood, T. P. (2009). The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. *Cognitive Neuropsychology*, *26*, 91–110. <https://doi.org/10.1080/02643290802425914>
- James, K. H., & Bose, P. (2011). Self-generated actions during learning objects and sounds create sensori-motor systems in the developing brain. *Cognition, Brain, Behavior: An Interdisciplinary Journal*, *15*, 485–503.
- James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, *44*, 2937–2949. <https://doi.org/10.1016/j.neuropsychologia.2006.06.026>
- James, K. H., & Maouene, J. (2009). Auditory verb perception recruits motor developing brain: An fMRI investigation. *Developmental Psychology*, *12*, F26–34. <https://doi.org/10.1111/j.1467-7687.2009.00919.x>
- James, K. H., & Swain, S. N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Psychology*, *14*, 1–6. <https://doi.org/10.1111/j.1467-7687.2010.01011.x>
- Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, *8*, 71–78. <https://doi.org/10.1016/j.tics.2003.12.002>
- Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology*, *36*, 763–787. <https://doi.org/10.1080/87565641.2010.549884>
- Kelly, S. D., Healy, M., Ozyurek, A., & Holler, J. (2014). The processing of speech, gesture, and action during language comprehension. *Psychonomic Bulletin & Review* <https://doi.org/10.3758/s13423-014-0681-7>
- Kersey, A. J., & James, K. H. (2013). Brain activation patterns resulting from learning letter forms through active self-production and passive observation in young children. *Frontiers in Psychology*, *4*, 567. <https://doi.org/10.3389/fpsyg.2013.00567>
- Krinzinger, H., Koten, J. W., Hennemann, J., Schueppen, A., Sahr, K., Arndt, D., . . . Willmes, K. (2011). Sensitivity, reproducibility, and reliability of self-paced versus fixed stimulus presentation in an

- fMRI study on exact, non-symbolic arithmetic in typically developing children aged between 6 and 12 years. *Developmental Neuropsychology*, 36, 721–740. <https://doi.org/10.1080/87565641.2010.549882>
- LeBarton, E. S., Goldin-Meadow, S., & Raudenbush, S. (2014). Experimentally-induced increases in early gesture lead to increases in spoken vocabulary. *Journal of Cognition and Development*. <https://doi.org/10.1080/15248372.2013.858041>
- Longcamp, M., Anton, J.-L., Roth, M., & Velay, J.-L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage*, 19, 1492–1500. [https://doi.org/10.1016/s1053-8119\(03\)00088-0](https://doi.org/10.1016/s1053-8119(03)00088-0)
- Macedonia, M., Muller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, 32, 982–998. <https://doi.org/10.1002/hbm.21084>
- Mix, K. S. (2010). Spatial tools for mathematical thought. In K. S. Mix, L. B. Smith, & M. Gasser (Eds.), *Space and Language*. New York: Oxford University Press.
- Novack, M., Congdon, E., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychological Science*, 25, 903–910. <https://doi.org/10.1177/0956797613518351>
- Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, 3, 359–400. [https://doi.org/10.1016/0885-2014\(88\)90021-4](https://doi.org/10.1016/0885-2014(88)90021-4)
- Piaget, J. (1952). *The Origins of Intelligence in Children*. New York, NY: International University Press.
- Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: Children learn about balance. *Developmental Psychobiology*, 40, 1059–1067. <https://doi.org/10.1037/0012-1649.40.6.1059>
- Poldrack, R. A., Pare-Blagoev, E. J., & Grant, P. E. (2002). Pediatric functional magnetic resonance imaging: Progress and challenges. *Topics in Magnetic Resonance Imaging*, 13, 61–70. <https://doi.org/10.1097/00002142-200202000-00005>
- Prado, J., Clavagnier, S., Otzenberger, H., Scheiber, C., Kennedy, H., & Perenin, M. T. (2005). Two cortical systems for reaching in central and peripheral vision. *Neuron*, 48, 849–858. <https://doi.org/10.1016/j.neuron.2005.10.010>
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–154. <https://doi.org/10.1080/713752551>
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576–582. <https://doi.org/10.1038/nrn1706>
- Pyke, A., Betts, S., Fincham, J. M., & Anderson, J. R. (2014). Visuospatial referents facilitate the learning and transfer of mathematical operations: Extending the role of the angular gyrus. *Cognitive, Affective, & Behavioral Neuroscience* <https://doi.org/10.3758/s13415-014-0317-4>
- Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, 15, 1779–1790. <https://doi.org/10.1093/cercor/bhi055>
- Saenz-Ludow, A., & Walgamuth, C. (1998). Third Graders interpretations of equality and the equal symbol. *Educational Studies in Mathematics*, 35, 153–187.
- Smith, L. B. (2005). Action alters shape categories. *Cognitive Science*, 29, 665–679. https://doi.org/10.1207/s15516709cog0000_13
- Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96, B1–11. <https://doi.org/10.1016/j.cognition.2004.07.004>
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereo-taxic atlas of the human brain*. New York: Thieme.
- Trofatter, C., Kontra, C., Beilock, S. L., & Goldin-Meadow, S. (2015). Gesture has a larger impact on problem-solving than action, even when action is accompanied by words. *Language, Cognition and Neuroscience*. <https://doi.org/10.1080/23273798.2014.905692>
- Wakefield, E. M., Hall, C., James, K. H., & Goldin-Meadow, S. (2018). Gesture for generalization: Gesture facilitates flexible learning of words for actions on objects. *Developmental Science*.
- Wakefield, E. M., & James, K. H. (2015). Effects of learning with gesture on children's understanding of a new language concept. *Developmental Psychology*, 51, 1105–1114. <https://doi.org/10.1037/a0039471>
- Wakefield, E. M., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*
- Wiedenbauer, G., & Jansen-Osmann, P. (2008). Manual training of mental rotation in children. *Learning and Instruction*, 18, 30–41. <https://doi.org/10.1016/j.learninstruc.2006.09.009>
- Yerys, B. E., Jankowski, K. F., Shook, D., Rosenberger, L. R., Barnes, K. A., Berl, M. M., . . . Gaillard, W. D. (2009). The fMRI success rate of children and adolescents: Typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. *Human Brain Mapping*, 30, 3426–3435. <https://doi.org/10.1002/hbm.20767>
- Yuan, W., Altaye, M., Ret, J., Schmithorst, V., Byars, A. W., Plante, E., & Holland, S. K. (2009). Quantification of head motion in children during various fMRI language tasks. *Human Brain Mapping*, 30, 1481–1489. <https://doi.org/10.1002/hbm.20616>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.