



Experimentally Induced Increases in Early Gesture Lead to Increases in Spoken Vocabulary

Eve Sauer LeBarton, Susan Goldin-Meadow & Stephen Raudenbush

To cite this article: Eve Sauer LeBarton, Susan Goldin-Meadow & Stephen Raudenbush (2015) Experimentally Induced Increases in Early Gesture Lead to Increases in Spoken Vocabulary, Journal of Cognition and Development, 16:2, 199-220, DOI: [10.1080/15248372.2013.858041](https://doi.org/10.1080/15248372.2013.858041)

To link to this article: <https://doi.org/10.1080/15248372.2013.858041>



Published online: 08 May 2015.



Submit your article to this journal [↗](#)



Article views: 1033



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 40 View citing articles [↗](#)

Experimentally Induced Increases in Early Gesture Lead to Increases in Spoken Vocabulary

Eve Sauer LeBarton, Susan Goldin-Meadow, and Stephen Raudenbush

The University of Chicago

Differences in vocabulary that children bring with them to school can be traced back to the gestures they produced at the age of 1;2, which, in turn, can be traced back to the gestures their parents produced at the same age (Rowe & Goldin-Meadow, 2009a). We ask here whether child gesture can be experimentally increased and, if so, whether the increases lead to increases in spoken vocabulary. Fifteen children aged 1;5 participated in an 8-week at-home intervention study (6 weekly training sessions plus follow-up 2 weeks later) in which all were exposed to object words, but only some were told to point at the named objects. Before each training session and at follow-up, children interacted naturally with caregivers to establish a baseline against which changes in communication were measured. Children who were told to gesture increased the number of gesture meanings they conveyed, not only during training but also during interactions with caregivers. These experimentally induced increases in gesture led to larger spoken repertoires at follow-up.

Children increase their spoken vocabularies at very different rates during the first few years of life (Bates et al., 1994; Fenson et al., 1994). The disparities in vocabulary size that result from these early differences appear to matter, as early vocabulary is a key predictor of school success (R. C. Anderson & Freebody, 1981; Dickinson & Tabors, 2001; O'Neill, Pearce, & Pick, 2004). One predictor of children's vocabulary development in the first few years is their early gesturing. Before children speak, most children communicate using gesture, and at these ages, advances in emerging language skills can be measured through both gestured and spoken modalities (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). Naturalistic studies have shown that children's early gesture use relates to later word learning in infancy and through preschool (Acredolo & Goodwyn, 1988; Iverson & Goldin-Meadow, 2005; Rowe & Goldin-Meadow, 2009b). For example, the greater the diversity of meanings a child conveys in gesture at the age of 1;2, the larger the child's vocabulary at school entry (Rowe & Goldin-Meadow, 2009a). As another example, the gesture–speech combinations children produce early in development predict the age at which they first produce two-word utterances (Iverson & Goldin-Meadow, 2005; see also Ozcaliskan & Goldin-Meadow, 2005).

Although these data underscore the point that gesture and speech form an integrated system, they do not bear on the degree to which one component of the gesture–speech system can impact

Eve Sauer LeBarton is now at the Johns Hopkins School of Medicine and the Kennedy Krieger Institute. Correspondence should be sent to Eve Sauer LeBarton, Center for Autism and Related Disorders, Kennedy Krieger Institute, 3901 Greenspring Ave., Baltimore, MD 21211, USA. E-mail: lebarton@kennedykrieger.org
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hjcd.

the other. Does gesture merely correlate with subsequent vocabulary development, or can it play a causal role in fostering that development? Gesture has been shown to play a causal role in older children learning to solve mathematical equivalence problems. Telling 9-year-old and 10-year-old children to gesture either prior to (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007) or during (Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Cook, & Mitchell, 2009) a math lesson facilitates their ability to profit from that lesson. In one case (Broaders et al., 2007), children were told to move their hands while explaining their responses to a set of math problems. Children complied and gestured, but the interesting result was that the gestures they produced when told to move their hands revealed knowledge about the problem that they had not previously revealed in either speech or their solutions to the problem. Having articulated this knowledge in the manual modality, the children were then more open to input from a math lesson.

In another study in which gesture was manipulated (Goldin-Meadow et al., 2009), children were not just told to move their hands, but they were shown which movements to make in relation to the problem. The movements instantiated the “grouping” strategy, a strategy that results in correct answers to the problem but one that the children did not produce prior to instruction in either gesture or speech. Children taught how to move their hands to instantiate a correct grouping strategy solved significantly more problems correctly after instruction compared with children taught to move their hands to instantiate a partially correct grouping strategy and children given no instructions at all about their hands. Importantly, children’s success after training was mediated by how well they had learned the grouping strategy (as measured by whether they produced this strategy in speech on the posttest after the lesson). The grouping strategy was not explicitly mentioned in speech by either teacher or child, nor was it mentioned in gesture by the teacher in any of the conditions. Grouping was found during instruction only in the gestures that the children were taught to produce, suggesting that their newfound math knowledge came from abstracting information from their own hands. Although we do not yet understand the precise mechanism underlying the effect that learners’ gestures have on their own learning, it seems clear that producing a hand movement with meaning that is relevant to a task primes learners to take in information that has the potential to facilitate change—in this case, positive change—on the task.

Our goal here is to explore whether gesture can play the same type of causal role for young children in the early stages of language learning. Children’s own actions have been shown to direct their attention to aspects of the environment important for acquiring particular skills (for review, see Rakison & Woodward, 2008). The gestures a child produces are, in fact, actions, and when produced in a naming context, they could increase attention to both the objects named and to the names themselves. Increasing attention in this way has the potential to be useful for early word learning, particularly because object names are among the first words learned and account for a majority of children’s early spoken lexicons (Fenson et al., 1994). In addition, increasing pointing gestures in labeling contexts could emphasize the pragmatic utility of pointing for the transmission of knowledge, including transmitting names for objects (Csibra & Gergely, 2009). A child’s own gesturing thus has the potential to affect the child’s word learning.

Although there is much observational research on gesture and language development during the first 2 years of life, very little experimental work has been done in this domain at these early ages. The one exception is a study by Goodwyn, Acredolo, and Brown (2000), which provides indirect evidence that child gesture may have causal effects on child language development. Goodwyn and colleagues trained parents to model gestures for their 11-month-old infants in

everyday interactions and to teach the children to use the gestures (training condition). They found that children in the gesture-training condition performed significantly better than controls on a variety of language measures assessed through age 3 years. Importantly, however, the study did not address the question that is the focus of the current study—was increased child gesture the *cause* of the experimental effects on child language? Children’s gesture was not directly observed, and effects of increases in child gesture on child speech were not tested. Thus, it is possible that the effects that parent gesture training had on child language operated through a mechanism other than child gesture.

To determine whether gesture plays a causal role in vocabulary learning, as in the math studies with older children, we need to go beyond observing the spontaneous gestures young language learners produce and begin to manipulate their gestures. We accomplish this goal here by showing children pictures of objects, pointing to one of the pictures while labeling the object, and then telling the child to put his or her pointing figure on the picture (e.g., “that’s a dress,” said while pointing at a picture of a dress, followed by “can you put your finger here?”). We compare this experimental condition (child and experimenter gesture [C&EG]) to two control conditions, one in which the experimenter labels the picture while pointing at it but does not instruct the child to point (experimenter gesture [EG]) and one in which the experimenter labels the picture and does nothing else (no gesture [NG]). Our goal was to change children’s pointing behavior in interactions with the experimenter and then examine whether this experimental manipulation had an effect on the children’s spontaneous interactions with their caregivers. The two control conditions resemble behaviors that parents typically do while interacting with their children and thus might not be special enough to have a ripple effect outside of the experimental context. In contrast, the experimental condition involves a somewhat unusual behavior (telling children to put their finger on a picture) that might be expected to receive notice not only from the child, but also from their caregivers. We therefore examined gesturing behavior during the naturalistic caregiver–child interactions during the course of 6 weeks for both caregiver and child.¹ There is reason to believe that a manipulation carried out in the lab can have far-reaching effects outside of the lab (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Smith and colleagues (2002) conducted a 9-week longitudinal study on 17-month-old children and found that children’s experiences with object naming based on shape in the lab had dramatic effects on the number of object names the children added to their vocabularies outside of the lab. Our hypothesis is that telling children to gesture in an experimental context will have a similarly powerful effect on their use of gestures and words outside of the experimental context.

Telling children to point at pictures of objects could influence word learning in a variety of ways. First, encouraging children to point at an object while the experimenter is labeling that object could focus their attention on the particular word–object link. This focused attention might be expected to increase the likelihood that children will learn the labels for the particular objects to which they pointed but have no effect on their word-learning skill overall. Alternatively, encouraging children to point at objects could focus their attention on the referential properties of pointing gestures in particular, and all gestures in general. Attending to gestures’ referential properties might be expected to improve word learning overall, although it might

¹A gesture intervention in which a gesture was modeled for the child and the child was then encouraged to use it has been shown to successfully increase gesturing outside of the training context in children with autism (Ingersoll, Lewis, & Kroman, 2007). Our study evaluates whether the strategy also works with typically developing children.

have little effect on learning the particular words to which the children were trained to point. Our goal was to test this second hypothesis. Note that this hypothesis depends on the pointing gesture being referential. There is, in fact, evidence for this claim from data on young language learners. As mentioned earlier, the gesture–speech combinations a child produces early in development predict when that child will begin to combine words into a sentence. More specifically, the age at which a child first conveys information in gesture that differs from the information conveyed in the accompanying speech (e.g., “mommy” + point at hat) predicts the age at which that child will produce her first two-word utterance (“mommy hat”). Importantly, it is not just the presence of gesture that signals the onset of two-word speech, but it is the meaning of the gesture taken in relation to the meaning of the accompanying word—the age at which a child first conveys the *same* information in gesture and the accompanying speech (e.g., “hat” + point at hat) does *not* predict the onset of two-word speech (Iverson & Goldin-Meadow, 2005). Because most of the gestures children produce during this early stage of language learning are pointing gestures, this finding can be taken as evidence that the pointing gesture serves a referential function for young children.

In addition to these two potential mechanisms that affect the word learner directly, telling children to point at pictures of objects could also have ripple effects that shape the learning environment, thus affecting word learning indirectly. For example, the more children gesture, the more opportunities their caregivers have to respond to those gestures with targeted input; if a child points at a bird, the child’s mother could respond by saying, “That’s a bird,” thus providing the child with timely input that might facilitate learning the word “bird” (Goldin-Meadow, Goodrich, Sauer, & Iverson, 2007; Golinkoff, 1986; Masur, 1982). As another example, an increase in child gesture could evoke an increase in caregiver gesture, which might naturally bring with it an increase in caregiver speech; that increase in caregiver speech might then have an impact on child word learning (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991).

In summary, observational research on the relation between gesture and language development and experimental research on the role gesture plays in learning in other cognitive domains give us reason to believe that gesture could play a role in facilitating the transition to spoken language. If so, telling children who are at the earliest stages of word learning to gesture ought to have an impact on the words they subsequently produce. At present, however, there is no evidence that we can alter the rate at which young children gesture. The goals of this article are thus twofold: a) to determine whether we can increase children’s gesturing through experimental manipulation, and if so, b) to determine whether this increase has a positive effect on the number of spoken words children produce in a naturalistic setting. We predict that if we can experimentally increase child gesture, we will also increase child speech.

METHOD

Participants

Fifteen infants (8 boys) with no known medical conditions and their primary caregivers participated in the study. Families were recruited through a university-maintained infant database. We enrolled children ($M_{\text{age}} = 1;5$; range = 1;4–1;6) who met the following criteria: a) monolingual

English speaker, b) spoken vocabulary of fewer than 50 words at the start of the study, as measured by the MacArthur-Bates Communicative Development Inventory (CDI; Fenson et al., 1993), and c) no multiword utterances. We chose to enroll children with a spoken vocabulary of fewer than 50 words because we wanted to study children at the earliest stages of word learning (Bates et al., 1994). We did not ask caregivers whether their child was exposed to a sign language (e.g., “baby sign”). However, we did observe 4 caregivers using signs at pretest (1 C&EG, 1 EG, 2 NG). In addition to our sample of 15 infants, there were 4 children (2 C&EG, 1 EG, 1 NG) who enrolled in the study but were unable to maintain weekly visits and were thus not included in this study.

Although we did not recruit children based on socioeconomic status (SES), families in our sample were from a limited SES range (as determined by primary caregiver education). All children in the study were of middle to high SES; primary caregivers had, at minimum, some education beyond a high school degree and, at maximum, a doctoral degree (7 with education beyond high school including a bachelor’s degree; 8 with education beyond a bachelor’s degree). Most children were Caucasian and all were Non-Hispanic (11 White, 1 Black or African American, 3 Mixed ethnicity).

Procedure

Children received in-home experimental training once a week for 6 weeks. Two weeks after the final training session, we observed unstructured caregiver–child interaction in an in-home follow-up session. Each training session began with a 30-min videotaped play session between the child’s primary caregiver and the child to establish a baseline against which to measure changes in the child’s communication over time. The same primary caregiver was observed at all six training sessions and at follow-up. Primary caregivers included 13 mothers, 1 father, and 1 grandmother. The same procedure was used for the caregiver–child interaction at follow-up as was used for each of the training sessions.

After each caregiver–child interaction, the experimenter played with the child in a videotaped 30-min structured play segment during which training was administered. During this time, the experimenter played with a standard set of toys that included the experimental stimuli (picture books). In addition to administering the experimental manipulation with the picture books (see Stimuli section), the child and experimenter participated in free play with the standard set of toys brought by the experimenter (e.g., balls, toy animals, and toy cars). If the child needed a break at any time from the picture books, we resumed play with the free-play toys, or rarely, the child took a break with a parent. When not engaging in the experimental manipulation, the experimenter kept labeling and gesture production to a minimum. In all experimental conditions, children interacted with the experimenter for 30 min. In some cases, the experimenter had difficulty engaging the child in the picture book activity, and in these cases, the experimenter–child interaction period lasted slightly longer than 30 min because many breaks were taken to play with the other toys. The same experimenter was used for all sessions and all conditions.

Children were randomly assigned to one of three conditions (five per condition), with an eye toward matching the conditions on child sex and vocabulary size: C&EG, EG, and NG. Caregivers completed the CDI Words and Gestures form within 1 week prior to our first visit. We used this CDI to enroll children in the study and assign children to conditions, and we used

TABLE 1
Pretest Child Speech and Gesture Measures by Condition

Condition	Age	Standardized Measures of Child Speech		Child Communication During Caregiver–Child Interactions	
		Pretest Child Spoken Vocabulary	Pretest Child Comprehension Vocabulary	Pretest Child Speech	Pretest Child Gesture
No Gesture ($n = 5$, 3 M)	516.80 ($SD = 31.25$; 491–567)	20.00 (14.30)	123.20 (48.52)	7.40 (6.77)	15.60 (10.16)
Experimenter Gesture ($n = 5$, 3 M)	505 ($SD = 9.75$; 490–514)	19.80 (14.96)	117.00 (107.46)	11.40 (4.72)	9.80 (2.59)
Child and Experimenter Gesture ($n = 5$, 2 M)	508 ($SD = 11.28$; 491–518)	25.60 (12.08)	188.00 (76.86)	9.40 (6.27)	10.00 (6.96)

Note. The table contains means for all pretest measures; standard deviations and ranges are provided in parentheses. The standardized measures of child speech were obtained from the MacArthur-Bates Communicative Development Inventories. Age was balanced across conditions.

the spoken vocabulary and comprehension vocabulary measures from the CDI as pretest measures. In addition, after completing the observations, we coded the speech and gesture that children produced during the caregiver–child interaction that took place prior to the first experimental session, and we used these scores as pretest measures as well.

Table 1 presents pretest measures for children in each condition, two standardized measures from the CDI completed by mothers prior to Session 1 (pretest child spoken vocabulary and pretest child comprehension vocabulary), and two measures from the caregiver–child interaction at Session 1 (pretest child speech and pretest child gesture). There were no significant differences across conditions in any pretest measures (all two-tail $p > .34$). However, our small sample size may lead to low power, which could limit our ability to detect differences on pretest measures. As a result, we controlled for pretest child speech in our analysis of indirect effects on child speech outcome; the pretest child speech measure was taken from the same observational context from which our speech outcome measure was taken and, thus, we believe, is the most appropriate covariate. In addition, as seen in Table 1, children in our experimental condition, C&EG, had higher caregiver-reported spoken vocabulary and comprehension vocabulary scores than did children in the other two conditions, EG and NG. Even though the difference was not significant, we thought it important to check whether taking these individual differences in caregiver-reported pretest vocabulary measures into account would affect our results. We first reran our analyses of indirect effects on child speech outcome including spoken vocabulary as a covariate. We next reran these analyses including comprehension vocabulary as a covariate. In both cases, the results were unchanged. Similarly, we reran analyses using pretest child gesture as a covariate and results were unchanged.

Experimental manipulation. During each training session, the child and experimenter looked at two six-page picture books. In all conditions, the experimenter provided a spoken label

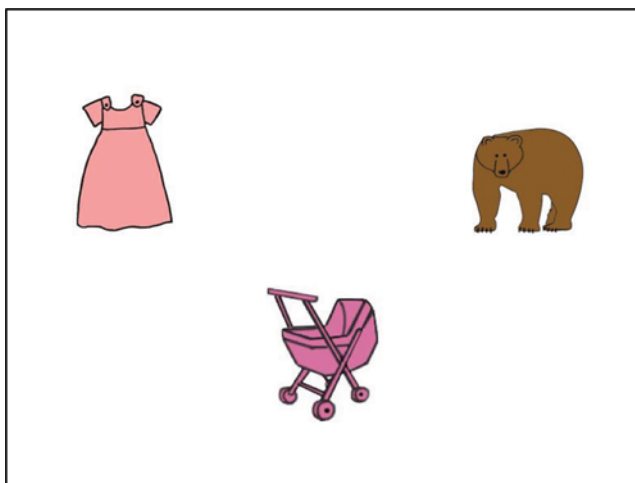
for one of the three pictures on each page (the target picture). The experimenter produced the label twice and paused between each labeling utterance (e.g., “*Look at the dress. That’s a dress.*”). All target labels were provided when the child was attending to the page.

In the experimental condition, C&EG, the experimenter pointed at the target picture while providing the spoken label for the picture and then asked the child to point to/put her finger on the picture. For example: *Experimenter Target Utterance 1: “Look at the dress” + EXPERIMENTER POINTS AT DRESS; Experimenter Gesture Instruction: “Can you do this?” CHILD POINTS AT DRESS; Experimenter Target Utterance 2: “That’s a dress” + EXPERIMENTER POINTS AT DRESS.*

We also had two control conditions in which children were not told to gesture. In the first control condition, EG, the experimenter pointed at the target while providing the spoken labels but did not ask the child to point. In the control condition, NG, the experimenter provided the spoken labels but did not point at the target and did not ask the child to point. Having two control conditions allowed us to isolate the effect of seeing versus doing gesture. We grouped the conditions into two groups as a function of instruction to gesture: told to gesture (C&EG) and not told to gesture (EG, NG). In all analyses, we predicted an effect of instruction to gesture.

To increase the child’s compliance with our instructions to gesture, the experimenter provided enthusiastic positive reinforcement (e.g., “Good job!”) following the child’s pointing gesture in the C&EG condition. Because children in the EG and NG conditions were not being encouraged to point, they did not receive this type of reinforcement. However, they did receive positive acknowledgement (e.g., nod or “yes”) to their responses. Further, in all conditions, the experimenter was positive and engaging during book reading to encourage the child’s participation in the task. In all conditions, if the child became distracted and wanted to stop playing with the book, the experimenter tried to reengage the child in the task, and if that failed, a break was taken during which the experimenter and child played with an item from the standard set of toys. Sometimes the child began to look away from the book before the experimenter completed producing the label for the target word. In this case, the label was repeated. During the 6 weeks, 720 target label utterances were planned for each condition (144 for each child). Including repeats, the number of target label utterances produced for each condition was: C&EG = 738, EG = 748, NG = 744. The mean number of target utterances produced to each child during the course of the study did not differ between conditions (group means, C&EG = 146.80, EG = 148.60, NG = 145.00; $p > .41$).

Stimuli. We designed four six-page picture books to use in training. The target labels and books were identical across conditions. We used labels that children might encounter in their everyday lives (e.g., “dress”) because the books were meant to serve as a medium within which we could teach children to gesture in a communicative context. For all children in all conditions, two books were presented at each session, and we alternated books between sessions (Books A and B were shown at Sessions 1, 3, and 5; Books C and D were shown at Sessions 2, 4, and 6). Each page of the book contained three colored line drawings (Figure 1). One picture on each page was the target, and the experimenter labeled only the target. All target words were whole-object labels (i.e., not labels for the parts of the object). None of the pictures appeared on more than one page or in more than one book. The target items were selected prior to beginning the study so that they would represent a spectrum of unknown and familiar words similar to the words young children are likely to encounter in everyday interactions. The target object

FIGURE 1 Sample stimuli page (target item = *dress*).

labels and nontarget object labels were of varying levels of difficulty—some items were object labels that most 16-month-olds knew (e.g., cat) and some were known by very few 16-month-olds (e.g., giraffe). We used CDI age norms for the proportion of 16-month-old children who comprehended each lexical item to assign a difficulty level to each lexical item (Dale & Fenson, 1996). We divided words into three groups from “easy” (known by many children) to “hard” (known by few children), with an equal distribution across the three groups. Keeping these items constant across children and conditions allowed us some control over what the children saw, heard, and did and also allowed us to explore whether any effects we might find were due to learning these words in particular, or to learning to use gesture in general. During the course of the study, children were given labels for 24 target items, each presented at three sessions.

We did not include assessments of how well the children learned the 24 words that we explicitly taught for two reasons. First, our focus was on whether the experimental manipulation would have an impact on the children’s overall word-learning skills (which we measured by assessing the number of words children used in conversation with their caregivers), not on their acquisition of these particular words. Second, we thought that including a test session probing the 24 words might encourage children to focus on the taught words to the exclusion of others. However, we were able to look at whether children in each group used the 24 target items in their spontaneous speech with caregivers during the follow-up session. Overall, children produced too few of the target items for us to systematically investigate use of the target words. Out of the 15 children, only 5 children produced target words in their spontaneous speech at follow-up, with a maximum of 2 words produced by any 1 child. Of the 5 children who produced at least 1 target word: 1 was in the NG condition, 1 was in the EG condition, and 3 were in the C&EG condition. To investigate whether our analyses of child speech outcome were influenced by the particular words taught during the training sessions, we ran our analyses predicting speech outcome again, but this time, we excluded all target words. Results were unchanged. Thus, we continue to include all children’s spontaneously produced words in our measure of child speech at follow-up.

Gesture Coding

Child-gesture-in-experiment measure. We coded all pointing gestures that children produced to the target pictures in response to the experimenter's target utterances for all six training sessions. The child's gesture could occur before, during, or after either of the two experimenter target utterances for that page, but the gesture was counted only if it pointed to the target item in the book. We then calculated the total number of target item presentations to which the child responded with a point. Because the child was given 72 opportunities to produce a pointing gesture during the course of the six sessions, possible scores ranged from 0 to 72.

Child-gesture-with-caregiver measures. Using previously developed criteria for coding communicative gesture (Iverson & Goldin-Meadow, 2005; Rowe & Goldin-Meadow, 2009a, 2009b), we coded all of the gestures that each child produced during 20 min of the caregiver-child interactions at Sessions 1 (pretest), 3, and 5 and at follow-up. Data were sometimes not available for the full 30 min (e.g., due to scheduling problems or video error); we therefore coded the first 20 min for each child.

Communicative hand movements that did *not* involve direct manipulation of objects (e.g., twisting a jar open) or a ritualized game (e.g., patty cake) were considered gestures (Goldin-Meadow & Mylander, 1984). Head nods and shakes were also considered gestures. Children produced three kinds of gestures: a) deictic gestures, which indicate concrete objects, people, or locations (e.g., pointing to a dog to indicate a *dog*); b) iconic gestures, which depict the attributes or actions of an object via hand or body movements (e.g., flapping arms to convey a *flying bird*); and c) conventional gestures, which have forms and meanings that are prescribed by the culture (e.g., shaking one's head side to side to convey *no*).

We coded each gesture for meaning and then calculated the number of unique gesture meanings (gesture types) that each child produced at each session. For example, a point at a dog was assumed to mean *dog* and thus counted as a gesture type. This procedure for attributing meaning to early child gesture has been used in previous studies and results in systematic patterns (Rowe & Goldin-Meadow, 2009a, 2009b). An independent coder blind to condition membership performed gesture coding on 10% of the videos. Percent agreement between coders was 86% for identifying gestures and 95% for assigning meanings to gestures.

Our measure of *pretest child gesture* is the number of gesture types the child produced when interacting with the caregiver at Session 1. The *child-gesture-with-caregiver* measure is the slope of the number of gesture types each child produced across Sessions 1, 3, and 5 and at follow-up. A slope was calculated individually for each child by performing child-specific regressions using session (Sessions 1, 3, and 5 and follow-up) as the predictor and gesture type as outcome. Each child's slope was the estimated regression coefficient resulting from this analysis.

Speech Transcription

We transcribed child speech during the same 20 min of caregiver-child interaction coded for gesture. Sounds that were reliably used to refer to entities, properties, or events (e.g., "doggie," "nice," "broken"), along with onomatopoeic sounds (e.g., "meow," "choo-choo") and conventionalized evaluative sounds (e.g., "oopsie," "uh-oh"), were counted as words. Morphologically inflected variants of words (e.g., "run," "running") were considered a single

type. An independent coder blind to condition membership coded half of the videos. We established reliability for our speech transcripts on 10% of the transcripts. Percent agreement between coders was calculated at the utterance level, and transcribers were reliable when percent agreement reached 90%.

Our measure of *pretest child speech* was the number of unique word types the child produced when interacting with the caregiver at Session 1. The *child-speech-with-caregiver-at-follow-up* measure, the speech outcome measure used in our analyses, was the number of unique word types the child produced when interacting with the caregiver at the follow-up session.

Analytic Plan

We analyzed our data in two parts. First, we examined instruction effects on child gesture (our mediator) and instruction effects on child speech (our outcome). Based on our a-priori predictions, we used the instruction-to-gesture group as our experimental instruction variable. Thus, analyses of instruction effects test differences between the: a) told-to-gesture group (the C&EG condition) and b) not-told-to-gesture group (i.e., both control conditions, EG and NG). Although we focused on differences between instruction-to-gesture groups, we did examine whether there were differences between the two control conditions and found no significant differences between the EG and NG conditions on any of our mediator and outcome measures (all two-tail $p > .21$). We consider possible differences between control conditions in our discussion of future directions.

Second, we tested our hypothesis that increases in child gesture (our mediator) lead to increases in child speech (our outcome). We did this by testing indirect effects of our instruction (instruction-to-gesture group) on child speech through child gesture. To account for variability that may be attributed to differences between control conditions, we included control group assignment as a residual contrast in our indirect effects analyses. The residual contrast was not significant in any of our models (two-tail $p > .23$).

In developing this analytic plan, we paid particular attention to two aspects of the study. First, we hypothesized an *indirect effect* of the experimental intervention on speech outcome. Because we were investigating causal effects of child gesture on child speech, we predicted that the intervention would impact child speech only to the extent that it impacts child gesture. Thus, analyses of simple effects of our intervention (comparison of group means on our outcome variable) would not adequately test our hypothesis. For instance, we could observe a total effect of intervention on speech outcome but not know whether it was child gesture that was causing the effects on child speech. Alternatively, we could *not* find an effect of intervention on speech outcome and not know whether individual differences in the potency of our intervention on child gesture masked gesture effects on child speech. Second, our sample size was small and thus may have been sensitive to violations of assumptions of traditional parametric statistics. To address both of these concerns, a) in addition to testing total effects, we tested our hypothesis by using indirect effects analysis, and b) in addition to using parametric statistics, we used permutation analyses when possible.

Statistical Analyses

We analyzed our data using both parametric and nonparametric methods and present results for both. To test the significance of our effects using nonparametric methods, we used permutation

analysis. For permutation analysis, we performed approximate randomization tests (Good, 2000; MacKinnon, 2008). Rather than resample all possible rearrangements of our data as in a true permutation test, we set our number of random resamples at 1,000 to approximate the complete randomization distribution (Hesterberg, Moore, Monaghan, Clipson, & Epstein, 2006). For all randomization tests: a) We randomly resampled without replacement in a manner consistent with the null hypothesis (e.g., randomly reassigning conditions); then b) for each of the 1,000 resamples, we calculated the statistic of interest (e.g., mean difference between groups); and finally, c) we used the values generated in each resample to construct the permutation distribution. The proportion of values greater than or equal to the values observed in our sample was the one-tail p value (Hesterberg et al., 2006). For all tests involving instruction-to-gesture effects, we permuted instruction-to-gesture by randomly assigned conditions and then collapsed the permuted conditions into instruction-to-gesture groups: told-to-gesture (C&EG condition) versus not-told-to-gesture (EG and NG conditions). Because our predictions were directional, we present one-tail p values for all analyses. When statistical significance differed for two-tail tests, two-tail p values are reported. Unless otherwise stated, all results that are significant at $\alpha = .05$ with a one-tail test are also significant using a two-tail test.

RESULTS

We hypothesized that telling children to gesture would increase child gesture and, through child gesture, increase child speech. We present findings for a) the total effects of our experimental manipulation (instruction to gesture) on children's gesture and speech, and b) the *indirect* effects of this experimental manipulation on child speech through child gesture. The total effects analyses in Part 1 do not include any covariates. However, we provide results from specification checks that include covariates for all Part 1 analyses. The indirect-effects analyses in Part 2 include pretest child speech as a covariate. As mentioned earlier, if we rerun indirect-effects analyses including comprehension vocabulary or pretest child gesture as a covariate, the results are unchanged.

Part 1: Instruction Effects on Child Gesture and Speech

Instruction effects on gesture. We first asked whether instructing children to gesture increased child gesture during the experiment (Figure 2). We found a significant effect of instruction to gesture on *child gesture in experiment*, using both nonparametric² (randomization test, $M_{\text{difference}} = 35.9$; $p = .002$) and parametric, $t(13) = 5.074$, $p = .009$, tests.³ As predicted, children in the told-to-gesture group ($M = 46.6$, $SD = 21.3$) produced more gestures during the experiment than did children in the not-told-to-gesture group ($M = 10.7$, $SD = 6.3$).

²If we use a standard nonparametric Mann-Whitney test for our four Part 1 analyses, we obtain similar results.

³As specification checks, we reran analyses of effects of instruction to gesture on child gesture in experiment, this time including covariates, and we obtained similar results: a) pretest child speech as a covariate ($B = 11.967$, $SE = 2.244$, $\beta = .815$, $p < .001$; randomization test, $p < .001$); b) pretest child gesture as a covariate ($B = 12.424$, $SE = 2.387$, $\beta = .846$, $p < .001$; randomization test, $p < .001$). For both analyses, we obtained similar results to what we would have obtained if we had also included the residual contrast described in Part 2.

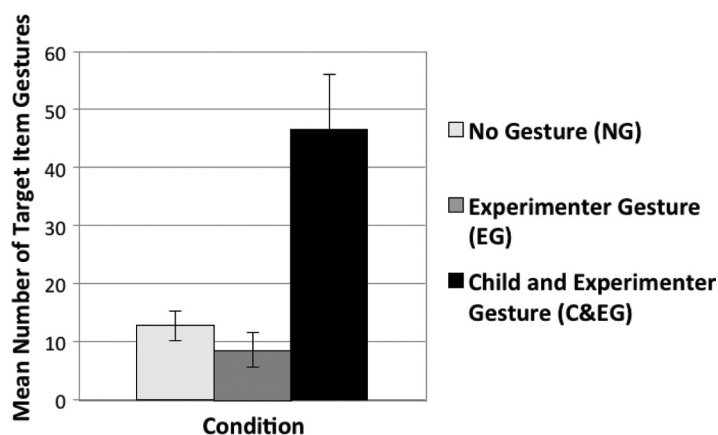


FIGURE 2 Child gesture in experiment. Mean number of gestures at targets produced by children in each condition during the six training sessions with the experimenter. Error bars represent standard errors of the mean.

We next asked whether this effect extended beyond the experimenter–child context to interactions with the caregiver. We first counted the number of gesture types the child produced at Sessions 1, 3, and 5 and at follow-up (Figure 3, top). We then calculated the slope of child gesture types from pretest to follow-up for each child, our *child-gesture-with-caregiver* measure (Figure 3, bottom), by calculating a regression line for each child using gesture types from Sessions 1, 3, and 5 and follow-up. There was a significant effect of instruction to gesture on child gesture with caregiver, using both nonparametric (randomization test, $M_{\text{difference}} = 1.902$, $p = .015$) and parametric, $t(13) = 2.531$, $p = .013$, tests.⁴ As predicted, children who were told to gesture ($M = 2.06$, $SD = 1.49$) had a greater increase in gesture with their caregivers during the course of the study than did children not told to gesture ($M = 0.16$, $SD = 1.31$).

To increase the stability of our gesture measure, in subsequent analyses, we combined the two child gesture measures (child gesture in experiment and child gesture with caregiver) into one *child-gesture-composite* score by averaging the z score of each. Child-gesture-composite scores were higher for the told-to-gesture group ($M = 0.95$, $SD = 0.64$) than for the not-told-to-gesture group ($M = -0.48$, $SD = 0.46$). Group differences were significant using both non-parametric (randomization test, $M_{\text{difference}} = 1.423$, $p = .001$) and parametric, $t(13) = 4.982$, $p < .001$, tests.⁵

As mentioned earlier, we did not find significant differences between our control conditions on our mediator and outcome measures. However, before continuing with our data analysis, we

⁴We reran analyses of effects of instruction to gesture on child gesture with caregiver, this time including covariates, and we obtained similar results: a) pretest child speech as a covariate ($B = 0.634$, $SE = 0.253$, $\beta = .575$, $p = .014$; randomization test, $p = .011$); b) pretest child gesture as a covariate ($B = 0.530$, $SE = 0.207$, $\beta = .480$, $p = .013$; randomization test, one-tail $p = .032$, two-tail $p = .068$). We obtained similar results to those we would have obtained if we had also included the residual contrast in each analysis.

⁵We reran analyses of effects of instruction to gesture on child gesture composite, this time including covariates, and obtained similar results: a) pretest child speech as a covariate ($B = 0.475$, $SE = 0.090$, $\beta = .810$, $p < .001$; randomization test, $p < .001$); b) pretest child gesture as a covariate ($B = 0.453$, $SE = 0.095$, $\beta = .773$, $p < .001$; randomization test, $p < .001$). We obtained similar results to those we would have obtained if we had also included the residual contrast in each analysis.

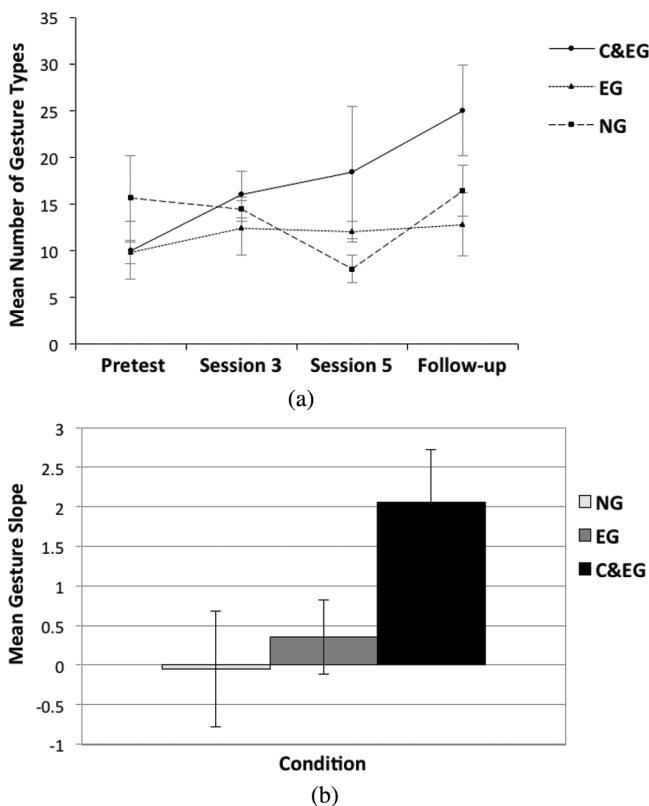


FIGURE 3 Child gesture with caregiver. Top: mean number of gesture types produced by children in each condition at Sessions 1 (pretest), 3, and 5 and follow-up. Bottom: growth in gesture types over sessions (mean slope) in each condition. Error bars represent standard errors of the mean. NG = no gesture; EG = experimenter gesture; C&EG = child and experimenter gesture.

used randomization tests of differences between means to explore whether the told-to-gesture condition differed from each of the control conditions. For *child gesture composite*, scores for the told-to-gesture condition (C&EG) significantly differed from both control conditions (EG, $M_{\text{difference}} = 1.410$, $p = .002$; NG, $M_{\text{difference}} = 1.438$, $p < .001$). Similar patterns were found for the component measures: child gesture in experiment (EG, $M_{\text{difference}} = 38$, $p = .001$; NG, $M_{\text{difference}} = 33.8$, $p < .001$); child gesture with caregiver (EG, $M_{\text{difference}} = 1.7$, $p = .052$; NG, $M_{\text{difference}} = 2.104$, $p = .012$). Thus, in the absence of evidence that our instruction operates differently for the two groups, we continued to follow our analytic plan of comparing groups based on the instruction-to-gesture group. However, in Part 2 analyses, we accounted for any potential variability between the two control conditions by including condition as a residual factor.

Instruction effects on speech. We next investigated whether child speech with caregiver at follow-up, which took place 2 weeks after the final experimental session, differed for instruction-to-gesture groups. At follow-up, children in the told-to-gesture group produced an average of 14.20 ($SD = 5.93$) word types and children in the not-told-to-gesture group produced

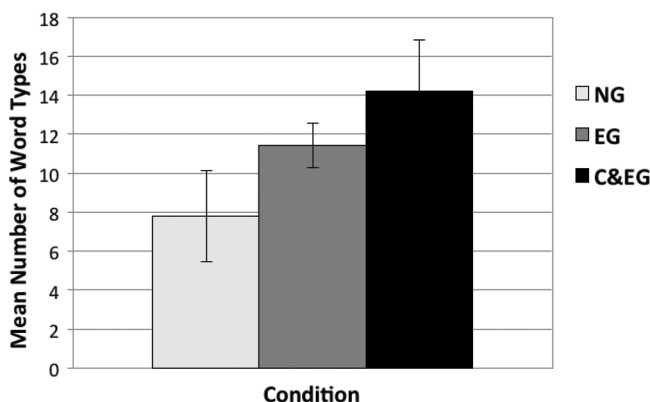


FIGURE 4 Child speech with caregiver at follow-up. Mean number of word types produced by children in each condition when interacting with caregiver during follow-up. Error bars represent standard errors of the mean. NG = no gesture; EG = experimenter gesture; C&EG = child and experimenter gesture.

9.60 ($SD = 4.35$; Figure 4). These differences were marginally significant, using both nonparametric (randomization test, $M_{\text{difference}} = 4.6$, $p = .056$) and parametric, $t(13) = 1.717$, $p = .055$, tests⁶. These differences were not statistically significant using a two-tail test (randomization test, $p = .12$; t test, $p = .11$).

Part 2: Indirect Effects on Child Speech through Child Gesture

We used indirect-effects analysis (MacKinnon, 2008) to explore whether our experimental manipulation has an *indirect effect* on child speech that operates through child gesture. Our hypothesis was that our randomly assigned experimental manipulation will impact child speech only to the extent that it impacts child gesture. Thus, we predicted that the effect of instruction to gesture (our treatment variable) on child speech with caregiver at follow-up (our outcome variable) will operate through child gesture composite (our mediator variable).⁷ As in our other analyses, we used both parametric and nonparametric tests. Nonparametric tests such as our randomization tests are useful with small sample sizes and can be used to test significance of indirect effects (MacKinnon, 2008). The *indirect-effects* analysis assumes that there are no unobserved confounders correlated with both the mediator and outcome (MacKinnon,

⁶If we include pretest child speech as a covariate in this analysis, we find an effect of instruction-to-gesture group on child speech with caregiver at follow-up ($B = 1.533$, $SE = 0.812$, $\beta = .430$, one-tail $p = .042$, two-tail $p = .083$; randomization test, one-tail $p = .062$, two-tail $p = .124$). If we instead include pretest child gesture as a covariate, there is a marginally significant effect ($B = 1.333$, $SE = 0.889$, $\beta = .374$, one-tail $p = .080$, two-tail $p = .160$; randomization test, one-tail $p = .083$, two-tail $p = .157$). Results for both analyses are similar if we also include the residual contrast.

⁷If we use the CDI toddler form, which was administered at follow-up, as our speech outcome measure in Path b, we find similar effects. Child gesture composite is a significant predictor of speech outcome, as measured by the CDI ($B = 41.555$, $SE = 22.521$, $\beta = .817$, one-tail $p = .0475$, two-tail $p = .095$; randomization test, $p = .002$). The Part 1 analysis of instruction effects is not significant, $t(13) = 0.394$, $p = .350$ (randomization test, $M_{\text{difference}} = 9.7$, $p = .366$).

2008). We addressed this assumption by controlling for potential confounders in the indirect-effects analysis.⁸

The indirect-effects analysis we used is similar to a mediation analysis but differs in the guidelines for evaluating the indirect effect. An indirect-effects approach requires a significant *Path a* (between treatment and mediator) and a significant *Path b* (between mediator and outcome, controlling for treatment) but does not require a significant relation between treatment and outcome (Hayes, 2009). The estimate of the indirect effect is the product ($a \times b$) where “*a*” is the unstandardized coefficient for the treatment in the *Path a* regression model and “*b*” is the unstandardized coefficient for the mediator in the *Path b* regression model (MacKinnon, Lockwood, & Williams, 2004; Preacher & Hayes, 2004). A significant indirect effect indicates that treatment and outcome are significantly related (albeit indirectly) and that the relation operates through the mediator.

We conducted our analysis in a series of three steps. In Steps 1 and 2, we performed a regression of child gesture composite on instruction to gesture (*Path a*) and then we performed a regression of child speech with caregiver at follow-up on child gesture composite, controlling for instruction to gesture (*Path b*). In Steps 1 and 2, we controlled for pretest child speech.⁹ We assumed there were no unobserved confounders correlated with both child gesture composite and child speech with caregiver at follow-up, while controlling for instruction to gesture and pretest child speech. In all of our regression models that included instruction to gesture as a factor, we also included a residual contrast factor. Because there were three conditions in the study, we needed two orthogonal contrasts to test our experimental effects. Our instruction-to-gesture factor, told to gesture (C&EG) versus not told to gesture (EG, NG), captured the effect of doing gesture versus not doing gesture. The residual contrast, EG versus NG, captured any differences between the two control conditions and thus tested the effect of seeing (but not doing) gesture versus no gesture at all. In Step 3, we calculated the product $a \times b$ from the values generated in Steps 1 and 2. In all steps, we tested significance using randomization tests but also reported significance from parametric tests.

In Step 1, we investigated whether instruction to gesture significantly predicted child gesture composite (*Path a*). Our regression model predicting child gesture composite included instruction to gesture, pretest child speech, and the residual contrast. In all regression analyses, the categorical variables were coded in the following manner: instruction to gesture (C&EG = 2, EG = -1, NG = -1) and residual contrast (C&EG = 0, EG = 1, NG = -1). Consistent with our findings in Part 1, we found that instruction to gesture predicts child-gesture-composite score, $B = 0.475$, $SE = 0.093$, $\beta = .810$, $p < .001$. We performed randomization tests of significance by permuting the regression model outcome variable (child gesture composite) and performing a regression analysis that included the permuted values of the outcome and the actual values of our predictor and covariates (M. J. Anderson & Legendre, 1999; Good,

⁸We also performed an instrumental variables analysis as a complementary method of testing indirect effects. In this analysis, random assignment is our instrument (in our study, instruction to gesture) and is thought to address the issue of omitted variables that could be confounders (Angrist & Krueger, 2001). We tested significance of the indirect effect using a randomization test, and both one-tail and two-tail tests were marginally significant (one-tail $p = .098$, two-tail $p = .099$). Though not conclusive, these findings suggest a causal effect of manipulated child gesture on child speech outcome, a result confirmed in our indirect-effects analysis.

⁹As reported earlier, if we instead control for our parent-reported pretest child language measures (pretest child spoken vocabulary and pretest child comprehension vocabulary), we obtain similar results.

2000). This method accounted for the possibility that the predictor variables were related. We found that instruction to gesture significantly predicts child gesture composite ($p < .001$)¹⁰ and thus meets our indirect-effects analysis criterion of a significant *Path a*.

In Step 2, we investigated whether child gesture composite relates to child speech with caregiver at follow-up, while controlling for our experimental manipulation, instruction to gesture (*Path b*). Our regression model predicting child speech with caregiver at follow-up included the following predictors: child gesture composite, instruction to gesture, pretest child speech, and the residual contrast. Child gesture composite predicted child speech with caregiver at follow-up, $B = 5.992$, $SE = 2.074$, $\beta = .984$, $p = .008$. We performed randomization tests of significance by randomly permuting our regression model outcome variable (child speech with caregiver at follow-up), and then we regressed our permuted outcome variable on the actual values of our predictors. Child gesture composite was a significant predictor of child speech with caregiver at follow-up ($p = .021$) and thus met the indirect-effects analysis criterion of a significant *Path b*. Children with higher child-gesture-composite scores produced more word types at follow-up than did children with lower scores (Figure 5, top).

In Step 3, we tested our prediction that there was an indirect effect of experimental manipulation (instruction to gesture) on outcome (child speech with caregiver at follow-up) that operates through child gesture (child gesture composite). We estimated the indirect effect by calculating the product $a \times b$ for our sample ($a \times b = 2.846$) and then testing significance with randomization tests. To do so, we recalculated *Paths a* and *b* for each iteration and then calculated the product $a \times b$ using the coefficients generated for *Paths a* and *b* in that iteration (MacKinnon, 2008). We generated the coefficients for *Path a* and *Path b* as follows: We permuted child gesture composite and child speech with caregiver at follow-up and used these permuted values in the *Path a* and *Path b* regression models presented earlier. We used actual values for the remaining variables in the models. We found a significant indirect effect of instruction to gesture on child speech with caregiver at follow-up through child gesture composite ($p < .001$; Figure 5, bottom).

To summarize, we were able to experimentally increase the rate at which children gestured and, by doing so, increase the number of different words they produced in spontaneous interactions with their caregivers. Our data permit us to ask one further question—did the caregivers play a role in increasing child gesture? Most of the caregivers watched the experimenter as she interacted with the child and thus might have, on their own, encouraged the children to gesture, or they may have increased their own gesture in response to the child's increase in gesture, thus reinforcing gesturing as a means of communication. We explore here whether any increases in caregiver gesture might have played a role in our effects on child gesture and child speech. If parents play a role in our effects, we would predict an effect of instruction to gesture on caregiver gesture and a relation between caregiver gesture and child gesture. However, if our effect is still operating through the child's own gesture production, we would find that child gesture remains a significant predictor of child speech when accounting for caregiver gesture.

To explore this possibility, we coded the gestures caregivers produced during the 20 min of caregiver-child interactions at Sessions 1, 3, and 5 and follow-up using the same system used for

¹⁰We also ran a randomization test using an alternative method of permuting the data that was similar to the one used in Part 1. We randomly permuted the instruction-to-gesture group but used actual values for all other predictors and outcome. Our findings are unchanged (one-tail $p < .001$).

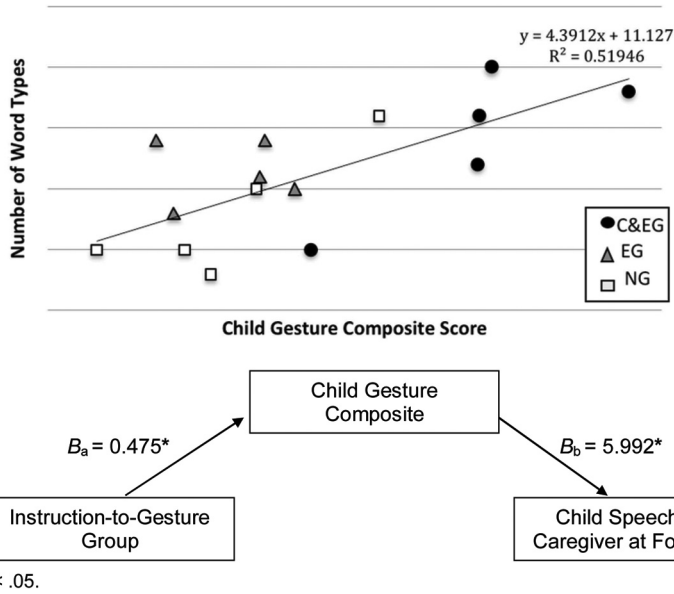


FIGURE 5 Predicting child speech. Top: mean number of word types children in each condition produced when interacting with the caregiver during follow-up, as a function of each child’s gesture composite. Bottom: indirect-effects analysis showing that the experimental manipulation (instruction to gesture) had an effect on child speech (child speech with caregiver at follow-up) through child gesture (child gesture composite). NG = no gesture; EG = experimenter gesture; C&EG = child and experimenter gesture.

the children and calculated the number of gesture types per session. We first examined whether differences in caregiver gesture at pretest may be underlying our findings. There were no significant differences between conditions in mean caregiver gesture at pretest (Session 1): C&EG = 20.40 ($SD = 2.61$); EG = 20.20 ($SD = 1.79$); NG = 21.20 ($SD = 6.02$); analysis of variance, $F(2) = 0.091, p = .457$; Kruskal-Wallis, $K(2) = 0.270, p = .437$. In addition, caregiver gesture at pretest did not relate to child gesture composite (p values for parametric tests and randomization tests $> .38$). Lastly, if we rerun our series of indirect-effects analyses using caregiver gesture at pretest as a covariate instead of pretest child speech, we obtain a similar pattern of results (all paths were significant at $p < .01$ using parametric and randomization analyses; the indirect effect was significant using the randomization test at $p < .001$).

We next calculated the slope of caregiver gesture production by performing caregiver-specific regressions using session (Sessions 1, 3, and 5 and follow-up) as the predictor and gesture type as outcome. Each caregiver’s slope was the estimated regression coefficient resulting from this analysis. We used slope of caregiver gesture production as our measure of *caregiver gesture*. Mean caregiver gesture slopes were: 2.30 ($SD = 1.29$) for C&EG; 0.05 ($SD = 1.08$) for EG; and -0.13 ($SD = 1.21$) for NG. We found a significant effect of instruction to gesture on caregiver gesture in a model that also included the residual contrast and caregiver gesture at pretest as covariates ($B = 0.767, SE = 0.211, p = .004$; randomization test, $p = .001$). Caregiver gesture also significantly related to our child-gesture-composite measure ($B = 0.387, SE = 0.104, p = .003$; randomization test, $p < .001$). Importantly, however, the increases in caregiver gesture

did *not* account for the increases in child speech: Including caregiver gesture as a covariate, child gesture composite still significantly related to child speech with caregiver at follow-up ($B = 4.951$, $SE = 1.732$, $p = .014$; randomization test, $p = .011$). The increases we see in caregiver gesture are nonetheless important as they suggest that we can, with little effort, engage caregivers in attempts to manipulate young children's gestures.

DISCUSSION

In summary, we found: a) We can increase infant (age 1;5) gesture through experimental manipulation, and b) this increase in child gesture has a positive effect on the number of different words children use in spontaneous interactions with their caregivers. We discuss each point in turn.

Increasing Child Gesture through Experimental Manipulation

Our study adds to the literature on early child gesture by demonstrating that it is possible to experimentally manipulate typically developing child pointing in an experimental context and that these experimental effects generalize to children's spontaneous gesturing in naturalistic interactions with caregivers. Compared with other studies (e.g., Goodwyn et al., 2000), our experimenter-administered intervention was relatively short and involved little time commitment—the intervention lasted 6 weeks, with one 30-min experimental session per week. The finding that gesture is malleable in children and can be manipulated by noncaregivers may have implications for the design of studies aimed at increasing child gestural communication not only in typically developing children, but also in children with developmental delays (e.g., Yoder & Warren, 2002). For both typically and atypically developing children, simply being able to increase infant gestural communication may have positive cascading consequences for learning (e.g., through the effect child gesture has on interactions with others, as will be discussed later). Lastly, although all children in the group that was told to gesture had larger increases in gesture than the control group children (with the exception of one child in the NG condition; Figure 5, top), there were individual differences in the potency of our intervention on child gesture. The fact that our manipulation did not have a uniform effect on child gesture (and that children in the control condition were not prevented from gesturing and thus could, and in some cases did, gesture) may account, at least in part, for the fact that the *indirect* effect of instruction through gesture was significant but the *direct* effect was only marginal. Future work is needed to investigate ways to increase the potency of our experimental manipulation.

Increases in Gesture Lead to Increases in Spoken Words

Our study provides evidence that gesture can be used to facilitate early spoken word use. We were able to experimentally increase the rate at which children gestured and, by doing so, increase the number of different words they produced in naturalistic interactions with their caregivers. Unfortunately, we did not have assessments of the children's comprehension vocabularies at the end of the study and thus do not know whether our manipulations affected the number of different words the

children had in their repertoires, or the number of different words they used when talking to their caregivers at home. Future work is needed to determine precisely where gesture has its effects.

Why did our instruction to gesture have an impact on the number of words children produced? Multiple mechanisms may underlie our findings. First, we found that the caregivers of children who were instructed to gesture increased their own gesturing during the course of the study, either because they discovered the importance of gesture from watching the experimenter or because they responded to increases in their child's gestures with increases in their own. An increase in caregiver gesture might bring with it an increase in caregiver speech. Increased speech could, in turn, have an impact on child vocabulary, particularly because the amount of talk parents direct to their children has been found to affect the size of the child's vocabulary (Huttenlocher et al., 1991).

A second possibility is that the increase in child gesture gives parents more opportunities to provide their children with targeted speech input. Parents' translations of their children's gestures (e.g., saying, "Yes, that's a dog" in response to a child's point to a dog) have been found to relate to child vocabulary (Goldin-Meadow et al., 2007; Golinkoff, 1986; Masur, 1982). The children in our study who were told to gesture increased the number of gestures they produced and thus gave their caregivers more opportunities for these translations. If caregivers respond to their children's gestures by providing a label for the object being pointed at, not only are they increasing the number of words they say to their children, but the words are produced at just the moment when the child is focused on the relevant object—a teachable moment.

These first two possibilities grow out of the fact that gesture is a communicative tool. Increasing the child's use of the tool has the potential to affect the kinds of communications and interactions the child has with others, which, in turn, can shape the child's language-learning environment. In other words, gesture affects learning indirectly through its impact on the learning environment.

The third possibility is that increasing child gesture affects learning directly by having an impact on child cognition. Experimental research shows that children's actions can influence cognitive development (Campos et al., 2000; Campos, Bertenthal, & Kermoian, 1992; Rakison & Woodward, 2008), as can their gestures (Broaders et al., 2007; Cook et al., 2008; Goldin-Meadow et al., 2009). In our study, gesturing to the target pictures in the context of labeling may have focused children's attention to objects in the environment, the labels accompanying them, and/or object–word relations (Clark, 1978; Goldfield & Reznick, 1990; Werner & Kaplan, 1963). Gesturing also gives children practice in referring to objects, which could encourage them to seek out information relevant to word learning. Children's active engagement in the bidirectional labeling context when told to gesture may also enrich their understanding of the function of gesture, which could have beneficial consequences for children's vocabulary development (Csibra & Gergely, 2009; Tomasello, Carpenter, & Liszkowski, 2007; Woodward & Guajardo, 2002; Yoon, Johnson, & Csibra, 2008).

Future Directions

As discussed earlier, we found individual differences in the magnitude of our effects on child gesture, leading us to ask whether there are other factors, in addition to our experimental instructions, that could have influenced child gesture in our experiment. For example, individual caregivers may have

facilitated our intervention. In all conditions, 80% of caregivers observed the experimenter interact with the child during at least one session and thus might have, on their own, begun to employ those same behaviors in their everyday interactions with their children. Caregivers in the told-to-gesture group did, in fact, increase their own gestures, which could serve to reinforce gesturing as a means of communication and thus enhance the impact of our manipulation. One possibility for future studies is to enlist parents, either explicitly (as was done in Goodwyn et al.'s, 2000, study) or implicitly (as may have happened in our study) as "experimenters" to administer gesture interventions. Positive relations have been observed between caregiver gesture and infant gesture (Namy, Acredolo, & Goodwyn, 2000; Rowe & Goldin-Meadow, 2009a), and future research with larger samples is needed to deepen our understanding of the role that caregivers are able to play in increasing child gesture. Considering factors other than direct instruction that could influence child gesture may be useful in thinking about different ways to translate our findings into interventions.

In addition, future studies are needed to examine potential differences between our control conditions. In our Part 1 analyses, we predicted larger increases in gesture for children who were told to gesture (C&EG condition) than for children who were not told to gesture (EG and NG control conditions). Our findings supported this prediction. However, our small sample did not permit us to fully explore differences between the two control conditions. Studies with larger samples are needed to investigate whether seeing an experimenter gesture (our EG condition) itself has an effect on learning. For example, Capone and McGregor's (2005) study of children ages 2;3 to 2;6 suggests that seeing gesture can increase performance on a word-learning task. It is worth noting, however, that their effect may have been through child gesture because a number of children in the gesture input condition spontaneously imitated the experimenter's gesture (Capone, 2007; see also Cook & Goldin-Meadow, 2006, who found, in older children, that experimenter gesture led to increased child gesture, which, in turn, led to improved performance on a math task).

Lastly, our findings offer a potential intervention for children with speech delays and those at risk for low vocabulary. Children's early gesture predicts speech delay (Capone & McGregor, 2004; Sauer, Levine, & Goldin-Meadow, 2010; Thal & Tobias, 1994); our findings raise the possibility that gesture can be used as a tool to prevent or lessen those delays. Intervention could also be targeted to children from low-SES families who, as a group, have smaller vocabularies than those of children from high-SES families (Hoff, 2006). Vocabulary differences that children from high-SES versus low-SES families bring with them to school can be traced, in part, back to how the children gestured at age 1;2, which, in turn, can be traced back to how their parents gestured to their children (Rowe & Goldin-Meadow, 2009a). These findings suggest that one means for reducing disparities in vocabulary may be to manipulate early gesturing in parents and children.

ACKNOWLEDGMENTS

We thank A. Evenson and G. Spharim for help in coding, as well as S. Beilock, J. Iverson, K. Kinzler, and S. Levine for comments on earlier drafts.

FUNDING

The research was supported in part by a grant from the National Institute of Child Health and Human Development to Susan Goldin-Meadow.

REFERENCES

- Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development, 59*, 450–466.
- Anderson, M. J., & Legendre, P. (1999). An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *Journal of Statistical Computation and Simulation, 62*, 271–303.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.
- Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *The Journal of Economic Perspectives, 15*, 69–85.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols*. New York, NY: Academic.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., . . . Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*, 85–123.
- Broaders, S., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136*, 539–550.
- Campos, J., Anderson, D., Barbu-Roth, M., Hubbard, E., Hertenstein, M., & Witherington, D. (2000). Travel broadens the mind. *Infancy, 1*, 149–219.
- Campos, J., Bertenthal, B., & Kermoian, R. (1992). Early experience and emotional development: The emergence of wariness of heights. *Psychological Science, 3*, 61–64.
- Capone, N. (2007). Tapping toddlers' evolving semantic representation via gesture. *Journal of Speech, Language, and Hearing Research, 50*, 732–745.
- Capone, N., & McGregor, K. (2004). Gesture development: A review for clinical and research practices. *Journal of Speech, Language, and Hearing Research, 47*, 173–186.
- Capone, N., & McGregor, K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research, 48*, 1468–1480.
- Clark, R. (1978). The transition from action to gesture. In A. Lock (Ed.), *Action, gesture, and symbol: The emergence of language* (pp. 231–257). New York, NY: Academic.
- Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7*, 211–232.
- Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*, 1047–1058.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148–153.
- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers, 28*, 125–127.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, MD: Brookes.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., . . . Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA: Singular.
- Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language, 17*, 171–183.
- Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20*, 267–272.
- Goldin-Meadow, S., Goodrich, W., Sauer, E., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental Science, 10*, 778–785.
- Goldin-Meadow, S., & Mylander, C. (1984). Gestural communication in deaf children: The effects and noneffects of parental input on early language development. *Monographs of the Society for Research in Child Development, 49*(3/4, Serial No. 207).
- Golinkoff, R. M. (1986). 'I beg your pardon?': The preverbal negotiation of failed messages. *Journal of Child Language, 13*, 455–476.
- Good, P. (2000). *Permutation tests: A practical guide to resampling methods for testing hypotheses* (2nd ed.). New York, NY: Springer.
- Goodwyn, S. W., Acredolo, L. P., & Brown, C. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior, 24*, 81–103.

- Hayes, A. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*, 408–420.
- Hesterberg, T., Moore, D. S., Monaghan, S., Clipson, A., & Epstein, R. (2006). Bootstrap methods and permutation tests. In D. S. Moore & G. P. McCabe (Eds.), *Introduction to the practice of statistics* (Chapter 14). New York: Freeman.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*, 55–88.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248.
- Ingersoll, B., Lewis, E., & Kroman, E. (2007). Teaching the imitation and spontaneous use of descriptive gestures in young children with autism using a naturalistic behavioral intervention. *Journal of Autism and Developmental Disorders*, *37*, 1446–1456.
- Iverson, J., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*, 367–371.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Psychology Press.
- MacKinnon, D., Lockwood, C., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*, 99–128.
- Masur, E. (1982). Mothers' responses to infants' object-related gestures: Influences on lexical development. *Journal of Child Language*, *9*, 23–30.
- Namy, L. L., Acredolo, L., & Goodwyn, S. (2000). Verbal labels and gestural routines in parental communication with young children. *Journal of Nonverbal Behavior*, *24*, 63–79.
- O'Neill, D. K., Pearce, M. J., & Pick, J. L. (2004). Relation between early narrative and later mathematical ability. *First Language*, *24*, 149–183.
- Ozcaliskan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, *96*, B101–B113.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717–731.
- Rakison, D., & Woodward, A. (2008). New perspectives on the effects of action on perceptual and cognitive development. *Developmental Psychology*, *44*, 1209–1213.
- Rowe, M. L., & Goldin-Meadow, S. (2009a). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, *323*, 951–953.
- Rowe, M., & Goldin-Meadow, S. (2009b). Early gesture selectively predicts later language learning. *Developmental Science*, *12*, 182–187.
- Sauer, E., Levine, S., & Goldin-Meadow, S. (2010). Early gesture predicts language delay in children with pre- or perinatal brain lesions. *Child Development*, *81*, 528–539.
- Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*, 13–19.
- Thal, D., & Tobias, S. (1994). Relationships between language and gesture in normally developing and late-talking toddlers. *Journal of Speech and Hearing Research*, *37*, 157–170.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*, 705–722.
- Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought*. New York, NY: Wiley.
- Woodward, A., & Guajardo, J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*, 1061–1084.
- Yoder, P. J., & Warren, S. F. (2002). Effects of prelinguistic milieu teaching and parent responsivity education on dyads involving children with intellectual disabilities. *Journal of Speech, Language, and Hearing Research*, *45*, 1158–1174.
- Yoon, J., Johnson, M., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 13690–13695.