Is She a Good Teacher? Children Learn to use Meaningful Gesture as a Marker of a Good Informant

Elizabeth M. Wakefield (ewakefield1@luc.edu)
Department of Psychology, Loyola University Chicago
1032 W. Sheridan Road, Chicago, IL 60660 USA

Eliza L. Congdon (elc6@williams.edu)
Department of Psychology, Williams College
25 Stetson Court., Williamstown, MA 01267 USA

Miriam A. Novack (miriam.novack@northwestern.edu)
Department of Psychology, Northwestern University
Swift Hall 102, 2029 Sheridan Road, Evanston, IL 60208 USA

Lauren H. Howard (lauren.howard@fandm.edu)
Department of Psychology, Franklin & Marshall College
P.O. Box 3003, Lancaster, PA 17604 USA

Abstract

To learn from others, children rely on cues (e.g., familiarity) to infer who will provide useful information. We extend this research to ask whether children will use an informant's inclination to gesture as a marker of whether they are a good person to learn from. Children (*N*=459, ages 4-12 years) watched videos in which actresses made statements accompanied by meaningful iconic gestures, beat gestures, or no gestures. After each trial, children were asked "Who do you think would be a good teacher?" (good teacher- experimental condition) or "Who do you think would be a good friend?" (good friend-control condition). Results show children *do* believe that someone who produces iconic gesture would make a good teacher over someone who does not, but this is only later in childhood and only if a child has the propensity to see gesture as meaningful. The same effects were not found in the good-friend condition.

Keywords: gesture; learning; informants

Introduction

Humans are fundamentally social creatures: we rely on those around us when we need to gather information. When we want to learn from others, we must decide *who* to ask, making choices about which people will be the most likely to provide us with useful information. Young children are especially sensitive to signs or signals that a social partner might be a good informant. For example, children prefer to learn information from individuals who appear knowledgeable, confident, and nice, versus ignorant, timid, and mean (see Harris, 2012). They are also more likely to trust information provided by individuals who they are familiar with (Corriveau & Harris, 2009) and those who are members of their social in-group (i.e., more "like them", Aboud, 2003). Here, we extend this prior research on the cues children use to identify a good informant by asking whether children also use an informant's communicative tendencies when choosing who to learn from. We focus specifically on an informant's inclination to gesture.

Gesture, in both conversational and instructional contexts has been found to support communication (e.g., Hostetter, 2011) and learning (e.g., Goldin-Meadow, 2015). Decades of research show that children are more likely to learn, generalize and retain knowledge when a teacher accompanies spoken instruction with gesture, compared to when she explains concepts through spoken instruction alone (Goldin-Meadow, 2015). We see this beneficial effect of gesture across development: gesture instruction improves 2-3 year-olds' understanding of the functions of novel toys (M. Novack, Goldin-Meadow, & Woodward, 2015), 3-4 year-olds' knowledge of symmetry (Valenzeno, Alibali, & Klatzky, 2003), 5-6 year-olds' ability to solve Piagetian conservation problems (Ping & Goldin-Meadow, 2008), and 8-10 year-olds' understanding of mathematical equivalence (Congdon et al., 2017). Instructional gestures are often *iconic* gestures, those that are representational of ideas or problem-solving strategies through visuo-spatial representations (McNeill, 1992). To illustrate how iconic gestures are used in instruction, consider the concept of liquid conservation – the idea that the amount of liquid is conserved across containers of various shapes. One strategy that teachers can use to help children understand conservation is to emphasize that a container has multiple dimensions. Ping and Goldin-Meadow (2008) manipulated whether a teacher expressed this concept through spoken instruction alone, "One of the glasses is taller and the other one is shorter, but the shorter glass is wider and the taller glass is skinnier. So it makes up for it," or whether she accompanied this spoken instruction with gestures, where the hands indicated the relative heights and widths of containers. Children learned more from instruction with gesture than from spoken instruction alone.

Although we know children learn more when a teacher uses gesture than when she does not, we do not know whether children are *aware* that a teacher who gestures is a good person to learn from. In naturalistic settings not all teachers gesture to the same extent, although gesture use is a characteristic of good teaching (Richland, 2015). Therefore, if a child wants to learn information, it might benefit them to choose an informant who *does* tend to use her hands when speaking. In the present study, we ask precisely this question. That is, we ask whether children identify a speaker as a good informant or teacher if that person produces meaningful, iconic gestures when speaking.

Finally, we suggest that individual differences in a child's general ability to recognize gestures as meaningful and communicative likely plays a role in whether they select an iconic gesturer as a good teacher. Previous work shows that while young children *can* see gesture as meaningful, this ability is fragile and emerges slowly in development. Using stimuli created by Novack, Wakefield & Goldin-Meadow (2016) that provided children with very few contextual cues, Wakefield, Novack and Goldin-Meadow (2017) asked how the ability to see movement as 'gesture' develops across early childhood. The researchers showed 4- to 9-year-old children videos in which an actor gestures the movement of placing balls into boxes, without physically touching any objects. Children were asked to describe the video and their responses were coded for whether they interpreted the 'empty-handed' movements as meaningful gestures (e.g., stating that the actor was *showing* how to put the balls into boxes) versus as 'meaningless movement' (e.g., stating that the actor was moving her hands back and forth near some balls). Results showed that older children were more likely to describe these 'empty-handed' movements as 'gesture' whereas younger children, were more likely to describe the movements as meaningless, however at all ages there were some children who described the movement as gesture. We anticipate that, along with age, children's individual propensity to see meaning in gesture-like movements may also impact their likelihood of identifying iconic gesture production as a cue of a good informant.

In the present study, we use an alternative-forced choice paradigm in which children are shown three pairs of short videos and are asked to choose which person in each pair of videos would be a good teacher (experimental condition) or which speaker would be a good friend (control condition). The three video pairs contrast 1) a speaker who produces iconic gestures with a speaker who does not gesture at all, 2) a speaker who produces iconic gestures with one who produces beat gestures, and 3) a speaker who produces no gesture with one who produces beat gestures. Beat gestures are gestures that emphasize parts of a spoken sentence, but do not express semantic information. Including a speaker who produces beat gestures controls for the possibility that children would see a person who gestures *at all* as a good informant, as we are specifically interested in children's attention to iconic, semantically meaningful gestures. In addition, the good friend control condition will help to determine the specificity of any reported effects. If, as predicted, children select a speaker who produces iconic gesture as a better teacher but not as a better friend, we can rule out the hypothesis that general likeability for someone who produces iconic gestures is driving the effect in our good teacher experimental condition. After the forced-choice paradigm, all participants are presented with a video drawn from Novack et al. (2016) as a measure of their propensity to see gesture as meaningful outside of the context of our video pairs. We predict that when children are making judgements about who would be a better informant, they will be more likely to choose the speaker who produces iconic gesture, but that this may depend on both a child's age and their propensity to see meaning in hand movements.

Method

Participants

Our final sample included data collected from 459 children between the ages of 4 and 12 years (252 females, 207 males) at a large science museum, (4 yrs: n=47; 5 yrs: n=60; 6 yrs: n=60; 7 yrs: n=59; 8 yrs: n=63; 9 yrs: n=65; 10 yrs: n=36; 11 yrs: n=43; 12 yrs: n=26). Parents provided informed consent, and children provided verbal assent. An additional 17 children were excluded from analyses for not completing the study (n=10), receiving parental assistance (n=5), or because they were outside the target age range (n=2). The task took 3-5 minutes to complete and children were given stickers for participating.

Stimuli

Videos. Eighteen, 6-second videos were chosen from a set of 162 videos for the present study, based on results of a norming study[1]. Each video showed one of six actresses from the waist up, dressed in a solid-colored shirt. The set of 18 videos included 3 videos per actress: a no gesture video, an iconic gesture video, and a beat gesture video. In all three video types, the actress made the same basic statement (e.g., "In fall, the leaves *drop* to the ground and I *build them up* in a pile"; see Figure 1). In no gesture videos, the actress kept her hands to her side. In iconic gesture videos, the actress produced two iconic gestures accompanying her statement. In the example above, the gesture produced simultaneously

[1] To ensure that any differences in participants' video choices would be based on the experimental manipulation, a norming study was first conducted. Ninety-four English-speaking adults participated in the study via Amazon Mechanical Turk. Participants watched 9 videos. Each video showed a different actress making a statement. Participants rated the actresses on a number of dimensions using a 7-point Likert scale (e.g., how friendly, approachable, attractive, intelligent actresses were). The videos were drawn from a set of 162 videos: 9 actresses each recorded 18 videos, which included 6 unique statements produced with each of the three gesture types (no gesture, beat gesture, iconic gesture). Based on these Likert responses we then paired actresses together who received similar ratings on all measures – 6 actresses were selected to create 3 actress pairs.

Figure 1: Examples of the three trial types presented during the alternative-forced-choice task.

with *drop* depicted the downward motion of leaves (moving the hands downwards while wiggling the fingers), and the gesture produced simultaneously with *build them* depicted building a pile of leaves (moving the hands towards each other and upward; palms-in). In beat gesture videos, the actress produced two beat gestures accompanying her statement. These rhythmic downwards hand motions were produced at the same points in the statement as the iconic gestures were performed. Each actress stated the same fact in her three videos, and the same audio file was used for all three videos, ensuring no differences in emphasis or inflection. Each actress made a different statement (see Figure 1).

For the study, videos were presented in pairs during an alternative-forced choice (AFC) task. Children saw three AFC trials, one of each of three trial types (no gesture vs. iconic gesture; no gesture vs. beat gesture; beat gesture vs. iconic gesture; see Figure 1). The same actresses were always paired together for an AFC trial, and a different actress pair appeared in each of the three trials, such that children never saw the same actress twice. Across participants, we varied which actress pair was used for each trial type, and within the pair, we varied which video type was shown for each actress. The order in which children saw the three trial types was also varied.

Empty-handed Movement Video. A 10-second video was drawn from a stimulus set created by Novack et al. (2016). The video showed the torso of a woman in front of a table with four balls (two orange and two blue) and two boxes (one orange and one blue) in front of her. In the video, the woman moves her hand over the inner blue ball with her left hand, and then over the blue box on her left; then she moves her hand over the inner orange ball with her right hand and then over the orange box on her right. These empty-handed movements are repeated with the outer balls. For additional details, see Novack et al. (2016).

Procedure

Children were invited to participate in the study while they were visiting a science museum. Those who agreed and whose parents signed a consent form were seated at a small table next to an experimenter, facing a wall to decrease distractions. Children wore headphones during the study to ensure the audio was clear. At the beginning of the study, children were told they would watch two short videos in which two different people would each tell them something. Depending on the condition children were randomly assigned to, they were told that after watching the videos they would be asked which of the people in the videos seems like they would be better at giving good information (good teacher-experimental condition) or which of the people in the videos seems friendlier (good friend-control condition).

The experimenter then played the first set of videos. After the videos ended, the experimenter advanced to a slide showing a picture of each actress from the videos (See Figure 1). She then prompted the child to answer, "Who do you think would be a good teacher" (good teacher-experimental condition) or "Who do you think would be a good friend?" (good friend-control condition). If the child did not respond, they were re-prompted, "If you had to pick one person that would be a good friend/good teacher, who would you pick?" This procedure was repeated for two additional video pairs, such that children made a selection for each type of video pair (no gesture vs. beat gesture; no gesture vs. iconic gesture; beat gesture vs. iconic gesture). The order of trials was randomized across participants.

After completing the three AFC trials, children were told there was one more video for them to watch and that they should pay close attention. The experimenter then played the empty-handed movement video and asked the child, "What happened in the movie?". Participants' responses were audio recorded for later transcription. Children were then thanked for their participation and offered stickers.

Coding empty-handed movement video responses. Participants' responses to the prompt, "What happened in the movie?" regarding the final empty-handed video were classified into categories based on the coding scheme used by Wakefield and colleagues (2017). The codes are described below:

(1) *Representational Response:* The movie is described in terms of movements representing (but not actually completing) external goals (e.g., "She was pretending to put the balls in different boxes").

(2) *Non-Representational Response:* Non-representational responses could be within three categories: external goal or movement-based goal responses, and other.

a. External Goal Responses: The movie is described in terms of actions completed on objects (e.g., "They put the balls in the same color box").

b. Movement-Based Goal Responses: The movie is described in terms of low-level spatiotemporal movements without mentioning a higher-level goal—the description is focused on the movement of the hands themselves (e.g., "he mostly only moved his hands around").

c. Other: The movie is described (a) without mentioning movement (e.g., "There were balls and they were different colors and they were blue and orange"), or (b) mentioning movement, but too ambiguous to assign a goal-oriented code (e.g., "someone doing something with balls").

Two researchers independently assigned a single code to all responses. Coders were blind to the condition and age of each participant. For the purpose of analysis, we considered whether children gave a representational goal response versus a non-representational goal response (collapsing across the external goal, movement-based goal, and other responses), and reliability was performed on this level of analysis. Coders agreed on 442 of 459 trials (96.0%), $\kappa = 0.95$. Any disagreements were discussed and resolved.

Results

Before exploring our main question, we consider the results of the empty-handed movement video question. Table 1 shows that overall, the proportion of children in each age group who gave a representational response increased with age, replicating Wakefield et al. (2017). A logistic regression confirmed that child age significantly predicted their likelihood to provide a representational response (0,1) to the empty-handed movement video ($\beta = 0.25$, $SE = 0.04$, $t = 5.02$, $p < .001$).

Table 1: Propensity to Interpret Gesture as Meaningful Across Development

Age (yrs)	Proportion of Children who Gave Representational Responses
4	0.07 (3 of 44)
5	0.15 (8 of 52)
6	0.33 (15 of 45)
7	0.64 (23 of 36)
8	0.70 (26 of 37)
9	0.81 (29 or 36)
10	0.71 (15 of 21)
11	0.72 (18 of 25)
12	0.63 (10 of 16)

Our main question was whether children are more likely to think that someone who produces semantically meaningful, iconic gesture while speaking will be a better informant or teacher than someone who produces a gesture that does not express information (a beat gesture) or someone who does not use gesture while speaking. To address this question, we considered two of the three trials from each participant: those in which one of the speakers produced iconic gesture and the other produced either no gesture or beat gestures. On average, collapsing across age, children performed similarly in both conditions, choosing the speaker who used iconic gestures about half the time in the good teacher condition ($M = 0.55$, $SE = 0.03$) and in the good friend condition ($M = 0.52$, $SE = 0.03$). However, we had anticipated that the age of a child and the propensity of that child to interpret gesture as meaningful might play a role in whether they used gesture as a cue for choosing a good informant. Therefore, we considered a mixed-effects logistic model that included not only condition (good teacher versus good friend), trial type (iconic gesture vs. no gesture; iconic gesture vs. beat gesture), age as a continuous measure, and a random effect of participant, but also a 3-way interaction between condition, age, and propensity to see gesture as meaningful (0, 1). The model revealed no main effect of condition ($\beta = 0.61$, $SE = 0.63$, $z = 0.96$, $p = .34$), trial type ($\beta = 0.01$, $SE = 0.15$, $z = 0.05$, $p = .96$), age ($\beta = 0.08$, $SE = 0.06$, $z = 1.34$, $p = .18$), or propensity to see gesture as meaningful ($\beta = 0.50$, $SE = 0.93$, $z = 0.53$, $p = .59$). However, an analysis of variance of the complex model revealed that there *was* a significant 3-way interaction between condition, age, and propensity to see gesture as meaningful, $\chi^2(1) = 3.93$, $p < .05$. We separated children based on condition, and found that this interaction was driven by differences between the good teacher and good friend conditions: Whereas a 2-way interaction of age and propensity to see gesture as meaningful predicted choice of the speaker who used iconic gesture for children in the good teacher condition ($\chi^2(1) = 5.16$, $p < .05$), there was not a significant interaction between these factors for children in the good friend condition ($\chi^2(1) = 0.14$, $p = .70$; Figure 2a).

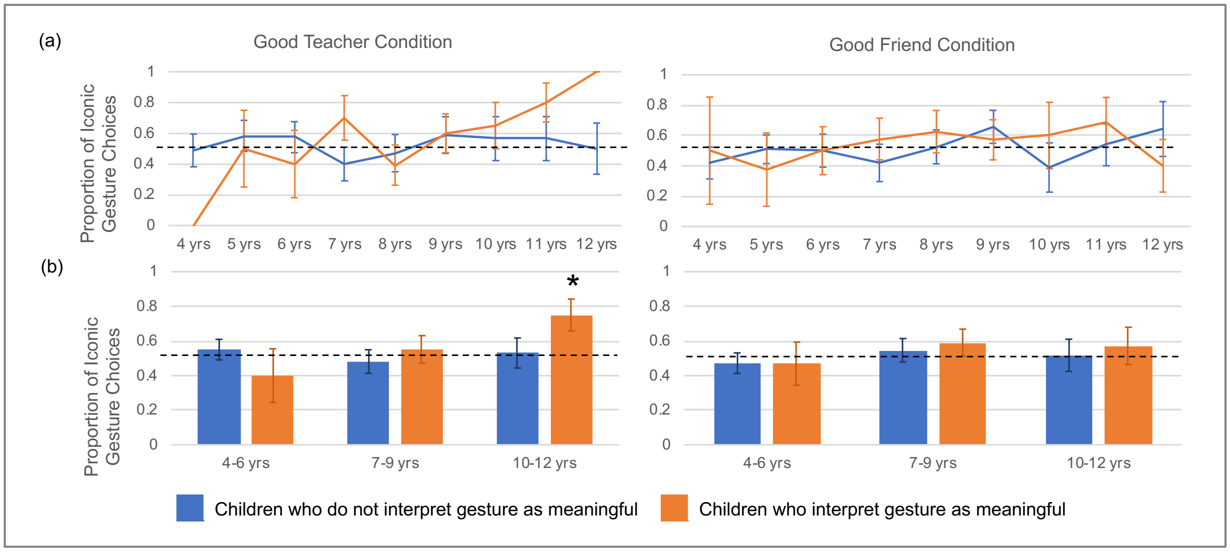

Figure 2: Proportion of iconic gesture speaker choice separated by condition, whether children interpreted gesture as meaningful in the empty-handed movement video, and age – (a) year and (b) binned-age-group. Error bars represent the standard errors of the proportions, calculated by the formula √[p(1-p)/n].

The results suggest a developmental change, in which a child's propensity to see gesture as meaningful increased the likelihood that they chose the speaker who used iconic gesture as a better teacher, but only after a certain age. However, this does not take into account the fact that choices are made within the context of an AFC task, where chance of making a random choice is 50%. We therefore also considered children's selections relative to chance performance. We separated children in the good teacher condition into age groups (4-6, 7-9, 10-12; see Figure 2b, children in the good friend condition are also binned for the sake of visual comparison). The choice to bin children into wider age groups was made because, as found by Wakefield and colleagues (2017), the number of children who interpret gesture as meaningful is relatively small, especially in the younger age groups, which makes an analysis in which children are separated by year-of-age underpowered. Wakefield and colleagues found that context began to affect children's likelihood to see gesture as meaningful around the age of 6-years, thus, we considered a 3-way split of the data that included this as one of the break-points. To obtain one test value per child, we averaged children's responses to the iconic gesture vs. no gesture and iconic gesture vs. beat gesture trials. This was justifiable, as we saw no evidence that children responded differently to these trials in the previous analysis.

We conducted 6 one-sample t-tests, asking whether the proportion of choice of the speaker who used iconic gesture as the better teacher was greater than chance (0.5), separating children by age group and whether they interpreted gesture as meaningful. To correct for multiple-comparisons, an alpha of $p<.008$ was considered statistically significant based on a conservative Bonferroni correction. In the youngest two age groups, we found no evidence that likelihood of choosing the speaker who used iconic gesture as the better teacher was above chance (4-6 year-olds who do not see gesture as meaningful: $t(69) = 1.04$, $p=.30$; 4-6 year-olds who do see gesture as meaningful: $t(9) = -1.50$, $p=.17$; 7-9 year-olds who do not see gesture as meaningful: $t(53) = -0.35$, $p=.73$; 7-9 year-olds who do see gesture as meaningful $t(38) = 0.78$, $p=.44$). However, in the 10-12 year old groups, children who interpret gesture as meaningful significantly choose the speaker who used iconic gesture above chance ($t(21) = 3.48$, $p=.002$). This was not the case for 10-12 year old children who do not see gesture as meaningful ($t(32) = 0.47$, $p=.64$), Together, these results suggest that by 10-12 years, children who have a propensity to see gesture as meaningful also believe that a speaker who produced meaningful gestures will make a good informant.

Discussion

In the present study, we explored whether children use iconic gesture as a cue that an adult would be a good informant. Our results show that children *do* believe that someone who produces iconic gesture while speaking would make a good teacher, but this does not occur until relatively late in childhood, and is only likely if a child has the propensity to see gesture as meaningful. Importantly, we found that children do not select a gesturer as a good teacher simply because they like them more: when children were given the prompt to think about which speaker would be a better friend, children showed no preferences for the speaker who used iconic gesture.

Compared to other cues that children use when evaluating potential informants, such as familiarity (Corriveau & Harris, 2009) or group membership (Aboud, 2003), which appear to emerge as early as three years of age (e.g., Koenig & Harris, 2005), the use of iconic gesture as a cue develops later in childhood. Why is this the case? One possibility is that gesture is more subtle than other cues: we are often explicitly unaware of when others gesture, even if gestures influence our understanding, impressions, and judgements of other speakers (Alibali et al., 1997). A second possibility is that the ability to use gesture as a marker of a good informant may rely on a child's representational capacities, which show a protracted development across childhood (Richland, Morrison, & Holyoak, 2006). We know that the ability to process relationally complex information is still developing before the age of 10 years (Richland et al., 2006), partially because of the protracted development of working memory (Morrison, Doumas, & Richland, 2011). Finally, it is worth noting that our use of gesture as a cue was relatively subtle in this paradigm. The gesture was not strictly necessary to understand the spoken language and there were only two instances of gesture within each video. Thus, the current study may serve as a conservative estimate of children's ability to use gesture as a cue to a good informant. Nevertheless, our results do suggest that the ability to see gestures as meaningful is necessary, but not sufficient, for selecting a person who produces iconic gestures as a good informant. It may be the case that both development of representational processes *and* experiences learning from meaningful gestures (two things that are often but not always correlated) are necessary to explicitly identify iconic gesture as a potentially useful source of information.

Our findings also have implications for how children learn from gesture. We know that teachers use gesture in the classroom when explaining concepts to students (e.g., Alibali & Nathan, 2012), especially if students appear to be struggling to grasp the concept (Alibali et al., 2013). For example, iconic gesture can visually depict how a scale would shift, when an elementary school teacher is explaining the concept of a balance. We also know that on average, seeing a teacher gesture leads to improved learning outcomes for students (e.g., Congdon et al., 2017; Singer & Goldin-Meadow, 2005; Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018). Researchers have begun to unpack *why* gesture helps learners, suggesting that gesture helps children because it allows them to follow along with and more fully understand a teacher's spoken instruction (Congdon et al., 2017; Wakefield et al., 2018), decreases working memory load (Cook, Yip, & Goldin-Meadow, 2012), and involves the motor system in the learning process (Wakefield, Congdon, Novack, Goldin-Meadow, & James, 2019). However, in addition to these general cognitive influences, a child's tendency to think of their teacher as a good informant may boost gesture's impact on early learning. In other words, our results suggest that, for older children, gesture might improve learning not only because it expresses useful information, but because children show a desire to learn from informants who *use* gesture. Future research on learning from gesture should consider the impact of this type of individual difference.

The current study raises several exciting mechanistic questions for future research. First, we do not yet know whether children view someone's propensity to gesture as a stable trait, similar to personality or intelligence. Although research has shown that gesture production rates do vary somewhat within an individual based on external environmental constraints, such as whether a communicative partner is knowledgeable in a given context (e.g., Alibali & Nathan, 2007; Hilliard, O'Neal, Plumert, & Cook, 2015), there is also good evidence that individual differences in gesture rates are relatively stable across time and across similar contexts (e.g., Chu, Meyer, Foulkes, & Kita, 2014; Hostetter & Potthoff, 2012). Future work could directly test whether children have any sense of this stability, and whether that measure partially or fully mediates the reported relation between age and propensity to select the iconic gesturer as a good teacher.

In addition, as mentioned earlier, our manipulation in the current study was subtle: we used iconic gestures that were complementary to speech, but did not add additional information not found in speech. Perhaps we would see a more robust effect if gestures provided supplemental information *not* found in the narrative, or information that disambiguated spoken information. Again, this is a question that could be addressed in future work.

Finally, it could be interesting to explore what additional traits children might take into consideration when choosing a gesturer as a worthy informant. For example, would child gesturers be considered more friendly, but less likely to be a good teacher, than adult gesturers? Would we expect other communicative qualities, such as accent or clarity, to override children's preferences for gesturers as teachers?

Open questions aside, we show that even though gesture is a subtle cue compared to those that are more traditionally considered in the informant literature, by age ten, children who see meaning in gesture can make use of this cue when deciding who to learn from. The present work makes a novel, important contribution to our understanding of how children make choices about informants.

Acknowledgments

Funding for this study was provided by Loyola University Chicago to E. Wakefield. The authors would like to thank Madeline Jurcev, Allison Haussler, Nema Shareef and Sasha Stojanovich for their help with data collection and coding, and undergraduate research assistants at Williams College for volunteering to be filmed for stimuli. The authors would also like to thank children and families who participated in this study while visiting the Museum of Science and Industry in Chicago, and the museum for graciously allowing external researchers to conduct studies that would not be possible in other settings.

References

Aboud, F. E. (2003). The formation of in-group favoritism and out-group prejudice in young children: Are they distinct attitudes? *Developmental Psychology, 39*, 48-60.

Alibali, M. W., & Nathan, M. J. (2007). Teachers' gestures as a means of scaffolding students' understanding:

Evidence from an early algebra lesson. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video Research in the Learning Sciences*. New York, NY: Lawrence Erlbaum Associates Inc.

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247-286.

Alibali, M. W., Nathan, M. J., Church, R. B., Wolfgram, M. S., Kim, S., & Knuth, E. J. (2013). Teachers' gestures and speech in mathematics lessons: Forging common ground by resolving trouble spots. *Mathematics Education, 45*, 425-440.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General, 143*, 694-709.

Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O'Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction, 50*, 65-74.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes, 27*, 594-610.

Corriveau, K. H., & Harris, P. L. (2009). Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Developmental Science, 12*, 188-193.

Goldin-Meadow, S. (2015). From action to abstraction: Gesture as a mechanism of change. *Developmental Review, 38*, 167-184.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Boston, PA: Belknap Press of Harvard University Press.

Hilliard, C., O'Neal, E., Plumert, J., & Cook, S. W. (2015). Mothers modulate their gesture independently of their speech. *Cognition, 140*, 89-94.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*, 297-315.

Hostetter, A. B., & Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture, 12*, 62-83.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261-1277.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: The University of Chicago Press.

Morrison, R. G., Doumas, L. A. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. *Developmental Science, 14*, 516-529.

Novack, M., Goldin-Meadow, S., & Woodward, A. (2015). Learning from gesture: How early does it happen? *Cognition, 142*, 138-147.

Novack, M. A., Wakefield, E. M., & Goldin-Meadow, S. (2016). What makes a movement a gesture? *Cognition, 146*, 339-348.

Ping, R. M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology, 44*, 1277-1287.

Richland, L. E. (2015). Linking gestures: Cross-cultural variation during instructional analogies. *Cognition and Instruction, 33*, 295-321.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*, 249-273.

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*, 85-89.

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*, 187-204.

Wakefield, E. M., Congdon, E. L., Novack, M. A., Goldin-Meadow, S., & James, K. H. (2019). Learning math by hand: The neural effects of gesture-based instruction in 8-year-old children. *Attention, Perception, & Psychophysics*, 1-11.

Wakefield, E. M., Novack, M., & Goldin-Meadow, S. (2017). Unpacking the ontogeny of gesture understanding: How movement becomes meaningful across development. *Child Development*.

Wakefield, E. M., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*.