Representational Gesture as a Tool for Promoting Verb Learning in Young Children

Elizabeth M. Wakefield, Casey Hall, Karin H. James, and Susan Goldin-Meadow

1. Introduction

The movements we produce or observe others produce can help us learn. Two forms of movement that are commonplace in our daily lives are actions, hand movements that directly manipulate our environment, and gestures, hand movements that accompany speech and represent ideas but do not lead to physical changes in the environment. Both action and gesture have been found to influence cognition, facilitating our ability to learn and remember new information (e.g., Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005; Casile & Giese, 2006; Chao & Martin, 2000; Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Cook, & Mitchell, 2009; Goldin-Meadow et al., 2012; 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; Wakefield & James, 2015). However, the two types of movement may affect learning in different ways. In previous work, the effects of action and gesture on learning have been considered separately (but see Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014). Our goal here is to directly compare children’s ability to learn from actions on objects versus gestures off objects. We consider this question in the realm of word learning, specifically, teaching children verbs for actions that are performed on objects. We also ask whether learning through these movements unfolds differently when movements are produced versus observed by a child. More broadly, our study is a first step in understanding how information is learned, generalized, and retained based on whether it is expressed through action or gesture.

*Corresponding author: Elizabeth Wakefield, Loyola University Chicago and University of Chicago, ewakefield1@luc.edu. Casey Hall and Susan Goldin-Meadow are affiliated with the University of Chicago. Karin James is affiliated with Indiana University. The authors would like to thank all of the children and families who participated, as well as Jessica Jankowski and Emily Kubota for their help with data collection. Finally, this work was made possible by a collaborative grant from the National Science Foundation (142224/1422329) to S. Goldin-Meadow and K. H. James, which also supported E. Wakefield and C. Hall.

The actions we do in relation to a task have been shown to affect how we process and remember that task. Furthermore, the ability to learn from action varies depending on the perspective of the learner. For example, people are more likely to recall an action if they have performed the action than if they have read a verbal description of the action (Engelkamp & Zimmer, 1989; Nilsson, 2000). Previous work has also shown that learning through doing action leads to quicker learning and better short-term retention than seeing the same action (Butler & James, 2013). Neuroimaging investigations have confirmed that information we learn through doing our own actions is processed differently from information learned through seeing others’ actions (James, 2010; James & Engelhardt, 2012; James & Maouene, 2009; James & Swain, 2011). For example, James and Swain (2011) asked children to learn words for actions on objects (e.g., twisting a piece of an object) either through doing the action on the object, or through seeing an experimenter perform the action on the object. In both cases, the actions were labeled with a novel word, which conformed to standard English verb morphology, as they were performed (e.g., ‘ratching’ said during the twisting). Children then participated in a functional Magnetic Resonance Imaging (fMRI) session during which they listened to the learned words and no actions were performed. Analyses revealed significantly stronger activation in frontal and parietal motor regions – regions associated with planning motor actions and the act of grasping – when children listened to words that they learned through doing actions compared to words learned through seeing actions. The movement experiences we have while learning a new word can thus affect how we process that word after learning.

Like action, doing (Cook et al., 2008; Goldin-Meadow et al., 2009) or seeing (Singer & Goldin-Meadow, 2005) a task-relevant gesture while learning a task facilitates success on that task. There is also some evidence that learning through doing gesture is more powerful than learning through seeing gesture (Goldin-Meadow et al., 2012). However, learning through gesture is rarely directly compared to learning through action. We therefore do not yet know whether the effect of seeing versus doing is equally strong in action and gesture.

Gestures are movements of the hand and thus are themselves actions. However, gesture and action are distinct in many respects. Critically, gestures do not have a direct effect on the world – producing a rotate gesture while mentally rotating an object does not actually involve touching and repositioning the object, as does physically rotating the object. As a result, there may be a greater disparity between doing versus seeing action than between doing versus seeing gesture. Doing either an action or gesture involves planning and executing a motor movement, whereas seeing either an action or gesture does not. However, through action, a learner also experiences tactile feedback when he performs an action on an object – an experience that is lacking when he learns through seeing someone else act. In contrast, neither doing nor seeing gesture involves this type of tactile feedback. Thus, if part of what drives the difference between doing and seeing action is the presence or absence of tactile feedback, we may see a greater disparity in children’s ability to learn through doing versus seeing
actions on objects, compared to doing versus seeing gestures off objects. On the other hand, if tactile feedback is not relevant to the learning differences found between doing and seeing, we would expect a main effect of doing versus seeing that does not interact with condition (action versus gesture).

Gestures can resemble actions, but they vary in how closely they mirror the actions they represent (e.g., a rotate gesture produced with a C-shaped hand simulating how the object would be held if it were rotated resembles the actual act of rotating more closely than a rotate gesture produced with a pointing hand). This is a second difference between action and gesture that is potentially relevant to learning. Gesture has the potential to play a unique role in learning as it is itself an action and can exploit the effects that action has on cognition (c.f., Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008). But 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. This selectivity may be useful in learning verbs for actions on objects. Generalizing verbs is notoriously difficult for children (Imai et al., 2008; Kersten & Smith, 2002; Seston, Golinkoff, Ma, & Hirsh-Pasek, 2009); children often think the object on which an action is demonstrated is part of verb meaning. Gesture could help children focus on the movement per se and infer that it is the referent of the novel verb. In contrast, action may lead learners to focus on the object along with the movement and thus infer that both are the verb’s referent. However, because gesture is removed from the object, and represents, rather than depicts, it may take longer to initially learn from gesture than from action.

Previous research has found that gesture can help children learn new words (e.g., Capone & McGregor, 2005; Goodrich & Hudson Kam, 2009), but understanding how this learning process unfolds through gesture vs. action is an important step in understanding gesture’s potential not only for learning new verbs representing actions on objects, but also for generalizing those meanings.

Our goal is thus to determine whether gesture helps children learn verbs for actions that are performed on objects (concrete actions) as effectively as action and, if so, whether this effect is conditioned by doing the gesture or action versus seeing an experimenter do the gesture or action. We hypothesize that children will be able to learn through both types of movement, but that the rate at which they learn may differ as a function of type of movement. We also hypothesize that children will learn better from doing movements than from seeing movements, but that there may be a greater disparity between doing and seeing when children learn from action than from gesture.

2. Method
2.1. Participants

Forty-eight children (24 males, 24 females) between the ages of 4.5 and 5.5 years (54 – 66 mo; $M = 58$ mo; $SD = 3.6$ mo) participated in the study. Participants represented a diverse sample from a large metropolitan city (50% Caucasian, 21% African American, 17% Mixed Race, 4% Asian, 2% American.
Indian, 6% Non-reporting) and came from predominately high SES backgrounds, with at least one caregiver who completed a bachelor’s degree in 79% of the households. Informed consent was obtained from a parent or guardian of each participant, and children received stickers during the study as a form of compensation. Seven additional children were excluded from analyses due to failure to complete the experiment or non-compliance during the session ($n = 6$) or inadequate knowledge of English ($n = 1$).

2.2. Materials

**Training Stimuli.** Eight novel objects were created for the study (see examples in Figure 1), modeled after those used by James and Swain (2011). Objects were brightly colored and approximately 12 x 8 x 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. In training videos (described below), one distinct one-handed action was performed on each object; each action could easily be represented by a distinct gesture (the same movement performed near, but not on, the object). These actions or gestures were given novel labels that followed standard English verb morphology — yocking, wilping, tiffing, sprocking, ratching, panking, nooping, leaming (James & Swain, 2011).

![Figure 1. Examples of training objects.](image)

**Training Videos.** Eight 5-second training videos were created. In each video, one of the novel 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 displayed a different action. Videos were soundless.

2.3. Procedure

Children were randomly assigned to one of two conditions: Action Training or Gesture Training. Children in each condition learned four novel words for actions that could be performed on objects through **doing** actions or gestures, and four words for novel actions that could be performed on objects through **seeing** actions or gestures. Training occurred over four rounds, and each round was followed by an assessment of children’s knowledge of the novel words.
Training was blocked by round type (seeing vs. doing), and counterbalanced, such that half of the children participated in four rounds of training through seeing movements (actions or gestures), followed by four rounds of doing movements (actions or gestures), and the other half of the children completed the training through four rounds of doing movements followed by training through four rounds of seeing movements. The procedure took 20-30 minutes.

**Training.** Children were told they were going to learn new words for movements they could do with their hands, using toys belonging to Arnie the alien. The experimenter explained that they would watch videos of Arnie doing the movements, and then learn words for what Arnie was doing. Regardless of condition, before a new word was introduced, children watched a training video of Arnie acting on an object. The purpose of these videos was to demonstrate that the objects could all be acted on; we assumed that children in the gesture condition would understand the movements with which they were trained as representations of actions (we know that adults and children are able to interpret gestures in this way, see Novack, Wakefield, & Goldin-Meadow, 2016; Wakefield, Novack, & Goldin-Meadow, in press). However, without the videos, children in the gesture condition might have interpreted the movements they saw or did as movements for their own sake, as opposed to movements intended to represent actions on objects.

Importantly, actions were not labeled during training videos so that children’s exposure to novel words was only in the gesture-word pairings (Gesture Training) or the action-word pairings (Action Training).

**Action Training.** Learning by seeing action. After children watched a training video, the experimenter performed the same action on the toy three times and then taught the word for the action. For example, when teaching the word ‘ratching’ the experimenter said, “Arnie was ratching [ACTION]. This is called ratching [ACTION]. The best way to learn a new word is to say it out loud. Can you say ratching?” Children were then asked to say ratching each time that the experimenter performed the associated action on the toy, which she did 5 times. Learning by doing action. After children watched a training video, the experimenter demonstrated the action Arnie did on the toy once, asked children to repeat the action two times, and then taught the word for the action. Children were asked to say the word and simultaneously produce the action five times.

**Gesture Training.** Training through seeing and doing gesture was comparable to action training, except that when introducing the movements, the experimenter performed gestures that used the same handshape and motion trajectory as the actions shown in the training videos; these gestures were produced near (but not on) the toys. 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

For the second, third, and fourth training rounds in both conditions, children only received a brief introduction to each word before they said the word while either seeing or doing the action or gesture (e.g., “Remember, this is called ratching [ACTION/GESTURE]. Can you say ratching and do your movement/while I do the movement five times?”).

**Assessment.** After each training round, the experimenter placed the four objects that were used during the round on the table in front of the child. All four objects were placed on the table at the same time, and the placement order of the toys was random. She then tested for understanding of each of the four novel words by prompting the child to perform the action associated with each word (e.g., “Can you show me ratching?”), a method that has been used in previous research (e.g., Childers, 2011).

### 3. Results

#### 3.1. Coding

Children’s responses were coded for correctness. Children were considered correct on a word if they performed the action that had been associated with that word during training. Although children would have been considered correct if they performed the correct action on any of the objects, children only produced correct actions on the object associated with each of the prompted words during training. Because the goal was to teach (using either action or gesture training) children actions that could be done on objects, gesture responses were not considered correct.

#### 3.2. Types of Movement Produced During Assessment

Before considering the effects of various factors on children’s performance, we first consider the types of movement children made in response to the prompt (e.g., “Can you show me ratching?”). Because we only considered children to be correct if they performed a correct action (i.e., gesture responses were not considered correct), it was necessary to establish that rates of action were the same across children in the Action Training and Gesture Training conditions. Otherwise, any differences that we found between Conditions could be based on children in the Gesture Training condition performing gestures rather than actions (recall that all children saw actions performed on each of the objects with which they were trained in the initial videos shown).

To evaluate whether children across conditions were equally likely to perform actions in response to the prompt, we analyzed the total number of

---

¹ The 2 children who were not compliant with instructions not to act on objects during gesture training were excluded from the study.
actions produced during assessment as a proportion of all responses given during assessment. Overall, children responded during assessment prompts with actions, gestures, points, or a verbal indication that they did not remember the relevant action. Children who received Action Training predominately produced actions (0.82), and occasionally produced points (0.18) or did not respond (<0.01). Children who received Gesture Training also predominately produced actions (0.73), and also produced points (0.22), action gestures (0.05) or did not respond (0.01). A between-subjects t-test revealed no significant difference between the proportion of actions produced during assessment by children in the Action Training and Gesture Training groups, t(46) = 1.00, ns. A paired samples t-test also showed that there was no effect of round type (seeing vs. doing), t(47) = 0.84, ns. Taken together, these findings suggest that children were equally likely to interpret the prompt ‘Can you show me ratching?’ as a request for them to perform an action on an object, regardless of whether they had been doing or seeing actions or gestures during training. This finding provides evidence that children in the Gesture Training group were interpreting the gestures as representations of actions performed on objects.

3.3. Effect of Training Condition and Round Type

Our main question was whether there would be differences in how well children learned through different types of movement training. Figure 2 displays the average proportion of problems children correctly answered on the first and final rounds of assessment for each training condition and for the doing and seeing rounds. Overall, children improved across training rounds, although it is clear from the graph that the words had not been fully learned by the end of the four rounds. For the purpose of analysis, we considered performance on individual problems: we used a binomial logistic regression model to test whether correct responses were predicted by Training Condition (Action Training, Gesture Training) and Round Type (Doing, Seeing). We also included Training Round (1 through 4) as a factor to determine whether children improved across the four training rounds, and controlled for Age and participant. We found that children performed significantly better after doing than after seeing movements (β = 0.30, SE = 0.12, z = 2.39, p < .05), and that children in the Action Training condition performed significantly better than children in the Gesture Training condition (β = 0.97, SE = 0.47, z = 2.04, p < .05). Importantly, a significant effect of training round indicated that all children improved across the training sessions, showing that learning did occur in each of the training conditions (β = 0.48, SE = 0.06, z = 8.28, p < .001). A second model

\(^2\) Age was included in the model because a preliminary model indicated that it was a significant predictor. Additional factors were shown to have no effect on performance in preliminary models, and were therefore not included (Gender: \(\beta = 0.86, SE = 0.45, z = 1.84, ns\); Order of Round Type: \(\beta = 0.72, SE = 0.46, z = 1.57, ns\)).
was run with an interaction term between Condition and Round Type; the interaction term was not significant (β = 0.06, SE = 0.25, z = 0.25, ns).

![Figure 2. Proportion of correct responses on the first and final rounds of training as a function of Condition (Action vs. Gesture Training) and Round Type (Doing vs. Seeing).](image)

Allowing for the fact that the verbs were not fully learned by any groups after four rounds, our findings point to two effects: (1) Children learn new verbs more quickly (given limited training) when they do the movements themselves than when they see others perform the movements; this effect does not depend on whether the verbs are learned through action or gesture. (2) Action training leads to faster acquisition of verbs (again given limited training) than gesture training; but given the increases in performance across the four trials, it is likely that children in all conditions would have learned the words if given enough time and training.

4. Discussion

Our goal was to investigate children’s ability to learn new verbs for actions performed on objects through different movement experiences. We were interested in how learning would unfold both when words were learned through action versus gesture experience, and when words were learned through doing these movements versus seeing these movements. We found that children made learning gains regardless of the type of training they were given. However, their knowledge of the taught verbs was significantly better after action experience
than after gesture experience. Moreover, they learned more from producing the movements themselves (actions or gestures) than from observing others’ movements (actions or gestures), and the disparity between learning through seeing and doing was not different across the conditions. The next sections consider these findings in relation to previous work on learning through doing vs. seeing movement, learning through gesture vs. action, and going beyond the information taught.

4.1. Effect of Doing Movement vs. Seeing Movement on Learning

Researchers have shown that children and adults learn 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). Our studies replicate these general findings. 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 learned words immediately after four rounds of training, we did not find an interaction between the type of movement used during training (action, gesture) and whether the movement was produced or observed (doing, seeing). Thus, the mechanism underlying children’s ability to learn information through producing rather than observing movement may be the same, be the movement action or gesture. This finding suggests that the power of learning through self-produced 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.

4.2. Effect of Action vs. Gesture Experience on Learning

In addition to varying whether children learned through doing versus seeing movement, we also varied whether children learned the novel words through action versus gesture experience. We found a main effect of movement type on children’s knowledge of the novel verbs after four rounds of training: Children performed significantly better if they learned through action experience than through gesture experience. As our dependent measure was children’s ability to perform the correct action (rather than the correct gesture) when prompted with one of the novel verbs, this finding could reflect task demands. If, however, our results were due to task demands, we might expect children to produce significantly fewer actions at test after learning through gesture than after learning through action. But children in the action and gesture conditions were equally likely to produce actions at test. Indeed, children in the gesture condition produced gestures only 5% of the time during test. This pattern suggests that children in the gesture condition understood the gestures they experienced during training as representations of actions that could be performed on objects.
Nevertheless, children who learned through gesture produced fewer correct actions at test than children in the action condition. Children thus seem to be able to learn some types of information more quickly through action than gesture. This effect could reflect the congruency between the behavior taught in the action condition and the behavior required at test (which was an action), but it could also stem from the fact that action provides a more complex sensory experience that leads to quicker learning than gesture (be it produced or observed by a learner).

4.3. Beyond Learning what is Taught: Gesture may Promote Generalization

Our results indicate that learning verbs for actions on objects through action can lead to success more quickly than learning the same verbs through gesture. An open question for future research is what exactly are children learning from action versus gesture. Our findings suggest that the children in our study interpreted the novel words they were taught as actions. In both conditions, children predominately produced actions on objects in response to the experimenter’s prompt: “Can you show me ratching?” There were cues to guide children towards interpreting the meaning of each novel word as an action on an object – in the initial directions, children were told they would learn words for “movements they could do with their hands.” However, it is an open question as to whether children tied the meaning of the action words they learned to the objects on which the words were learned. Because we did not give children an opportunity to generalize ratching to other objects, our data do not allow us to distinguish between a child who understood ‘ratching’ as a twisting movement that can be performed on any object that afforded twisting, and a child who understood ‘ratching’ as a twisting movement that must be performed on the training object — children of both types would produce the twisting movement on the training object when asked to demonstrate ‘ratching’. Previous work has shown that children often have difficulty generalizing verbs because they believe the object on which an action is demonstrated is just as important to verb meaning as the action (Kersten & Smith, 2002). Given the properties of action and gesture, gesture may lead to more flexible learning than action, allowing children to extend the action label they learned to new objects.

4.4. Conclusions

Our study was designed to understand how different forms of movement experience affect children’s ability to learn words for actions on objects. We ask about the effects of learning through action versus gesture, and the effects of doing versus seeing action or gesture, all within the same study. Our findings add weight to previous work showing that children learn better through doing movement than seeing movement, and we show, for the first time, that this effect is the same whether the movement is an action on an object or a gesture off an object. Although children seem to learn action words more quickly
through action than through gesture, gesture is nevertheless a viable teaching tool, one that has the potential to promote generalization (see Novack et al., 2014).

References


