

Gesture for generalization: gesture facilitates flexible learning of words for actions on objects

Elizabeth M. Wakefield^{1,2} | Casey Hall² | Karin H. James³ | Susan Goldin-Meadow²

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

²Department of Psychology, The University of Chicago, Chicago, Illinois, USA

³Department of Psychological and Brain Sciences, Indiana University, Bloomington, Indiana, USA

Correspondence

Elizabeth Wakefield, Department of Psychology, Loyola University Chicago, Chicago, IL 60660, USA.
Email: ewakefield1@luc.edu

Funding

NSF BCS (1422224/1422329)

Abstract

Verb learning is difficult for children (Gentner, 1982), partially because children have a bias to associate a novel verb not only with the action it represents, but also with the object on which it is learned (Kersten & Smith, 2002). Here we investigate how well 4- and 5-year-old children ($N = 48$) generalize novel verbs for actions on objects after doing or seeing the action (e.g., twisting a knob on an object) or after doing or seeing a gesture for the action (e.g., twisting in the air near an object). We find not only that children generalize more effectively through gesture experience, but also that this ability to generalize persists after a 24-hour delay.

1 | INTRODUCTION

We learn about the world by doing our own movements and by observing the movements of others. Two forms of movement that we engage in daily are *actions* on objects—movements of the hand that directly manipulate our environment—and *gestures*—movements of the hand that represent ideas and accompany speech. Both actions (e.g., Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005; Casile & Giese, 2006; Chao & Martin, 2000; James, 2010; James & Atwood, 2009; James & Gauthier, 2006; James & Maouene, 2009; James & Swain, 2011; Longcamp, Anton, Roth, & Velay, 2003; Longcamp, Tanskanen, & Hari, 2006; Pulvermüller,

2001) and gestures (e.g., Congdon et al., 2017; Cook, Mitchell, & Goldin-Meadow, 2008; Cook, Yip, & Goldin-Meadow, 2010; Goldin-Meadow, Cook, & Mitchell, 2009; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014; Singer & Goldin-Meadow, 2005; Valenzeno, Alibali, & Klatzky, 2003) have been shown to facilitate our ability to learn and retain new information. However, even though action and gesture both facilitate learning, they may do so in different ways. We test this hypothesis by asking whether children generalize new words differently if they are taught those words through actions versus gestures.

Gestures are movements of the hands and, in this sense, resemble the actions we use to directly manipulate our environment. But gestures differ from actions in a number of respects, and these differences have the potential to affect how we learn from action versus gesture. First, actions can have a direct effect on the world, but gestures do not—producing a *rotate* gesture does not reposition the object; only physically rotating the object has this effect. Second, the form that an action takes is determined by the object on which it is performed; gesture form is not as constrained. As a result, gestures can vary in how closely they mirror the actions they represent. For example, a *rotate* gesture produced with a C-shaped hand simulating how an object would be held if it were rotated and a *rotate* gesture produced with a pointing hand (providing no information about the object) can both

RESEARCH HIGHLIGHTS

- Children can learn the meaning of new words through action or gesture.
- But children *generalize* the meaning of words more effectively after learning through gesture than through action, suggesting that gesture and action aid learners via distinct mechanisms.
- These generalization effects occur whether actions and gestures are produced or observed by children during learning.
- These generalization effects persist after a 24-hour delay.

be used to represent the same rotating action. Gesture thus has the potential to play a unique role in learning as it, like action, is a form of movement and may therefore exploit the facilitative effects that action has on cognition (see Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008)—action establishes rich, sensori-motor representations for the information with which it co-occurs (James & Swain, 2011) and, in so doing, has the potential to improve learning (Kontra, Lyons, Fischer, & Beilock, 2015) and recall (Butler & James, 2013). But, at the same time, because gesture only *refers* to changes that can be made in the world, it can selectively highlight components of action that are relevant to a particular situation and therefore potentially go beyond the effects that action has on cognition.

Gesture's ability to highlight relevant components of action may make it an ideal teaching tool for promoting *generalization*—that is, for extending knowledge gained through an initial set of examples to novel situations. Because gesture includes some aspects of the exemplar to which it refers and omits others, it can de-emphasize aspects that are specific to the exemplar, while highlighting aspects that extend beyond the exemplar and are at the core of the to-be-learned concept. The *rotate* gesture produced with a pointing hand focuses attention on the rotation movement, while de-emphasizing the object that is being rotated and the specific hand movements necessary to grasp and act on that object. As a result, gesture may make it easier for children to transfer newly acquired knowledge beyond the initial learning context because potentially irrelevant particulars of the learning context are de-emphasized. Indeed, there is evidence from Novack and colleagues (2014) that learning through gesture can facilitate a more general understanding of a math concept than learning through action. They found that 9- and 10-year-old children learned a new mathematical concept equally well after performing actions versus gestures during the lesson. However, children were significantly more likely to generalize what they had learned—that is, to solve problems that were in a different format from the problems used during the lesson—if they had learned through gesture experience than if they had learned through action experience.

Gesture may be especially useful for instruction in domains in which generalization is particularly challenging. Here, we consider one of these domains: verb learning (Gentner, 1982; Gleitman, Cassidy, Nappa, Papfragou, & Trueswell, 2005; Gleitman & Gleitman, 1992). Children have difficulty generalizing verbs to new contexts (Imai et al., 2008; Kersten & Smith, 2002), even at age 8 (Seston, Golinkoff, Ma, & Hirsh-Pasek, 2009). Part of this difficulty may stem from children's focus on objects. An object-focus can be useful for noun learning, but it has the potential to hurt verb learning—focusing on the object associated with an action could encourage children to assume that a word intended to refer to that action performed on or by *any* object refers instead to the action performed on or by that *particular* object (see Kersten & Smith, 2002). When a child makes this inference, it can be argued that the child has not fully understood the verb. One technique that has been found to facilitate appropriate generalization in verb learning is to teach children new verbs through varied exemplars (Behrend, 1995; Childers, 2011; Childers, Heard, Ring, Pai, & Sallquist, 2012;

Forbes & Farrar, 1995), although not *too* varied (see Childers et al., 2016; Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008). Varying exemplars may be useful because it allows children to use cross-situational statistics to link words to their referents, as has been shown with action verbs (Scott & Fisher, 2012). We suggest that gesture can provide an additional avenue through which children can learn to generalize verbs.

Previous research on gesture and word learning has found that gesture can help children learn words for objects (Capone & McGregor, 2005) and words for actions that are performed on objects (de Nooijer, van Gog, Paas, & Zwaan, 2013; Wakefield, Hall, James, & Goldin-Meadow, 2017), as well as actions that are not performed on objects (Goodrich & Hudson Kam, 2009). In each of these studies, children demonstrated at test that they could pair the novel word with its referent; that is, that they had learned the word. However, they were not asked to generalize the word to new contexts. We therefore know little about the breadth of the word meaning that the children had acquired.

To our knowledge, only Mumford and Kita (2014) have explored whether gesture can help children generalize the meaning of verbs. They showed 3-year-old children a scene in which, for example, an actor placed felt pieces in a cloud shape. The experimenter then provided a novel word along with one of two gestures: a gesture highlighting the actor's manner of movement (i.e., depicting how the hand moved to make the cloud), or the end-state of the materials that the actor moved (i.e., tracing the shape of the cloud in the air); a third base-line group heard only the word with no gesture. At test, children were presented with two videos simultaneously and asked to point to the video that showed the novel verb. In one video, the actor's manner was the same but the end-state had changed (she made a different shape out of felt). In the other video, the end-state was the same but the actor's manner had changed (she used different movements to create a cloud). Children who saw manner gestures were more likely to choose the video preserving the actor's manner at test than children who saw end-state gestures or no gestures (who were equally likely to choose the video preserving end-state). These findings suggest that gestures produced in a word-learning context can influence the meaning attributed to the word, and thus how the word is generalized.

Our study extends Mumford and Kita's (2014) work in a number of ways. First, we ask whether gestures produced in a word-learning context influence whether children learn the word as a label for an action (a potential verb) or as a label for an object (a potential noun). Gesture may encourage children to attribute action meaning (rather than an object meaning) to a word simply because gesture does not involve physical manipulation of objects—this separation of movement and object could help young children overcome whatever bias they might have to focus on objects during verb learning. Performing a gesture that represents an action on an object *near* that object might then have a different effect on word learning than actually performing the

action on the object. Second, Kita and Mumford (2014) considered performance immediately after a word had been introduced. We ask whether children are able to generalize words after they have demonstrated that they learned the word in its initial instantiation; we also rule out effects of working memory by testing continued knowledge of the new words after a short delay and after a 24-hour delay. Finally, Kita and Mumford (2014) examined the impact of *seeing* another's gestures on word learning. We ask whether *doing* one's own gestures has the same impact on generalizing as seeing another's gestures. Previous work has shown that learning through doing (Cook et al., 2008; Goldin-Meadow et al., 2009) or seeing (Singer & Goldin-Meadow, 2005) a task-relevant gesture while learning the task facilitates success on that task, and gesture can help learners retain information long after its use (Cook et al., 2008; Macedonia & Klimesch, 2014). We also know, for both gesture (Goldin-Meadow et al., 2012) and action (e.g., James, Humphrey, & Goodale, 2001; James et al., 2002; James & Swain, 2011; Kontra, Goldin-Meadow, & Beilock, 2012; Kontra et al., 2015; Longcamp et al., 2006; Longcamp, Zerbato-Poudou, & Velay, 2005), that learning is more robust when movements are performed, rather than observed, and neuroimaging results suggest that, in the case of action, this robustness is due to changes in the recruitment of sensori-motor regions during the learning process and upon subsequent recall of learned information (e.g., James, 2010; Longcamp et al., 2003). However, we do not yet know whether learning through doing vs. seeing differentially impacts *generalization* of a learned concept.

Our main goal is thus to determine whether gesture is more effective than action in helping children to generalize verbs representing actions performed on objects (e.g., understanding that if performing a twisting motion on an object is called *ratching*, this label should also be applied to the same twisting motion performed on a new object). If part of the difficulty in generalizing verbs comes from focusing too much on the object used during learning, children may be better able to generalize (i.e., to apply the verb to actions performed on new objects) after learning through gesture than after learning through action. Finding evidence to support this possibility would not only shed light on the differences between how children learn through action versus gesture, but would also suggest a novel way of supporting verb learning that is not dependent on teaching through multiple unique exemplars. We also vary whether children initially learn these new words through doing or seeing actions or gestures: knowing whether there are differences in how children generalize from their own motor experiences versus watching others perform actions or gestures has implications, not only for understanding mechanism, but also for implementing educational practice.

2 | METHOD

2.1 | Participants

Forty-eight children (20 males, 28 females) between the ages of 4.5 and 5.5 years (54–66 mos.; $M = 57.4$ mos.; $SD = 5$ mos.)

participated in the study. This age range was chosen because children at this age still struggle with verb generalization (Seston et al., 2009), and are old enough to complete a 30-minute task. Participants represented a diverse sample from a large metropolitan city (50% Caucasian, 6% Mixed Race, 4% Asian, 2% American Indian, 38% non-reporting) and came from predominately high SES backgrounds, with at least one caregiver who completed a bachelor's degree in 58% of the households. Informed consent was obtained from a parent or guardian of each participant. Five additional children were excluded from analyses for failing to complete the second session of the experiment, three from the action condition and two from the gesture condition. Children completed the experimental sessions individually at their school during the regular school day.

2.2 | Materials

2.2.1 | Training stimuli

Four training objects were used as stimuli (see Figure 1a) and each child was taught a word for an action that could be performed on each of the four objects (four words in total). Objects were a solid color and approximately $12 \times 8 \times 6$ cm in size. Objects were composed of three primary shapes, making them sufficiently complex to afford at least two distinct actions that were not obvious from their appearance alone, and were based on object designs used previously by James and Swain (2011) and Wakefield et al. (2017). In videos (described below), a distinct, one-handed action was performed on each object. Each of these actions could easily be represented by a gesture (i.e., the same movement performed near, but not on, the object). Only actions that did not result in a change in end-state were used so that end-state would not differ between the two conditions. We chose simple actions so that the entire movement could naturally be described with one verb, rather than choosing a series of novel movements that might naturally be described using multiple, separate verbs. These actions or gestures were assigned one of four novel labels that followed standard verb morphology—*ratching*, *tiffing*, *leaming*, *yocking* (James & Swain, 2011). The objects were paired so that the two objects in the pair could each afford the same two actions (e.g., the orange and purple objects, pictured in Figure 1, could each afford both twisting and squeezing). However, children were taught labels for *only one* of the two actions that could be performed on an object (e.g., a child learned



FIGURE 1 Objects used during training, with trained actions and associated words. Set 1: *ratching* (twisting); *tiffing* (squeezing). Set 2: *yocking* (pulling); *leaming* (pressing). Objects within each set afford the same two actions

that *ratching* meant twisting using the orange object, and that *tiffing* meant squeezing using the purple object).

An additional four objects were created to serve as novel stimuli in a generalization assessment (e.g., the blue object in Figure 4). These objects were similar in design to the training objects in that they were composed of three primary shapes, similarly sized, and brightly colored. However, the generalization objects were distinct from the training objects in shape; there were thus no shape-matches across any of the objects used in the study. Each of the generalization objects could afford one trained action and one novel action. Novel actions were not taught during the session.

2.2.2 | Videos

Four 5-second training videos were created. In each video, one of the objects was displayed against a white background. A hand painted blue and green (children were told the hand belonged to “Arnie”, an alien) entered the frame, reached towards the object, performed an action on the object, and left the frame. The starting and ending states of the objects were identical, and each video showed a different action. Videos did not include sound.

Sixteen additional videos were created for the generalization test, two videos for each of the eight objects (four trained and four novel objects), one for each of the two actions that the object could afford. These videos were similar to the original four videos, except that the hand was not painted (Figure 4).

2.3 | Procedure

Children were randomly assigned to one of two conditions: action or gesture, and participated in two 20–30-minute videotaped experimental sessions. During the first session, children were taught four new words through either action or gesture (depending on condition). In both conditions, children learned two words through self-produced movements and two words through observing an experimenter’s movements; the order of *doing* versus *seeing* was counterbalanced, and children’s knowledge of each set of words was assessed immediately after they were trained on that set. Thus, condition (action; gesture) was a between-subjects factor, and round type (doing; seeing) was a within-subjects factor. Following training, children completed a generalization assessment. A flow-chart of Day 1 is shown in Figure 2. This procedure was used twice; once for each set of two objects. In the second session, which occurred 24 hours later, children were tested on their memory for both sets of words learned on the previous day using the same procedure for assessing generalization. The procedure for each section is outlined in more detail below.

2.3.1 | Day 1: Training

Children were told they were going to learn new words for movements that they could do with their hands, using toys belonging to “Arnie, the alien”. The experimenter explained that

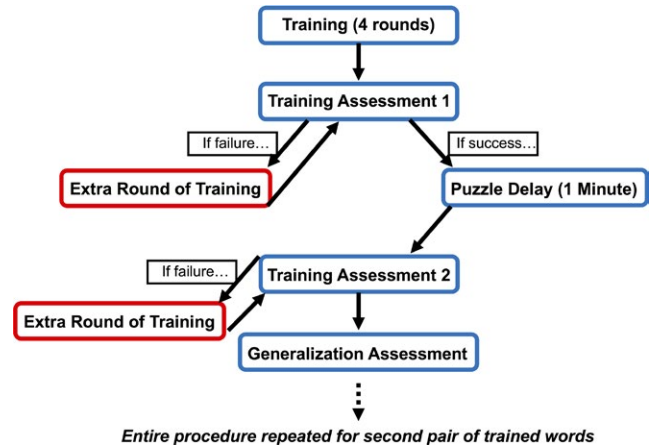


FIGURE 2 Procedure used for each pair of trained words on Day 1. Children learned one pair of words through *doing* action or gesture, and one pair of words through *seeing* action or gesture. The order of doing and seeing was counterbalanced. Twenty children needed at least one extra round during training through *doing* action or gesture, and 32 children needed at least one extra round during training through *seeing* action or gesture. Number of rounds needed was included as a factor in all analyses

they would watch one video of Arnie doing a movement, and then learn a word for what Arnie was doing. Before a new word was introduced, children watched a video of Arnie acting on an object, regardless of the experimental condition to which they had been assigned. The purpose of these videos was to demonstrate that the objects could all be acted on. We know from previous work that 4- and 5-year-old children are able to interpret gesture as a representation of action (Wakefield, Novack, & Goldin-Meadow, 2017). But, without the videos, children in the gesture condition might have interpreted the gestures on which they were trained as movements for their own sake, as opposed to movements intended to represent actions on objects. Importantly, actions were not labeled during the videos so that children’s exposure to the novel words was *only* in the gesture–word pairings (gesture condition) or the action–word pairings (action condition).

Action Training (see Figure 3a)

Children in this condition learned two words by seeing an action and two by doing an action.

Learning by seeing action After children watched a video, the experimenter performed the action shown in the video on the toy three times, and did *not* say the verb. For example, the experimenter said, “Did you see what Arnie did? Look what I can do [ACTION], I’ll do it again [ACTION] and one more time [ACTION].” The child experienced the action a total of three times. Next, the experimenter taught the word for the action, saying the word while simultaneously producing the action. For example, when teaching the word *leaming*, the experimenter said, “Arnie was leaming [ACTION]. This is called leaming [ACTION]. The best way to learn a new word is to say it out loud. Can you

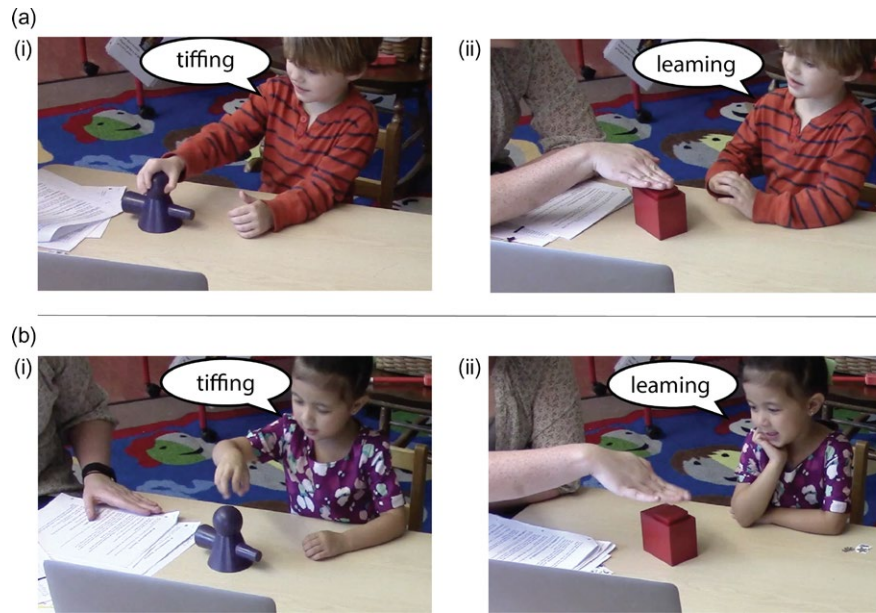


FIGURE 3 Example of word-learning procedure. (a) Action training (top row). i. Child learns the word “tiffing” through *doing an action* on an object while saying the word; ii. Child learns the word “leaming” through *seeing an action* performed on an object by the experimenter while the child says the word. (b) Gesture training (bottom row). i. Child learns the word “tiffing” through *doing a gesture* near, but not on, an object while saying the word; ii. Child learns the word “leaming” through *seeing a gesture* performed by the experimenter near an object while the child says the word

say leaming? Can you say leaming every time I do the movement?” Children were then asked to say *leaming* each time that the experimenter performed the associated action on the toy, which she did five times.

Learning by doing action After children watched a video, the experimenter demonstrated the action Arnie did on the toy once, without saying the verb, and asked children to repeat the action two times (e.g., “*Did you see what Arnie did? Look what I can do [ACTION]. Can you do that? [child produced ACTION] Can you do that one more time? [child produced ACTION].*” The child experienced the action a total of three times. Next, the experimenter taught the word for the action, saying the word while simultaneously producing the action. For example, when teaching the word *tiffing*, the experimenter said, “*Arnie was tiffing [ACTION]. This is called tiffing [ACTION]. The best way to learn a new word is to say it out loud. Can you say tiffing? Can you say tiffing every time you do the movement?*” Children were asked to say the word and simultaneously produce the action five times.

Gesture Training (see Figure 3b)

Children in this condition learned two words by seeing a gesture and two by doing a gesture. Training through seeing and doing gesture was comparable to action with one exception—when introducing the movements, the experimenter performed gestures that used the same handshape and motion trajectory as the actions shown in the videos, but the movements were produced near (and not on) the toys.

Learning by seeing gesture After children watched a video, the experimenter performed a gesture that represents the action

shown in the video near the toy shown in the video three times and did *not* say the verb. For example, the experimenter said, “*Did you see what Arnie did? Look what I can do [GESTURE], I’ll do it again [GESTURE] and one more time [GESTURE].*” The child experienced the gesture a total of three times. Next, the experimenter taught the word for the action represented by the gesture, saying the word while simultaneously producing the gesture. For example, when teaching the word *leaming*, the experimenter said, “*Arnie was leaming [GESTURE]. This is called leaming [GESTURE]. The best way to learn a new word is to say it out loud. Can you say leaming? Can you say leaming every time I do the movement?*” Children were then asked to say *leaming* each time that the experimenter performed the associated gesture on the toy, which she did five times.

Learning by doing gesture After children watched a video, the experimenter demonstrated a gesture that represents the action Arnie did on the toy once, without saying the verb, and asked children to repeat the gesture two times (e.g., “*Did you see what Arnie did? Look what I can do [GESTURE]. Can you do that? [child produced GESTURE] Can you do that one more time? [child produced GESTURE].*” The child experienced the gesture a total of three times. Next, the experimenter taught the word for the action represented by the gesture, saying the word while simultaneously producing the gesture. For example, when teaching the word *tiffing*, the experimenter said, “*Arnie was tiffing [GESTURE]. This is called tiffing [GESTURE]. The best way to learn a new word is to say it out loud. Can you say tiffing? Can you say tiffing every time you do the movement?*” Children were asked to say the word and simultaneously produce the

gesture for the action five times. Occasionally, a child learning through *doing* gesture would initially perform the action, rather than the gesture (i.e., the child would do the movement on the object). In this case, the experimenter said, “let’s just use our hands like I showed you”, and children were generally compliant.¹

Children completed four training rounds. In the second, third, and fourth training rounds in both conditions, children received a brief introduction to each word before they said the word while either seeing or doing the action or gesture:

Learning through seeing action/gesture The experimenter placed a training object in front of the child and said, “Remember, we go like this [ACTION/GESTURE] and this is called leaming [ACTION/GESTURE]? Remember, the best way to learn a new word is to say it out loud? Can you say leaming each time I do the movement? Let’s do it 5 times.” The experimenter then proceeded to perform the action or gesture while the child simultaneously said the word five times.

Learning through doing action/gesture The experimenter placed a training object in front of the child and said, “Remember, we go like this [ACTION/GESTURE] and this is called tiffing [ACTION/GESTURE]? Remember, the best way to learn a new word is to say it out loud? Can you say tiffing each time you do the movement? Let’s do it 5 times.” The child then proceeded to perform the action or gesture while the simultaneously saying the word five times.

If children did not learn the words after four rounds of training (as measured through two assessments described below), they completed subsequent rounds until they did succeed on the assessments. We controlled for the number of rounds that each child needed in order to succeed in all analyses.

2.3.2 | Day 1: Training assessments

Assessments occurred after completion of four rounds of training. Thus, assessments for the pair of words learned through *doing* action or gesture were completed immediately after four rounds of training through doing, and assessments for the pair of words learned through *seeing* action or gestures were completed immediately after four rounds of training through seeing (Training Assessment 1, see Figure 2). Children were tested on their knowledge of the words in two complementary ways to ensure that the words had been successfully learned, and had to succeed on both assessments before moving on to the generalization assessment. First, children’s recall of the words was tested by showing them videos of Arnie performing each of the four trained actions, and asking them to label the movement Arnie was performing (e.g., “What’s Arnie doing here?”). Children were considered correct only if they provided the correct label for each action. Second, children’s comprehension of the words was tested by placing the two toys in the pair on the table in front of the children in a random order, and asking them to produce each of the movements they had learned (e.g., “Can you show me ratching?”). Children were

considered correct if they performed the correct movement. Although it was possible to perform the correct movement using the toy on which it had *not* been originally trained, no child chose this option.

Children who failed the Training Assessment 1 (i.e., failed the labeling and/or production assessment) received an additional training round and were then assessed again. This procedure (alternating between an additional round and assessment round) was repeated until children succeeded on both assessments (the labeling and production assessment). Once children successfully completed the assessments, they were asked to work on a puzzle for 1 minute, and were then tested again on their ability to perform the correct actions (Training Assessment 2, see Figure 2) before moving on to the generalization test. This hiatus was used to ensure that children were not simply holding the meaning of the new words in working memory long enough only to succeed on the immediate assessments.

Most children in both the action and the gesture conditions required at least one additional round of training for at least one word (0.88 of children in the action condition; 0.96 of children in the gesture condition). The number of additional training rounds that a child experienced was not predicted by condition ($\beta = 0.85$, $SE = 0.78$, $z = 1.09$, $p = .28$). As noted earlier, the number of rounds that each child needed in order to succeed on the training assessments was included as a factor in all analyses.

2.3.3 | Day 1: Generalization assessment

After succeeding on the training assessments, children were presented with the two objects on which they had just been trained, along with two new objects, each of which could afford one of the trained actions. The experimenter then introduced the generalization test to the child, “I’m going to ask you about the words we’ve been learning by showing you more videos. You’ll see all these toys in the videos so you can take a minute to play with them.” The goal of this procedure was to avoid distractions that might result from having the child’s first exposure to the new toys occur during the generalization test.

We then tested children’s ability to generalize the words they had learned through an alternative-forced-choice task. Trials were blocked by word. Before assessing each word, the experimenter demonstrated the movement associated with the word through either action or gesture, depending on the child’s condition (e.g., “Now, I’m going to ask you about ratching; remember this [ACTION or GESTURE] is how we learned ratching.”). For each trial, children were shown two videos simultaneously in which two different actions were performed, each on a different object. The videos played on a loop, side-by-side, on a 13-inch MacBook Air via Microsoft PowerPoint, and children were asked to point to the video that exemplified the word they had learned (e.g., “Can you point to ratching?”). Children were given no feedback on whether or not their response was correct.

For each of the two words they learned, children completed 10 trials designed to test three levels of generalization: *Type 1: learned*

Generalization Trial Types (*ratching* - twisting, taught on orange object)

"See how two movements are happening? Can you point to which one is *ratching*?"

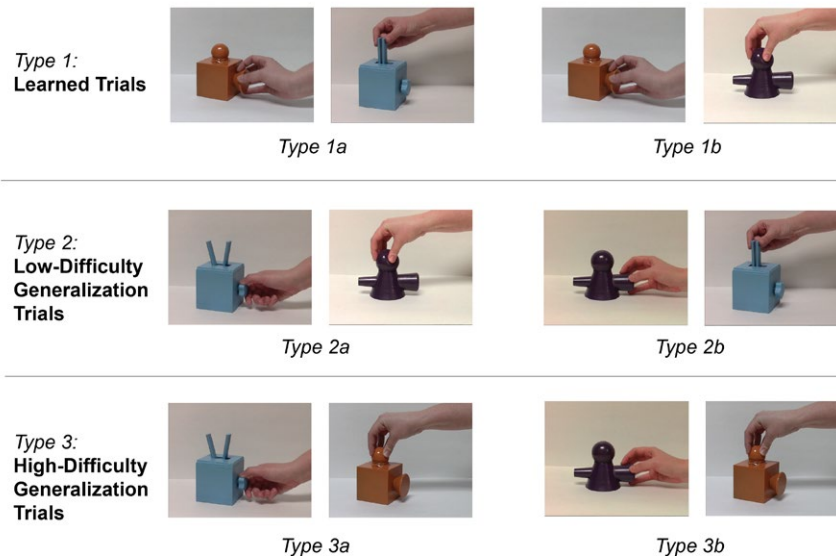


FIGURE 4 Example of generalization trials, testing for understanding of the word *ratching* (twisting). For presentation purposes, the correct choice in the above examples is always the video on the left.

(i.e., no generalization) trials, Type 2: low difficulty generalization trials, and Type 3: high difficulty generalization trials. We illustrate the different generalization types using the word *ratching* taught on the orange toy (see Figure 4).

On *learned* trials, children did not have to generalize what they had learned because one of the two videos showed the correct action on the object on which it had been taught (e.g., *ratching* performed on the orange toy on which it had been taught). In these trials, the second, distractor video either showed (a) a completely new action (one that the child had not learned a label for) performed on a new object that had not been used during training (Type 1a in Figure 4); or (b) the other action that the child had learned during that round performed on the object on which it had been taught (e.g., *tiffing* performed on the purple object on which it had been taught; Type 1b). These video pairs served as another check to ensure that children had learned the words during training and understood the task directions. Because all children were successful on the training assessments before they went on to do the generalization assessments, we predicted that the children would be correct on the learned trials in both the action and gesture conditions (e.g., we expected that, when asked "Can you point to *ratching*?" on a learned trial, children in both conditions would select the video showing the *ratching* movement performed on the orange object on which the word had been taught).

In both the *low* and *high* difficulty generalization trials, the correct action (i.e., the action associated with the word being tested) was *not* shown on the original object. In other words, in these trials, children would *not* see *ratching* performed on the orange object. Rather, *ratching* was performed on a novel blue object, which had not been used at all during training, or it was performed on the

purple object on which *tiffing* had been taught. Thus, to arrive at the correct answer, children had to have some understanding that the word could flexibly be applied to the trained action performed on various objects.

What differentiated the *low* and *high* difficulty generalization trials was whether the distractor video choice displayed the object on which the word had been taught (e.g., the orange object when testing the word *ratching*). In the *low* difficulty trials, the object on which the word had been taught was *not* present (Type 2a and Type 2b in Figure 4), but in the *high* difficulty trials, it was (Type 3a and Type 3b in Figure 4). Moreover, the object on which a word had been taught was shown with a *different* action (*tiffing* = squeezing, in the case of the orange object) being performed on it. If children considered the *object* on which the word was trained to be more important to the word's meaning than the action, they should choose the incorrect distractor video in the *high* difficulty trials (i.e., the orange object). Alternatively, if children interpreted the word as a label for the action per se, they should make the correct choice (i.e., *ratching* performed on a different toy) despite the presence of an object match. Only children who interpreted the word they had been taught as a label for an action will correctly choose the video showing the *action* associated with the word, rather than choosing the video showing the *object* originally associated with the word.

The order of the video pairs was semi-randomly determined for each tested word: (1) the same object did not appear in the same position (e.g., on either the left or right side of the screen) for two trials in a row. (2) The trial type tested (learned, low difficulty generalization, high difficulty generalization) varied for each pair so that the same trial type was never tested for two sequential trials. (3) The side on which the correct video was presented was counterbalanced.

After completing the generalization assessment for the first pair of trained words, children were allowed to choose a sticker before they began training for the second pair of words. The training and generalization assessment procedures were repeated for this second set of words, resulting in 40 video pairs tested over the entire session.

2.3.4 | Day 2: Retention assessment

At the beginning of the second session, children were given a recall retention assessment and a comprehension retention assessment. In the recall test, children were shown the videos for the words they had learned the day before, and asked to label each of the actions. In the comprehension test, objects were placed on the table in front of the child (separately for each pair) and children were asked to perform the action associated with each word (e.g., "Can you show me *ratching*?").

2.3.5 | Day 2: Generalization assessment

The same generalization assessment procedure used on Day 1 was repeated on Day 2. Children were again given a chance to play with all four toys for the first set of trained words, followed by the alternative-forced-choice task; this procedure was repeated for the second set of trained words.

3 | RESULTS

3.1 | Learning and retention

Analyses were conducted using R studio (R CoreTeam, 2016). Before considering performance on generalization, we asked whether type of training experience affected: (1) the number of rounds needed to perform successfully on the Day 1 training assessments; and (2) retention of words on the Day 2 assessments.

A general linear model with number of rounds children took to reach criterion as the dependent measure, and with round type (doing; seeing) and condition (action; gesture) as independent measures, revealed that children required significantly more rounds to learn through seeing movements ($M = 5.13$, $SD = 1.04$) than doing movements ($M = 4.60$, $SD = 0.82$; $\beta = 0.52$, $SE = 0.19$, $z = 2.71$, $p = .008$).² However, there were no significant differences in the number of rounds that children required to learn through action ($M = 5.46$, $SD = 0.94$) versus gesture ($M = 5.35$, $SD = 0.99$; $\beta = 0.15$, $SE = 0.19$, $z = 0.76$, $p = .45$).

Children were tested on their ability to remember the four learned verbs on Day 2 after a one-day delay. The original videos were shown one at a time and children were asked to label each of the movements. Recall rates were very low across participants; only three of 48 children remembered more than half of the words learned during training. Nevertheless, children were more likely to recall words that were learned through doing ($M = 0.79$ of 2, $SD = 0.71$) than through seeing ($M = 0.50$ of 2, $SD = 0.68$; $\beta = 0.68$, $SE = 0.31$, $z = 2.15$, $p = .03$). Again,

there was no effect of action vs. gesture experience on recall (action: $M = 1.25$ of 4, $SD = 0.85$; gesture: $M = 1.33$ of 4, $SD = 1.01$; $\beta = 0.10$, $SE = 0.31$, $z = 0.31$, $p = .75$). After the recall test, children were given the objects they had been taught on the day before, and asked to perform the different actions (e.g., "Can you show me *ratching*?"). Performance was much higher on this measure, but there was still no effect of condition (action: $M = 3.58$ of 4, $SD = 0.88$; gesture: $M = 3.29$ of 4, $SD = 1.12$; $\beta = 1.04$, $SE = 1.40$, $z = 0.74$, $p = .46$) and now no effect of round type (doing: $M = 1.73$ of 2, $SD = 0.68$; seeing: $M = 1.71$ of 2, $SD = 0.68$; $\beta = 0.17$, $SE = 0.58$, $z = 0.29$, $p = .77$). Together, these results suggest that seeing vs. doing a movement can, on some measures, have an impact on how well a word for that movement is learned and retained. Interestingly, however, at least in this paradigm, whether the movement is an action performed directly on an object, or a gesture performed in the air, has no discernable effect on learning or retention.

3.2 | Generalization

Our main question was whether children's ability to generalize would differ as a function of experience: doing action, seeing action, doing gesture, or seeing gesture. Figure 5 displays the average proportion correct on each of the three generalization types as a function of condition and round type for Day 1 (left graph) and Day 2 (right graph). On both days, children were more accurate on learned trials (Day 1: $M = 0.88$, $SD = 0.32$; Day 2: $M = 0.85$, $SD = 0.35$) than on low difficulty generalization trials (Day 1: $M = 0.67$, $SD = 0.46$; Day 2: $M = 0.73$, $SD = 0.44$), and were least accurate on high difficulty generalization trials (Day 1: $M = 0.49$, $SD = 0.50$; Day 2: $M = 0.63$, $SD = 0.47$). Across generalization types, children were equally accurate whether trained through doing movements (Day 1: $M = 0.68$, $SD = 0.44$; Day 2: $M = 0.74$, $SD = 0.42$) or seeing movements (Day 1: $M = 0.69$, $SD = 0.43$; Day 2: $M = 0.74$, $SD = 0.43$). Finally, and again across generalization types, children who learned through gesture (Day 1: $M = 0.72$, $SD = 0.42$; Day 2: $M = 0.79$, $SD = 0.40$) were more accurate than children who learned through action (Day 1: $M = 0.64$, $SD = 0.43$; Day 2: $M = 0.68$, $SD = 0.44$).

To statistically assess these effects, we first conducted mixed-effects binomial logistic regression models, with accuracy on each generalization trial (0, 1) entered as the outcome variable.³ We considered a simple-effects and complex model for each day. In simple-effects models for both days, condition (action; gesture), round type (doing; seeing), and generalization trial type (learned; low difficulty; high difficulty) were entered as predictors, controlling for the number of rounds needed for children to reach criterion, and entering participant as a random effect. On Day 2, the simple-effects model also controlled for children's performance on the immediate generalization test (Day 1), as well as their performance on the word recall test (Day 2). For both days, the more complex model contained all of the terms used in the simple model, as well as an interaction term between our two experience factors: condition and round type. Results from these complex models indicated that there was no significant interaction between condition and round type predicting performance on either day (Day 1: $\beta = 0.29$, $SE = 0.22$, $z = 1.32$, $p = .19$, OR = 1.34, 95% CI: 0.87,

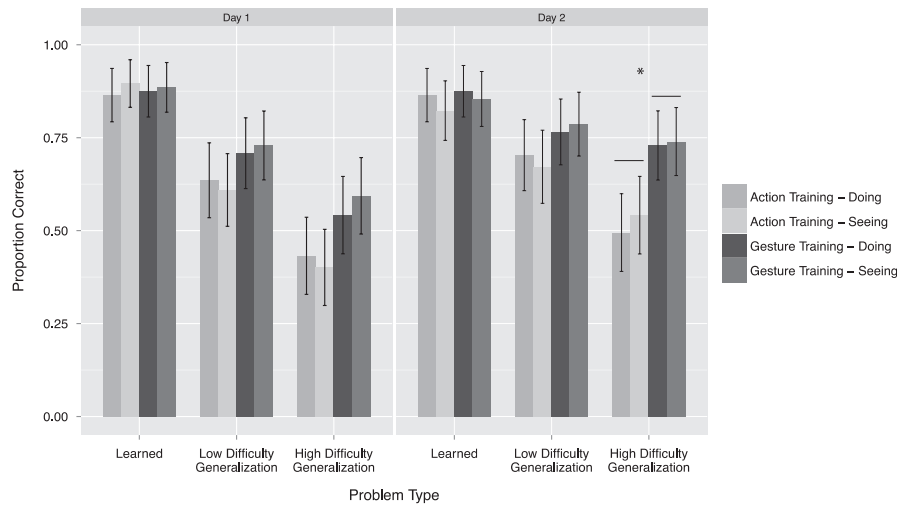


FIGURE 5 Proportion of correct responses on the three different types of generalization problems (learned, low difficulty, high difficulty) as a function of condition (action vs. gesture) and round type (doing vs. seeing). Graph on the left presents results from Day 1; graph on the right presents results from Day 2. As indicated by the asterisk, there was a significant effect ($p < .05$) of condition on Day 2 for high difficulty generalization problems. There was also a significant main effect of condition ($p < .05$) on Day 1 of condition: children learning through gesture outperformed children learning through action

2.06; Day 2: $\beta = 0.08$, $SE = 0.24$, $z = 0.33$, $p = .74$, $OR = 1.08$, 95% CI: 0.68, 1.73, and model comparisons confirmed that the complex model was not a significant improvement over the simple effects model on either day (Day 1: $\chi^2 = 1.73$, $p = .19$; Day 2: $\chi^2 = 0.11$, $p = .74$). Given these findings, we only interpret the simple-effects models.

On both days, there was no main effect of round type on generalization assessments: children performed equally well after receiving doing vs. seeing experience (Day 1: $\beta = 0.04$, $SE = 0.11$, $z = 0.39$, $p = .70$, $OR = 1.04$, 95% CI: 0.84, 1.29; Day 2: $\beta = 0.02$, $SE = 0.12$, $z = 0.18$, $p = .86$, $OR = 1.02$, 95% CI: 0.81, 1.29).

We found a more complex pattern for condition. On Day 1, children performed significantly better on generalization assessments after learning through gesture than after learning through action ($\beta = 0.73$, $SE = 0.34$, $z = 2.13$, $p = .03$, $OR = 2.08$, 95% CI: 1.07, 4.04). But, on Day 2, this effect was not significant ($\beta = 0.16$, $SE = 0.32$, $z = 4.92$, $p = .61$, $OR = 1.17$, 95% CI: 0.63, 2.20). However, training experience could have had different effects on the three types of generalization trials (learned, low difficulty, high difficulty). Indeed, we predicted that gesture would be likely to lead to more flexible understanding of a verb than action; if so, children in the gesture condition might have an advantage over children in the action condition on the more difficult generalization trial types. To test this possibility, we ran two additional models containing an interaction term between problem type and condition. Model comparisons indicated that this more complex model was *not* a better fit for Day 1 data ($\chi^2 = 2.10$, $p = .35$), but was a better fit for Day 2 data ($\chi^2 = 9.40$, $p < .01$).

As in the simple-effects model, the complex model for Day 2 revealed non-significant main effects of round type ($\beta = 0.02$, $SE = 0.12$, $z = 0.18$, $p = .86$, $OR = 1.02$, 95% CI: 0.81, 1.29) and condition ($\beta = 0.53$, $SE = 0.34$, $z = 1.60$, $p = .11$, $OR = 1.70$, 95% CI: 0.87, 3.31). However, an analysis of variance of the complex model revealed a significant

interaction between condition (action, gesture) and generalization trial type (learned, low difficulty, high difficulty), $\chi^2(2) = 9.50$, $p < .01$. Follow-up analyses showed that children in both conditions (gesture and action) performed equally well on learned trials ($\beta = 0.10$, $SE = 0.56$, $z = 0.18$, $p = .86$, $OR = 1.10$, 95% CI: 0.37, 3.31) and low difficulty generalization trials ($\beta = 0.56$, $SE = 0.42$, $z = 1.32$, $p = .19$, $OR = 1.75$, 95% CI: 0.77, 3.99). However, children in the gesture condition performed significantly better than children in the action condition on the high difficulty generalization trials ($\beta = 1.43$, $SE = 0.70$, $z = 2.05$, $p = .04$, $OR = 4.18$, 95% CI: 1.06, 16.47). We conducted additional post-hoc t tests to better understand this difference, and found that children who learned through action were at chance when answering these high difficulty generalization trials ($t(23) = 0.25$, $p = .80$), whereas children who learned through gesture performed significantly above chance when answering these trials ($t(23) = 3.57$, $p < .01$).

Recall that we framed the task for all participants in terms of *movement* rather than *objects* (e.g., we told children they would be learning words for *movements* they could do with their hands; we began each generalization trial with “see how two *movements* are happening?”). This framing was sufficient to drive children in both the action and gesture conditions away from consistently picking the object match in the generalization trials. However, children who had learned a word through action had an incomplete understanding of the meaning of the word as a label for an action. In contrast, children who had learned a word through gesture were not distracted by the object match and thus appeared to interpret the word as a label for an action that can be flexibly applied to more than one object.

3.3 | Potential effect of indexical cues

However, it is possible that the experimenter produced different types of indexical cues across the gesture and action conditions; if

so, these cues could account for the differences in generalization patterns that we found in gesture vs. action. For example, if the experimenter held the object more in the action condition than in the gesture condition and, in particular, was more likely to be holding the object while saying “*this is ratching*”, this co-occurrence could have been a subtle cue that the object was the appropriate referent for the word, not the movement. Although the experimenter was told to perform the movement as she said “*this is verb-ing*” (thus highlighting the connection between verb and movement) in all conditions, she might have introduced variability into these word–movement pairings, variability that could then have had an effect on children’s generalization performance.

To address this possibility, we first isolated all of the instances during which either the experimenter or the child mentioned the to-be-learned word. Within these instances, we coded whether the experimenter (or the child) held or stabilized the object with one hand while producing the action or gesture with the other hand.⁴ We then took the number of object handlings produced by the experimenter (or child) as a proportion of the total number of instances in which the word was mentioned. We found that the experimenter stabilized the object about half of the time when training children in the action condition ($M = 0.56$, $SD = 0.21$), significantly more often than when training children in the gesture condition ($M = 0.14$, $SD = 0.18$, $\beta = 0.42$, $SE = 0.08$, $t = 5.30$, $p < .001$). Similarly, children in the action condition stabilized the object about half of the time ($M = 0.51$, $SD = 0.30$), significantly more often than children in the gesture condition ($M = 0.11$, $SD = 0.26$), $\beta = 0.40$, $SE = 0.12$, $t = 3.50$, $p < .01$). To test whether these differences in object handling could account for our condition effects, we asked whether object handling predicted performance on the high difficulty generalization trials (the trials on which we found the biggest effects). We found that neither experimenter object handling ($\beta = 0.22$, $SE = 0.32$, $t = 0.70$, $p = .49$) nor child object handling ($\beta = 0.22$, $SE = 0.22$, $t = 0.99$, $p = .33$) predicted performance on the generalization trials.

Another indexical behavior that could have influenced children’s generalization is eye gaze. For example, looking at the object when the word is uttered could subtly indicate that “*this*” in “*this is ratching*” refers to the object. To explore this possibility, we coded whether the experimenter looked at the object, the child, or some other location as the word was uttered; we did the same coding for the child (i.e., whether the child looked at the object, the experimenter, or some other location). We then calculated the proportion of instances when the word was uttered that the experimenter (or the child) looked to each of these locations, and asked whether patterns of looking differed across conditions. We found no differences across conditions in where attention was allocated for the experimenter: looking to object vs. child (object: $M = 0.55$, $SD = 0.30$ vs. child: $M = 0.42$, $SD = 0.30$) in the action condition, compared to the gesture condition (object: $M = 0.71$, $SD = 0.16$, vs. child: $M = 0.25$, $SD = 0.16$). Nor did we find differences across conditions for the child: looking to object vs. experimenter (object: $M = 0.83$, $SD = 0.09$ vs. experimenter: $M = 0.10$, $SD = 0.08$) in the action condition, compared

to the gesture condition (object: $M = 0.76$, $SD = 0.11$ vs. experimenter: $M = 0.11$, $SD = 0.07$). Nevertheless, we asked whether looking patterns predicted performance on the generalization assessment, focusing again on high difficulty generalization trials. We found that neither the experimenter’s visual attention (looking to object: $\beta = 2.35$, $SE = 2.36$, $t = 1.00$, $p = .30$); looking to child: $\beta = 2.52$, $SE = 2.37$, $t = 1.06$, $p = .30$) nor the children’s visual attention (looking to object: $\beta = 0.48$, $SE = 0.74$, $t = 0.64$, $p = .53$; looking to experimenter: $\beta = 0.91$, $SE = 1.02$, $t = 0.29$, $p = .38$) predicted performance on the generalization trials. Thus, neither object handling nor eye gaze can account for the differences we find in generalization performance.

3.4 | Effects of extending action to multiple objects

Previous work suggests that children are more likely to generalize a verb when they are given variable input—seeing an action produced with a variety of objects improves a child’s ability to learn a word for that action (Behrend, 1995; Childers, 2011; Childers et al., 2012; Forbes & Farrar, 1995). In our study, children saw each action produced on only one object during training. However, after training, but before the generalization assessment, children were given the opportunity to play with the objects used during training and the new objects introduced for the first time in the assessment videos. The children could therefore have performed the same action on a variety of objects and, in so doing, might have given themselves the type of experience with multiple objects that could promote generalization. We asked whether this self-driven extension of action to multiple objects predicted generalization outcomes and, if so, whether it was related to condition. An initial analysis revealed no differences between the two conditions in number of objects explored (action: $M = 5.79$ objects (out of 8) $SD = 2.79$; gesture: $M = 5.83$ objects (out of 8) $SD = 2.68$; $\beta = 0.04$, $SE = 0.79$, $z = 0.05$, $p = .96$).

However, further analyses showed that the range of objects on which a child performed an action had implications for that child’s performance on both days of testing. For the action association with each prompted word, we coded the number of objects on which a child performed that action during the period of free object exploration. If children did not perform the action associated with a trained word on any object, or performed the action only on the object on which the word was taught, they received a score of 0. If they extended the action to one of the other objects that could afford the action, they received a score of 1. Finally, if they extended the action to both of the objects that could afford the action, they received a score of 2.

We asked whether children’s object exploration affected how they processed trained words immediately after training and after a 24-hour delay. To do so, we first ran a binomial logistic regression to determine whether the likelihood of getting a trial correct during the generalization assessment on Day 1 was predicted by condition (action; gesture), an interaction of problem type (learned; low difficulty; high difficulty) and extending action to other objects (using a continuous measure), controlling for participant. If extending an



action to additional objects had an effect, we assumed that the effect would be particularly strong for the low or high difficulty generalization trials, rather than the learned trials, and therefore predicted that problem type would interact with our measure of extending action to other objects. An analysis of variance of the regression model revealed a significant interaction of trial type by extending action, $\chi^2(2) = 14.81, p < .001$. Extending action *negatively* impacted children's performance on learned trials; that is, the more objects to which children extended the action, the *less* likely they were to correctly answer learned trials ($\beta = 0.62, SE = 0.27, z = 2.17, p < .05, OR = 1.86, 95\% CI: 1.10, 3.16$). We found no effect of extending action on performance on either low difficulty ($\beta = 0.03, SE = 0.14, z = 0.22, p = .83, OR = 1.03, 95\% CI: 0.78, 1.36$) or high difficulty ($\beta = 0.18, SE = 0.18, z = 1.00, p = .32, OR = 1.19, 95\% CI: 0.84, 1.70$) generalization trials.

Interestingly, extending action to multiple objects had a different effect on children's generalization performance after a 24-hour delay. An identical model was run using Day 2 generalization data, and again, an analysis of variance of the regression model revealed a significant interaction of trial type by extending action ($\chi^2(2) = 18.26, p < .001$): An increase in number of extensions had no effect on performance on learned trials ($\beta = 0.45, SE = 0.28, z = 1.60, p = .11, OR = 1.56, 95\% CI: 0.91, 2.71$) or low difficulty generalization trials ($\beta = 0.07, SE = 0.16, z = 0.43, p = .67, OR = 1.07, 95\% CI: 0.78, 1.47$), but *did* positively predict performance on high difficulty generalization trials ($\beta = 0.58, SE = 0.21, z = 2.77, p < .01, OR = 1.78, 95\% CI: 1.18, 2.70$). Thus, extending an action associated with a specific object may initially destabilize children's knowledge of the word so that it becomes *less* linked to the original object (thereby negatively affecting performance on the learned trials); but this destabilization might lead to better generalization after a delay. Importantly, there was still an overall effect of condition over and above the finding that action extension predicted correct answers on high difficulty generalization trials after a delay—children who learned through gesture outperformed children who learned through action ($\beta = 1.42, SE = 0.71, z = 2.00, p < .05, OR = 4.14, 95\% CI: 1.03, 16.63$), suggesting that gesture can play a unique role in helping children generalize, above and beyond having multiple examples from which to learn.

3.5 | What were children *really* learning during training?

Our study was designed to test children's ability to learn and generalize novel verbs. We introduced the idea that the to-be-learned words referred to movements on objects at the very beginning of the procedure, and we reinforced this idea throughout the study. For example, the prompt for the generalization test items always started with "See how two movements are happening?" and the novel verbs were presented with the *-ing* ending, which typically signals a verb in English. Nevertheless, it is possible that children thought the words they were learning referred to specific objects; if so, they ought not have generalized the words to new actions, but rather restricted

them to the object on which the action was originally performed. However, the retention data suggest that children did learn the words as labels for actions. When asked to show the experimenter one of the verbs during the retention test on Day 2 (e.g., show me *ratching*), out of 192 responses elicited from children, only *one* response was in the form of a point towards an object.

4 | DISCUSSION

Our goal was to investigate the impact that acting out the meaning of a verb, compared to gesturing about its meaning, has on how the verb is learned and generalized. With respect to learning, we found that children needed the same number of rounds to learn a verb through action as they needed to learn the verb through gesture (see Wakefield et al., 2017, for a more thorough study of initial verb learning through action and gesture). But we found a different pattern with respect to generalization. Children who learned a verb through gesture were *better* able to extend that verb to novel objects than children who learned the verb through action. We also investigated the impact that producing one's own movements (actions or gestures) has on learning and generalization, compared to observing someone else's movements. We found that, for both action and gesture, children needed fewer rounds to learn a verb, and later recalled more of the verbs, after producing the movements themselves than after observing others produce the movements. However, the pattern again differed for generalization. Producing movements (either actions or gestures) was no more, and no less, effective in promoting generalization than observing others produce the movements. The next sections consider these findings in relation to previous work, first, on learning through producing vs. observing movement and, then, on learning through gesture vs. action.

4.1 | Effect of producing movement vs. seeing movement on learning

Previous work suggests that children and adults can learn and retain more from producing their own movements than from watching another produce the same movements, both when learning through action (e.g., Butler & James, 2013) and when learning through gesture (Goldin-Meadow et al., 2012), although there are situations in which learning through self-produced movement produces similar benefits for learners as learning through observed movement (e.g., Feyereisen, 2009). Our studies add another context to the literature in which producing one's own movements provides greater benefits for learners than observing another's movements.

However, unlike previous work, we examined the effects of learning through action and gesture *within the same study*, allowing us not only to replicate, but also build upon, previous literature. Although null results must be interpreted with caution, it is interesting that when predicting how well children recalled words after a 24-hour delay, we did not find an interaction between the type of movement used during training (action, gesture) and whether the movement was produced or observed. Thus, the mechanism underlying children's ability to

learn and retain information through producing rather than observing movement may be the same, whether the movement is action or gesture. This finding suggests that the power of learning through movement does not rely solely on tactile experience with objects in the world. Rather, the effectiveness of self-produced movement may lie in engaging the body in the learning experience more generally. This finding complements results supporting the idea that engaging the motor system in the learning process has a significant effect on how information is stored and processed (e.g., James et al., 2001, 2002; James & Swain, 2011; Kontra et al., 2012, 2015; Longcamp et al., 2005, 2006), although more work is needed to adequately characterize the similarities and differences in neural changes that accompany learning through self-produced versus observed actions and gesture.

4.2 | Effect of action vs. gesture experience on generalization

Our results also indicate that learning through action, as opposed to gesture, can lead to less flexibility in extending what has been learned. On both days of testing, we found that children who received gesture training were better able to generalize what they had learned than children who received action training. Interestingly, the previous literature has suggested that learning a word with respect to only one exemplar can restrict a child's ability to generalize the word; exposing children to multiple exemplars, particularly if they are perceptually similar to each other, can help support generalization in word learning (Childers et al., 2012, 2016; Haryu, Imai, & Okada, 2011). Our findings indicate that gesture can also promote generalization in verb learning, using a route that does *not* involve multiple exemplars. Notably, we found that gesturing during verb learning helps children generalize, above and beyond experiencing multiple exemplars: Gesture continued to predict generalization, even when children's self-directed object exploration (which also predicted generalization) was added to the model. These results suggest that gesture can be a viable method for teaching new verbs. Moreover, learning through gesture can encourage generalizing a newly learned verb to a new context more robustly than does learning through action (see also Novack et al., 2014). We suggest that gesture supports generalization in verb learning because gestures are not performed on objects and, as a result, separate the movement component of an action from the particular object on which it is performed. This advantage becomes particularly apparent if children are given a delayed test, perhaps because they have more time to consolidate what they learned from the lesson. Gesture may also introduce more variability into the learning environment than action does. We consider these two possibilities in the next sections.

4.2.1 | Gesture separates movement from object, highlighting important aspects of action in verb learning

Paying attention to object properties is a crucial part of children's early learning experiences. From a young age, children have a shape

bias and thus use the shapes of objects to determine the category those objects fall within. Focusing on object properties, such as shape, is useful for noun learning. However, verb learning involves taking a step back from a particular object and focusing on the movement properties of an action independent of the object on which it is performed—a difficult process for young children (e.g., Behrend, 1995; Forbes & Farrar, 1995; Kersten & Smith, 2002). For example, Kersten and Smith (2002) found that children treated the object and motion involved in a verb-learning paradigm as equally important factors to consider when extending verbs. We hypothesized that gesture has the potential to help children overcome this bias simply because gesture separates the object being acted on from the motion. In verb learning, gestures can represent the action associated with a verb without directly tying that motion to an object. Our data support this hypothesis not only because gesture is beneficial overall, but also because the largest difference between children's performance after action vs. gesture occurred in the most difficult type of generalization trials, trials in which the *incorrect* choice is a video showing the object on which the prompted verb was initially taught, but with a different action. By perceptually separating the motion associated with the verb from the object used during training, gesture can help children attend to the motion and learn a broader meaning of the verb, rather than narrowly associating the verb with the object. We recognize that there are specialized verbs that should be applied more narrowly to a specific object-verb pairing (e.g., "sweep", "gallop", although even these verbs can be extended metaphorically to other objects). However, given children's difficulty in learning to separate the meaning of a verb from its associated object, we believe that there could be educational benefits to considering gesture as a learning tool that can help children generalize their knowledge of verbs.

4.2.2 | Gesture may add variability to the learning environment

A large literature suggests that some degree of variability in input is useful in helping children create abstract categories. For example, in early noun learning, showing a child multiple instances of a dog facilitates the acquisition of a "dog" schema, allowing the child to extend the label appropriately to any number of types of dogs (e.g., Balaban & Waxman, 1997; Waxman & Klibanoff, 2000). Other work shows that children naturally produce variability when learning to print letters of the alphabet, and that this variability supports their ability to develop letter categories (Li & James, 2016). In accounts of verb learning, researchers have made similar claims, suggesting that seeing the same action performed on different objects can help children recognize the core feature that is similar across exemplars and that allows generalization (Gentner, 2003; but see Maguire et al., 2008, who find that varying the person who performs an action can be detrimental to learning to generalize a verb for that action).

Although we did not calculate how much variability there was in the action vs. gesture forms the children produced during our study, it is likely that the children's gestures were more variable than their actions. Gesture is performed without an object in hand

and is thus not subject to the constraints that dictate action (e.g., affordances of the object, which determine an action's grasp size, force, etc.). The gestures the children produced and saw are thus likely to vary more than the actions they produced and saw. For example, the range of handshapes and amount of force that a child used during his five repetitions of the gesture for *ratching* is likely to be greater than the range used during his five repetitions of the action for *ratching*. If variable input is useful for verb learning, gesture may be providing more variability from which to learn than action. If so, the mechanism that underlies gesture's ability to support generalization may be similar to the mechanism by which children learn to generalize after seeing the same action produced with multiple exemplars. Whether gesture is actually more variable than action could be further explored by analyzing the movement properties of gesture and comparing them to action.

4.3 | Conclusions

Our study demonstrates that gesture instruction goes beyond action instruction to promote generalization in word learning—it increases the likelihood that a verb referring to an action will be extended beyond the original context in which the action was performed. Action engages learners in a specific context. Gesture, by contrast, provides a distancing from objects, as well as a variability of form, that may encourage a flexible interpretation of a word's meaning. Gesture and action may thus aid learners through overlapping, yet distinct, mechanisms.

ACKNOWLEDGEMENTS

This work was supported by a collaborative grant from NSF BCS (1422224/1422329) to Goldin-Meadow and James. The authors would like to thank Emily Kubota for her help with data collection, as well as the children and families who participated and the schools at which the study was completed.

ENDNOTES

¹ Of the 24 children in the gesture condition, six children never had to be corrected, seven children needed to be corrected on the first word during the first round of training, and 11 children needed to be corrected on both of the words during the first round of training. After these reminders, children produced gestures rather than actions. The number of times a child had to be reminded did not predict performance ($\chi^2(2) = 0.38, p = .69$).

² Although we controlled the number of times the child said the word so that it was the same in all conditions, it is possible that the timing between speech and movement differed when seeing vs. doing the movement. This difference might then be responsible for the fact that learning required fewer trials in doing vs. seeing trials. To examine this possibility, we coded the initial four rounds of training in a subset of videos for the number of times children said the trained word at the same time as they produced the movement in doing rounds, and compared it to the number of times they produced the word at the same time as they observed the movement in seeing rounds. We found that word and movement co-occurred most of the time in both the gesture condition (seeing: $M = 5.08; SD = 0.35$; doing: $M = 4.96; SD = 0.79$) and the action condition (seeing: $M = 5.01; SD$

$= 0.10$; doing: $M = 5.08; SD = 0.28$). There were no statistically significant differences in number of word–movement co-occurrences as a function of round type (seeing vs. doing, $\beta = 0.03, SE = 0.05, t = 0.44, p = .66$) or condition (gesture vs. action condition, $\beta = 0.03, SE = 0.06, t = 0.43, p = .68$).

³ Preliminary analyses revealed that there were no effects of gender ($\beta = 0.04, SE = 0.33, z = 0.12, p = .91$), age ($\beta = 0.01, SE = 0.03, z = 0.28, p = .78$), order of rounds ($\beta = 0.16, SE = 0.33, z = 0.49, p = .63$), or receptive vocabulary ($\beta = 0.01, SE = 0.01, z = 0.65, p = .52$) on overall performance. These factors were therefore not included in the analyses.

⁴ The camera in our experiment was set up to capture whether the child was producing gestures or actions, not necessarily whether the experimenter or child stabilized or looked at the object during these behaviors. Nevertheless, we were able to code these properties on half of the videos (12 of 24 participants in the gesture condition, 12 of 24 in the action condition).

REFERENCES

- Balaban, M.T., & Waxman, S. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64*, 3–26.
- Behrend, D.A. (1995). Processes involved in the initial mapping of verb meanings. In M. Tomasello & W.E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 251–273). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Beilock, S.L., Lyons, I.M., Mattarella-Micke, A., Nusbaum, H., & Small, S.L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences, 105*, 13269–13273.
- 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.
- Calvo-Merino, B., Glaser, D.E., Grezes, J., Passingham, R.E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249.
- Capone, N.C., & McGregor, K.K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research, 48*, 1468–1480.
- Casile, A., & Giese, M.A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology, 16*, 69–74.
- Chao, L.L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage, 12*, 478–484.
- Childers, J.B. (2011). Attention to multiple events helps 2 1/2-year-olds extend new verbs. *First Language, 31*, 3–22.
- Childers, J.B., Heard, M.E., Ring, K., Pai, A., & Sallquist, J. (2012). Children use different cues to guide noun and verb extensions. *Language Learning and Development, 8*, 233–254.
- Childers, J.B., Parrish, R., Olson, C.V., Burch, C., Fung, G., & McIntyre, K. (2016). Early verb learning: How do children learn how to compare events? *Journal of Cognition and Development, 17*, 41–66.
- 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.
- Cook, S.W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*, 1047–1058.
- Cook, S.W., Yip, T.K., & Goldin-Meadow, S. (2010). Gesturing makes memories that last. *Journal of Memory and Language, 63*, 465–475.
- de Nooijer, J.A., van Gog, T., Paas, F., & Zwaan, R.A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychologica, 144*, 173–179.
- Feyereisen, P. (2009). Enactment effects and integration processes in younger and older adults' memory for actions. *Memory, 17*, 374–385.

- Forbes, J.N., & Farrar, M.J. (1995). Learning to represent word meaning: What initial training events reveal about children's developing action verb concepts. *Cognitive Development*, 10, 1–20.
- Gentner, D. (1982). Why are nouns learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj (Ed.), *Language development: Language, thought and culture* (Vol. 2, pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.
- Gleitman, L.R., Cassidy, K., Nappa, R., Papfragou, A., & Trueswell, J.C. (2005). Hard words. *Language Learning and Development*, 1, 23–64.
- Gleitman, L.R., & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science*, 1, 31–35.
- Goldin-Meadow, S., Cook, S.W., & Mitchell, Z. (2009). Gestures gives children new ideas about math. *Psychological Science*, 20, 267–271.
- Goldin-Meadow, S., Levine, S.C., Zinchenko, E., Yip, T.K., Hemani, N., & Factor, L. (2012). Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science*, 15, 876–884.
- Goodrich, W., & Hudson Kam, C.L. (2009). Co-speech gesture as input in verb learning. *Developmental Science*, 12, 81–87.
- Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development*, 82, 674–686.
- Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., & Golinkoff, R.M. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, 79, 979–1000.
- James, K.H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science*, 13, 279–288.
- 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.
- James, K.H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, 44, 2937–2949.
- James, K.H., Humphrey, G.K., & Goodale, M.A. (2001). Manipulating and recognizing virtual objects: Where the action is. *Canadian Journal of Experimental Psychology*, 55, 111–120.
- James, K.H., Humphrey, G.K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M.A. (2002). "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers*, 34, 383–390.
- James, K.H., & Mawouene, J. (2009). Auditory verb perception recruits motor developing brain: An fMRI investigation. *Developmental Psychology*, 12, F26–F34.
- James, K.H., & Swain, S.N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Psychology*, 14, 1–6.
- Kersten, A.W., & Smith, L.B. (2002). Attention to novel objects during verb learning. *Child Development*, 73, 93–109.
- Kontra, C., Goldin-Meadow, S., & Beilock, S.L. (2012). Embodied learning across the life span. *Topics in Cognitive Science*, 4, 731–739.
- Kontra, C., Lyons, D.J., Fischer, C., & Beilock, S.L. (2015). Physical experience enhances science learning. *Psychological Science*, 26, 737–749.
- Li, J.X., & James, K.H. (2016). Handwriting generates variable visual output to facilitate symbol learning. *Journal of Experimental Psychology: General*, 145, 298–313.
- 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.
- Longcamp, M., Tanskanen, T., & Hari, R. (2006). The imprint of action: Motor cortex involvement in visual perception of handwritten letters. *NeuroImage*, 33, 681–688.
- Longcamp, M., Zerbato-Poudou, M.T., & Velay, J.L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, 119, 67–79.
- Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education*, 8, 74–88.
- Maguire, M.J., Hirsh-Pasek, K., Golinkoff, R.M., & Brandone, A.C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, 11, 628–634.
- Mumford, K.H., & Kita, S. (2014). Children use gesture to interpret novel verb meanings. *Child Development*, 85, 1181–1189.
- 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.
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, 5, 517–524.
- R CoreTeam (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Scott, R.M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, 122, 136–180.
- Seston, R., Golinkoff, R.M., Ma, W., & Hirsh-Pasek, K. (2009). Vacuuming with my mouth? Children's ability to comprehend novel extensions of familiar verbs. *Cognitive Development*, 24, 113–124.
- Singer, M.A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, 16, 85–89.
- Valenzeno, L., Alibali, M.W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28, 187–204.
- Wakefield, E.M., Hall, C., James, K.H., & Goldin-Meadow, S. (2017). Representational gesture as a tool for promoting word learning in young children. In *Proceedings of the 41st Annual Boston University Conference on Language Development*, Boston, MA.
- Wakefield, E.M., Novack, M., & Goldin-Meadow, S. (2017). Unpacking the ontogeny of gesture understanding: How movement becomes meaningful across development. *Child Development*, <https://doi.org/10.1111/cdev.12817>
- Waxman, S., & Klibanoff, R.S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, 36, 571–581.

How to cite this article: Wakefield EM, Hall C, James KH, Goldin-Meadow S. Gesture for generalization: gesture facilitates flexible learning of words for actions on objects. *Dev Sci*. 2018;e12656. <https://doi.org/10.1111/desc.12656>