Developmental Psychology
2004, Vol. 40, No. 6, 1059–1067

0012-1649/04/$12.00 DOI: 10.1037/0012-1649.40.6.1059

More Gestures Than Answers: Children Learning About Balance

Karen J. Pine and Nicola Lufkin
University of Hertfordshire

David Messer
South Bank University

This research extends the range of domains within which children's gestures are found to play an important role in learning. The study involves children learning about balance, and the authors locate children's gestures within a relevant model of cognitive development—the representational redescription model (A. Karmiloff-Smith, 1992). The speech and gestures of children explaining a balance task were examined. Approximately one third of the children expressed one idea in speech and another in gesture. These children made significantly more learning gains than children whose gestures and speech matched. Children's gestures were an indicator, at pretest, of readiness to learn and of cognitive gains. The authors conclude that children's gestures provide crucial insight into their cognitive state and illuminate the process of learning and representational change.

Developmental psychologists, when assessing children's knowledge, have long recognized the importance of listening carefully to all that a child says. Methods of statistical and discourse analysis have been refined to ensure accuracy and reliability when coding children's speech. However, when a child is asked to explain a problem-solving task, there is no way of ensuring that the child's explanation is a reliable indicator of all that the child knows. In this article, we argue that although precision in attending to children's speech is important, children's articulated speech is just one component of the communicative channel. Children also convey a substantial proportion of their knowledge through another mode—the hand gestures that accompany their speech. Attending to children's gestures as well as their speech, we argue, offers an additional window into the mind of the child and more accuracy when assessing children's understanding (see also Goldin-Meadow & Alibali, 2002).

The theoretical and empirical motivation for this article comes from work in two areas that address children's knowledge representations. The first comes from research investigating gesture production in children, particularly in problem-solving contexts such as conservation and in understanding mathematical equivalence (e.g., Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Perry, Church, & Goldin-Meadow, 1988). These studies provide the methodological, as well as some theoretical, impetus for this investigation into children's gestures. A unique contribution of the study presented here is that it extends previous empirical work by exploring children's gestures within a domain not previously investigated, using a balance beam task. The second body of work of relevance here focuses on Karmiloff-Smith's (1992) representational redescription (RR) model. This model is invoked because it provides a definitive and unique account of both nonverbal and verbal representations in development and has also been applied to children acquiring a concept about balance. Furthermore, it accounts for how children can have more knowledge than they are able to talk about and provides a developmental account of nonverbal knowledge. Therefore, it offers a very plausible theoretical framework within which to understand children's gestures. This study is the first to locate gestural knowledge of the balance beam task within the RR model.

Gestures as a Reflection of Knowledge

When children, and adults, are asked to explain something, they frequently gesture with their hands. These gestures are usually spontaneous and produced without conscious awareness. Research is increasingly focusing on what children's gestures can tell us about their thoughts because children may not always accurately explain what they know. One reason for this may be that the child lacks the necessary linguistic competence to produce an explanation. This is particularly true of children with language impairments, who have been found to express more sophisticated knowledge in gesture than in speech (Evans, Alibali, & McNeil, 2001). In contrast, children without any linguistic impairment may nonetheless possess knowledge that they cannot express verbally, that is, knowledge that is implicit or encoded in a visual or spatial format (Goldin-Meadow & Alibali, 1994; Karmiloff-Smith, 1992); it is this knowledge that may be communicated in gestural form.

Research has shown that it is possible for experienced coders to assign meaning to the gestures that children produce when solving problems. Even when making independent assessments, observers have been found to be able to interpret reliably the meaning of a gesture in the same way. This indicates that there is consistency in the way that different children gesture when they are given the same task to explain.

It is interesting to note that children have been found to sometimes convey different information in their gestures compared with that expressed in their spoken explanations (Alibali & Goldin-Meadow, 1993; Garber, 1998). For example, on the Piagetian

---

Karen J. Pine and Nicola Lufkin, Department of Psychology, University of Hertfordshire, Hertfordshire, England; David Messer, Department of Psychology, South Bank University, London, England.

This research was supported by Grant R000 22 3855 from the Economic and Social Research Council to Karen J. Pine and David Messer.

Correspondence concerning this article should be addressed to Karen J. Pine, Department of Psychology, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB, England. E-mail: K.J.Pine@herts.ac.uk

conservation task of liquid quantity, height and width are the two key variables. Some children have been found to talk about only one variable, the height of the containers, but to accompany this with a gesture that indicates the other variable, width (Church & Goldin-Meadow, 1986). These gesture–speech mismatches have been found to indicate children who are in transition between one knowledge state and another, and there is empirical evidence to suggest that this indicates a readiness to learn.

Goldin-Meadow and her colleagues (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Goldin-Meadow & Alibali, 2002) proposed that gesture is a reliable index of transitional knowledge and that mismatches between gestures and speech reflect openness to instruction. Children whose spoken explanations depict different information from that conveyed by their hand gestures have been found to be significantly more likely to profit from instruction than children whose speech and gestures match (see Goldin-Meadow & Alibali, 2002, for a summary). A child who produces two ideas concurrently, one in speech and another in gesture, is indexing cognitive instability or variability. Cognitive variability has been shown to be a reliable indicator that the knowledge system is in a state of transition and is ready to undergo change (Siegler, 1996). When Perry et al. (1988) presented children with mathematical equivalence problems and asked those children who were failing to explain how they arrived at their answers, some children displayed some understanding of the nature of equivalence in their gestures, although not in their spoken explanations. These children, it was later found, were more likely to improve after instruction than children who did not produce mismatches. Similarly, Church and Goldin-Meadow (1986) found that younger children who produced gesture–speech mismatches on a conservation task were also more open to instruction. The phenomenon, Goldin-Meadow and Alibali (2002) concluded, “is not tied to one age, nor to one task” (p. 83).

Therefore, spontaneously produced gestures are more than just paralinguistic features, and the evidence from these training studies suggests that their role is not merely a communicative one. Gestures are clearly an integral component of the cognitive process and can have an effect on thought itself (McNeill, 1992). They have the potential to give insights into the multidimensional nature of the child’s thinking processes and may reflect knowledge that the child has but that is not yet well developed enough to express verbally. Further support for this notion comes from the finding that blind babies gesture from birth, and the gestures they produce resemble, in both form and content, those produced by sighted children (Iverson & Goldin-Meadow, 1997). As these children have never seen other people gesture, this implies that gestures are important to the speaker as well as to the listener. In sighted people, gestures are not always produced when others are watching; for example, adults frequently gesture while speaking on the telephone. How gestures help the speaker is explored further when we ask whether gesturing not only reflects but also enhances cognitive capacities. First, however, we introduce a new problem-solving domain within which to explore these issues and extend previous findings.

Extending the Study of Children’s Gestures to a New Domain

Balancing tasks have frequently been used as a testing ground for theories of cognitive development (e.g., Halford, 1993; Inhelder & Piaget, 1958; Karmiloff-Smith, 1992; Karmiloff-Smith & Inhelder, 1974; Pine & Messer, 1998, 1999, 2000; Siegler, 1976). Having used the balance beam task as a means of investigating children’s implicit and explicit representations in previous work, we are able to draw on this body of work to extend the range of domains within which gesture is studied. To date, empirical studies have been conducted looking at children’s gestures when carrying out a conservation task (Alibali, Flevares, & Goldin-Meadow, 1997; Alibali, Kita, & Young, 2000; Breckinridge-Church, Kelly, & Lynch, 2000; Church & Goldin-Meadow, 1986; Goldin-Meadow & Momeni-Sandhofer, 1999) or when solving mathematical problems (Alibali, 1995; Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999; Alibali & Goldin-Meadow, 1993; Garber, Alibali, & Goldin-Meadow, 1998; Perry, Breckinridge-Church, & Goldin-Meadow, 1992), with a few studies focusing on children’s counting (Alibali & DiRusso, 1999; Graham, 1999) and the Tower of Hanoi problem (Garber, 1998; Garber & Goldin-Meadow, 2002). Because balancing studies have consistently and for many years provided a reliable test bed for theories about cognition, it is also informative to understand the role of children’s gestures within this domain and to integrate them into a theoretical framework of how cognitive change occurs.

The balance beam task used in our studies involves the child balancing wooden beams (similar to a wooden ruler) on a simple support, or fulcrum, and was originally reported in Karmiloff-Smith and Inhelder (1974). The beams are either symmetrical, having a wooden block at each end, or asymmetrical, with a block at one end only. The symmetrical beams balance by placing them onto the fulcrum at their midpoint. The asymmetrical beams have to be placed off-center onto the fulcrum to balance. Hence, two variables are involved in completing this task successfully—weight and distance—because distance from the fulcrum has to be adjusted according to the amount of weight at each end of the beam.

It has been particularly striking to us as investigators how children seem compelled to gesture with their hands when explaining the balance beam task. Also notable is the fact that certain gestures consistently appear to accompany descriptions of particular aspects of the task. For example, when talking about distance or length, children invariably gesture with a flat hand, palm down, and a sweeping movement from the fulcrum outward. Weight, or “heaviness,” is usually indicated by closing all the extended fingers in a downward pincer movement while making an up and down movement of the hand over one end of the beam. Children explaining the center position of a beam often point to the middle with the index fingers of both hands close together. The aim of this research is to establish, by examining more closely the videotapes of children explaining this task, a rigorous and valid coding scheme for the gestures produced. A second aim is to see whether, using a pretest–posttest paradigm, children also produce gesture–speech mismatches on this task and, if so, whether these are predictive of later learning.

The Representational Redescription Model

Despite producing some illuminating findings about the crucial role children’s gestures play in their learning, previous research has not consistently located gestures within any particular model of cognitive development. Our work aims to address this shortfall by considering gestures within the RR model (Karmiloff-Smith,

1992) because it is one that places significant emphasis on children's nonverbal knowledge. The RR model states that much of children's knowledge begins in an implicit, nonverbal format. Development then involves the gradual emerging into consciousness of this knowledge, via various redescriptions, until it is finally available for verbal report at the last explicit (or E3) level in the model. The RR model has been applied to many domains of children's thinking, from learning science to drawing, mathematics, and theory of mind. The microdomain to be considered here involves children's understanding of balance and makes use of the balance beam task used in previous studies and originally described in the work of Karmiloff-Smith and Inhelder (1974).

Many young children (4–5 years of age) are good at the balance beam task and can balance both types of beam successfully. However, they frequently are unable to explain their success or to state any of the rules underlying the concept of balance. In fact, when asked to explain how they managed to get a particular beam to balance, these children will typically say, "I don't know, I just did it!" They are, according to the RR model, at the implicit level. They have achieved behavioral mastery but their knowledge is still represented in a procedural, nonverbal format.

From these implicit procedures, the cognitive system detects regularities and begins to abstract a central tendency, or rule, to begin building an abstract knowledge system. This is the process of RR postulated to occur in response to an internal drive for understanding and cognitive control. Initially, a theory or rule is abstracted from the implicit procedures, and this is evident in the first explicit level of the RR model, Level E1. On the balance beam task, rule-driven Level E1 behavior is seen in the majority of 6–7-year-old children; for example, Pine and Messer (1999) found that of 168 five- to nine-year-olds tested on the balance beam, 80 children held a "center theory" as described by the model. When given the beams to balance, these children are successful with the symmetrical beams but not with the asymmetrical beams. They place all types of beam onto the fulcrum at their midpoint and, when the asymmetrical beams tip off, will often dismiss these as "impossible to balance." The children have abstracted a rule that all things balance in the middle and they overapply this rule to all instances, causing them to fail on the asymmetrical beams (Karmiloff-Smith & Inhelder, 1974). The RR model states that this level is still nonverbal, although we have found Level E1 to consist of two levels. One is when the child's center theory is nonverbal (this we have termed the *abstraction nonverbal level*), and another is when the child can articulate his or her theory (the *abstraction verbal level*; Pine & Messer, 1999, 2000, 2003). These have been found to be two distinct levels empirically, with different levels of conscious access and receptivity to instruction (Pine & Messer, 2000, 2003).

After Level E1, Karmiloff-Smith (1992) claimed that the representations are redescribed again, this time into Level E2 format, when knowledge becomes consciously accessible but is not yet coded in a linguistic format (nor may it ever be). This level appears to be driven less by empirical data and more by the need to incorporate into the model the feasibility of some consciously accessible spatial, kinesthetic, or other nonlinguistically coded representations (Karmiloff-Smith, 1992, p. 23). To date, the empirical evidence for Level E2 knowledge is limited, and it still remains the most speculative aspect of the model.

Indeed, we would question whether a consciously accessible but nonverbalizable representation could follow E1, given that we have found children who can verbalize their E1 knowledge. However, it has been suggested that children's gestures could reflect other spatial nonverbal knowledge that is not available for speech (Goldin-Meadow & Alibali, 1994). Furthermore,

> because the representational formats underlying gesture are mimetic and analog, rather than discrete, gesture permits speakers to represent ideas that lend themselves to these formats (e.g. shape, size, spatial relationships)—ideas that, for whatever reason, may not be encoded in speech. (Goldin-Meadow, 2002, pp. 1400–1401)

The final level in the RR model is Level E3, characterized by success and verbal explanation. On the balance beam task, children who have reached this level can balance both types of beam on the fulcrum and explain how they balance. In their verbal explanations, these children show understanding of the compensatory function of the weight and distance variables. They might say, when explaining how they balanced an asymmetrical beam, for example, "I had to make this side much longer because it doesn't have as much weight as the other side, so the longer side here will make up for the extra weight here." This type of E3 representation also allows greater cognitive flexibility (implicit-level procedures, though successful, are context bound and inflexible) and the transfer of knowledge to other domains (Pine & Messer, 2003); in short, at Level E3, the RR process is complete and the child has full conceptual understanding.

The RR model is silent about the role of gestures in this redescription process from implicit procedures to conceptual understanding. However, if, as the evidence presented here suggests, gestures reflect ideas that are not well formed enough to express in words, then it seems highly likely that emerging knowledge that is at one of these levels will be conveyed in the children's gestures. Goldin-Meadow and Alibali (2002) pointed out that "gestures convey knowledge that learners have not yet integrated into their explicitly acknowledged view of a problem" (p. 82). Thus, in children whose knowledge has not yet reached Level E3 but is still in the process of emerging, gestures may provide a window into their thoughts. We set out to test this by reanalyzing videotapes of children explaining the balance beam task when their knowledge had not reached Level E3 but had been coded at one of the earlier levels of representation.

This earlier study (Pine & Messer, 2000) set out to test the effect of different types of intervention on children's representations about the balance beam task. Pine and Messer (2000) found significant improvements when children observed an adult modeling the correct solution and were encouraged to explain what they saw. In general, children in this condition showed greater learning gains than children who simply observed a model but did not produce an explanation. However, although the overall difference between the groups was significant, there were nonetheless some children in the more facilitative condition that failed to improve. There were also a number of children in the less facilitative condition that seemed to show some learning gains. In this reanalysis of the videotapes, we set out to see whether the gestures that the children produced at pretest could have been predictive indicators of the children's differential ability to benefit from the intervention. We therefore submitted the videotaped data to detailed reanalysis using a computer-based observer system and transcribing all the gestures and speech the children produced at pretest. If children produced gesture–speech mismatches at pretest and if this indicates a readiness to learn, then these children would be expected to benefit more from the intervention (even when the intervention conditions are less than optimal) than children whose gestures and speech matched at pretest. In summary, the aims of this study were (a) to establish a reliable and valid coding scheme for assigning meaning to the gestures that children produce when explaining the balance beam task, (b) to identify gesture–speech mismatches on the balance beam task and verify empirically whether these predict later learning, and (c) to locate gestural knowledge within a theoretical framework of knowledge representation and development on the basis of the RR model.

Method

Design

In the original study, a pretest, treatment phase, and posttest design was used. The treatment phase had two conditions with a between-subjects design: the observe only (OO) and the observe and explain (OE) conditions. A further between-subjects variable to be investigated here is whether the child was discordant (produced gesture–speech mismatches) or concordant (produced gesture–speech matches) at pretest. The dependent variable is the amount of pretest to posttest improvement in ability to balance beams or in representational level.

Participants

One hundred forty children from two Hertfordshire, England, mixed infant–junior schools participated in the pretest. They ranged in age from 5 to 9 years. There were 61 boys (mean age = 84.8 months) and 79 girls (mean age = 83.83 months).

Materials

For the pre- and posttest, eight wooden balance beams were used. There were four symmetrical beams—two without blocks, one with a block on either end, and one made by overlapping two flat beams. All of these balanced at their geometric center. There were also four asymmetrical beams—one with one block at one end and three with two blocks at each end but varying in thickness and length. All of these balanced off-center.

For the treatment phase, the experimenter modeled one symmetrical beam (without blocks) and one asymmetrical beam (with two blocks at one end). A Panasonic VHS video camera was used to record all sessions, and the gestures identified in these videotaped data are the subject of the data analysis reported here.

Procedure

In the original study, the children were taken individually to a quiet area of the school. They were seated at a table next to the experimenter. After introductions, they were told, “Today we are going to be talking about balancing and playing some balancing games. Do you understand what ‘balancing’ means? What do we mean when we say that something balances?” This was to introduce the context of the task and to ensure that children had encountered the term *balance* before. The experimenter then explained that they would be trying to get some wooden beams to balance on the fulcrum, which was indicated to the children. Children were told that the aim was to make each beam stay level on the fulcrum so it did not tip off to one side or the other.

Pretest

The fulcrum was placed before the child, and each child was asked, “Can you see if you can make the beams stay level on this bar here? That is, make them balance without falling off?” The child attempted the beams one at a time, and the experimenter encouraged the child to give explanations about how each beam balanced or, if it would not, the reason why not. This was done by asking the child after success, “How is that one balancing?” “What do you have to do to make it balance?” or “How did you do that?” Similarly, if a child failed to balance a beam, questions such as, “Why won’t it balance?” “What did you do to try and get it to balance?” or “Do you think it can be balanced?” were posed.

Treatment Phase

Having attempted each of the beams, the child was then randomly assigned to one of the two conditions. In all, 53 children experienced the OO condition and 47 experienced the OE condition.

Observe only. In this condition, when the pretest had been completed, the experimenter told the child, “Now I am going to balance some beams and I would like you to watch carefully how I do it, then you can have another turn at balancing them.” The experimenter showed the child how to balance the symmetrical beam and the asymmetrical beam. The child was not invited to comment or attempt to balance the beams.

Observe and explain. In this condition, the experimenter told the child, “Now I am going to balance some beams and I would like you to watch carefully and try to tell me how each one balances on the bar. Then you can have another turn at balancing them.” The experimenter showed the child how to balance the two beams and invited the child to comment on how this was done. During this session, the children did not themselves attempt to balance the beams.

Posttest

The child was once again asked to balance each of the beams on the fulcrum, as in the pretest, and questions were asked to probe the child’s understanding and to encourage explanations. There was then a short debriefing session in which the experimenter answered any questions the children had, praised them, and then thanked them. Each child’s performance during the session was recorded on a data sheet by the experimenter and was also videotaped. Analysis of a child’s balance beam performance at pre- and posttest enabled the classification of each child into one of the following representational levels (with their correspondence to the original Karmiloff-Smith, 1992, levels in parentheses):

> Implicit (I): The child is able to balance at least two of each type of beam (symmetrical and asymmetrical) but has no consistent strategy for balancing or for initially placing a beam onto the fulcrum. In addition, the child is unable to offer an explanation for his or her success (e.g., “Don’t know” or “I just did it”) or explanations fail to include a mention of the relevant variables of weight and distance.
>
> Implicit transition (transition from I to E1): The child is able to balance no more than one of each type of beam but places all beams onto the fulcrum around their midpoint. Explanations are similar to those at the implicit level (see above).
>
> Abstraction nonverbal (E1): The child is able to balance at least two symmetrical beams but fails on all or all but one of the asymmetrical beams. There is clear evidence of a center strategy, with all beams being placed onto the fulcrum at their midpoint. The child may state that asymmetrical beams cannot be balanced but does not explain a center theory.
>
> Abstraction verbal (E1): Performance is equivalent to the abstraction nonverbal level (see above), but explanations include reference to the center strategy (e.g., “You have to put it in the middle”).
>
> Explicit transition (transition from E1 to E3): The child is able to balance at least two of each type of beam and is able to explain a strategy for balancing both types. For example, the child might say, “You have to put this in the middle,” for a symmetrical beam or, “You

> have to put this one a bit more over to the side," for an asymmetrical beam. However, there is no explanation of the function of the two relevant variables of weight and distance.
>
> Explicit E3 (E3): The child is able to balance at least two, and usually all, of each type of beam and explanations include reference to the compensatory function of the two variables of weight and distance. For example, the child might say, "This side's got more weight on so I make this side longer so that it has the same weight."

(In the study described, children at Level E3 did not continue after pretest, as there was little scope for improvement. However, the coding scheme for speech and gestures was based on the performance of these children.)

This system of classification is derived from Karmiloff-Smith's (1992) RR model, with modifications based on empirical findings from our own research with over 300 children (Pine & Messer, 1998, 1999, 2003). The implicit and explicit E3 levels correspond to those identified by Karmiloff-Smith. Level E1 has been replaced by two levels: abstraction verbal and abstraction nonverbal. Additional transition levels have been identified through previous work and incorporated into the model, although no evidence for Level E2 has been found. This was described by Karmiloff-Smith as a level of representation prior to reaching Level E3 that is consciously accessible but nonverbalizable. Our studies have shown that verbalization of some knowledge is evident by the time children are approaching Level E3, indeed from the abstraction verbal level onward. These revised levels have been the subject of validation by two independent raters, with interrater reliability exceeding 90% (Pine & Messer, 1999). Longitudinal testing has confirmed the hierarchical ordering of the levels, with children tending to progress from the implicit through the abstraction levels and on to Level E3 as their understanding of the task becomes more explicit (Pine & Messer, 2003).

Coding of Gestures

In the original study, the children were classified at a level of representation at pretest, and improvement was assessed via a pretest to posttest change in representational level, from a lower to a higher level within the scheme outlined above. The OO condition was found to produce improvement in 50% of the children, and the OE condition produced improvement in 70% of the children. It was concluded that the OE condition was more helpful to the children in terms of learning gains on the basis of analysis of the children's performance, explanations at pre- and posttest, and being at a higher level of representation at posttest than at pretest. For this study, the videotapes were reanalyzed with the focus on the children's gestures at pretest to see whether these shed further light on the mechanisms producing cognitive change. After all, half the children in the less facilitative condition still improved and almost a third of the children in the more helpful condition failed to improve. By looking at the gestures the children produced, we aimed to explain the differential ability to benefit from the intervention.

Devising the Coding Scheme

The existing videotapes were analyzed with the aim of producing a valid and reliable coding scheme for the gestures produced by the children. Initially, children were identified who were able to balance all types of beams and who were producing the correct verbal explanations, that is, children at the highest level (E3) in the RR model's coding scheme. There were 41 children at this level in total. They completed the pretest only and were then excluded from the rest of the experiment, as further learning was unlikely. Their correct verbal explanations were found to consistently include descriptions of one or more of three variables: weight, distance, and middle. In other words, the children spoke about the weight or weights on either side of the beam, the distance of the weights from the fulcrum (particularly with an asymmetrical beam that was placed off-center), or the importance of placing a beam in the middle (particularly for symmetrical beams). The gestures that regularly accompanied these verbal descriptions were identified as follows:

> Weight: The child closes all fingers together so they are pointing downward and moves the hand up and down over one end of the beam.
>
> Distance: The child moves his or her hand from the fulcrum to the end of the beam. The child's hand is either completely flat or clenched, with one finger pointed out. The palm faces downward, and the movement is from side to side.
>
> Middle: The child points at the middle of the beam on the fulcrum with one or both hands.

Reliability was established by having a second observer view a proportion (approximately 45%) of the videotaped sessions and code the children's gestures that accompanied correct explanations. The first rater identified a range of concordant explanations within these sessions, that is, where the spoken explanation given matched the gesture the child produced. Ten of these explanations included middle, seven included weight, and nine included distance. Interrater reliability was 98% agreement between the two coders for describing gestures that most frequently accompanied correct explanations. This coding scheme was then used to code the remaining 99 participants' gestures.

Identifying Discordant and Concordant Children

Next, the focus shifted to the 99 remaining children as they attempted each beam at pretest. These children were at one of the other levels in the classification system, that is, not yet at Level E3, and many were unable to balance some or all of the beams. All videotaped data were transcribed for gesture and speech according to any variables (weight, middle, or distance) that were being conveyed. Speech was coded by turning off the video picture and only listening to the audio portion of the tape to ensure that the experimenter could not be influenced by the presence of any gestures. The videotape was then rerun with the picture on but the sound turned off, and the children's gestures were coded in isolation from their spoken explanations. The next step involved comparing the spoken variable with the gestured variable and classifying the children as either concordant (gesture–speech matching) or discordant (gesture–speech mismatching).

Children were classified as concordant if their gestures were coded as expressing the same information as their speech.[1] Children were categorized as discordant if they produced a gesture that was coded differently from their speech. For example, a discordant child might explain that a beam balances "because it is in the middle" but might accompany this with a weight or distance gesture. We hypothesized that children classified as discordant would make a greater improvement on the balance beam task from pretest to posttest than would the children who were classified as concordant.

Results

Children were classified as discordant if they produced at least one gesture–speech mismatch at pretest or concordant if their gestures and speech matched. Of the 99 children who completed all parts of the study, 36 were classified as discordant and 63 as concordant. In this section, the effect of concordance and discordance on improvement is measured first as a change to a higher level of representation from pre- to posttest and then as an increase

---

[1] Children were also classified as concordant if they spoke without gesturing (as in Goldin-Meadow, Nusbaum, Garber, & Church, 1993), although there were only 3 such children in this study.

in the number of asymmetrical beams balanced at pre- and posttest. In the third section, we examine the relationship between producing concordant or discordant gestures and the child's level of representation.

The Effects of Gesture–Speech Match and Mismatch on Improvement in Pretest to Posttest Representational Level

Here, improvement is measured by a change to a higher level of representation from pretest to posttest and is a dichotomous categorical dependent variable (improve–not improve). Almost 50% of the concordant children improved, and of the smaller group of discordant children, 78% improved.

These frequencies were analyzed using chi-square tests, and a significant association between concordant–discordant gestures at pretest and improvement at posttest was found, $\chi^2(1, N = 99) = 6.98$, $p < .01$. However, the children who produced concordant and discordant gestures had not all experienced the same intervention between pre- and posttest. Of the 63 concordant children, 29 had experienced the OE condition and 34 had experienced the OO condition. Of the 36 discordant children, 18 had been in each condition.

Therefore, it was of interest to examine the association between pretest concordance–discordance and posttest improvement according to each of the conditions. A log-linear analysis of the three-way term Condition × Classification × Improve–Not Improve was conducted but did not reach significance ($\chi^2 = 0.04$, *ns*). Next, chi-square analyses of the separate associations between condition and improvement and classification and improvement were conducted. Table 1 shows the number of concordant and discordant children improving and the condition experienced between pre- and posttest. Children in the OE condition were more likely to improve than not. This endorses the finding of the earlier study (Pine & Messer, 2000) that this condition was most likely to be associated with improvement. Chi-square analysis revealed no association between concordance–discordance and improvement; both types of children were likely to improve under this more facilitative condition, $\chi^2(1, N = 47) = 2.40$, *ns*. However, even in this group, 83% of the discordant children improved compared with 62% of the discordant group. Of interest, this differed for the OO condition—the condition found to be less helpful in bringing about improvement. Although only 50% of the children improved in this condition, among the discordant children, this figure reached 72%, but for the concordant children, it was just 41%. Chi-square analysis found that this association between concordance–discordance and improvement was a reliable one, $\chi^2(1, N = 52) = 4.11$, $p < .05$. This suggests that the OO condition was least likely to help those children who were concordant, and children who were discordant fared better in the face of the less than optimal intervention.

Table 1
Number (and Percentage) of Discordant–Concordant Children Who Improved From Pre- to Posttest in Representational Level According to the Condition Experienced

Posttest improvement	Observe and explain		Observe only	
	Concordant	Discordant	Concordant	Discordant
Improve	18 (62%)	15 (83%)	14 (41%)	13 (72%)
Not improve	11 (38%)	3 (17%)	20 (59%)	5 (28%)

Measuring Improvement by Gains in Ability to Balance Asymmetrical Beams

The above analyses measured whether children improved from pre- to posttest in terms of the level of representation at which they were classified. It was also considered of interest to look at a behavioral measure that yielded quantitative data, that is, the number of asymmetrical beams balanced at pre- and posttest. Because these beams are the most difficult for children to balance when they first attempt the task, this is the measure that shows the greatest scope for improvement. For example, many children balance symmetrical beams at pretest but fail to balance asymmetrical beams because they try to get them to balance at the center rather than positioning them off-center. After seeing these beams balanced, and particularly after being encouraged to talk about how they balance, many children subsequently succeeded with some or all of the asymmetrical beams at posttest.

A mixed analysis of variance was conducted, with the number of asymmetrical beams balanced at pre- and posttest as the repeated-measures dependent variable and pretest classification (concordant, discordant) and condition (OO, OE) as the between-subjects factors. The means for each group are given in Table 2. There was a main effect of pretest to posttest change, $F(1, 95) = 116.40$, $p < .01$, with all children showing pretest to posttest improvement in ability to balance asymmetrical beams. There was a main effect of classification, $F(1, 95) = 4.75$, $p < .01$, but no main effect of condition, $F(1, 95) = 0.00$, *ns*. There was a reliable Pretest–Posttest × Condition × Classification interaction, $F(1, 95) = 4.59$, $p < .05$, indicating that children's ability to benefit from the intervention condition differed according to whether they were classified as concordant or discordant at pretest. The OE condition was most likely to be associated with improvement, as was being discordant at pretest, and it appeared that the worst case for children was to be concordant at pretest and experience the OO condition.

Association Between Representational Level and Concordance–Discordance

Of further interest are the level of representation at which the children were classified and the type of gesturing associated with each of the levels. One prediction from the characterization of the levels of the RR model is that once a child has reached a level, then concordance would be achieved, that is, with a unitary level of representation. Children in transition between levels, however, might be expected to be more discordant, as they consider more than one hypothesis and possibly entertain multiple representations before settling on one. Because our coding scheme included two transition levels, it was of interest to see whether these produced the highest rates of discordance. The explicit transition level did feature twice as many discordant as concordant children, a pattern that was the converse of that found overall and that fitted with our prediction. This was not the case with the implicit transition level, however, where there were three times as many concordant as discordant children. Although the numbers at some of the levels were low, with the majority of children being at one of the abstraction levels (either verbal or nonverbal), it can be seen from

Table 2
Mean Number of Asymmetrical Beams Balanced by Children at Pre- and Posttest According to Gesture–Speech Classification and Condition

	Concordant			Discordant		
Condition and test	*n*	*M*	*SD*	*n*	*M*	*SD*
Observe and explain	29			18		
Pretest		1.24	1.18		1.88	1.36
Posttest		2.89	1.29		3.55	1.04
Observe only	34			18		
Pretest		1.44	1.30		1.22	1.16
Posttest		2.44	1.46		3.55	0.78

Table 3 that discordance could be detected at all levels and was not just confined to the transition levels. A chi-square analysis of these frequencies confirmed that there was no reliable association between representational level and concordance–discordance, $\chi^2(4, N = 99) = 5.15$, *ns*.

Discussion

One of the aims of this study was to establish a reliable and valid coding scheme for assigning meaning to the gestures that children produce when explaining the balance beam task. We found that children's verbal and gestural explanations fell into three distinct categories, relating to three dimensions of the task: weight, distance, and middle. High interrater reliability confirmed that these could be reliably coded, and the coding scheme was used to classify 99 children's speech and gestures on the balance beam pretest.

A further aim was to identify gesture–speech mismatches and verify empirically whether these predict knowledge change in response to being exposed to someone modeling the solution. Having coded children's speech and gestures at pretest, it was possible to identify children who, when producing an explanation about the task, indicated one variable in speech and another in gesture. These children were classified as discordant. Children whose speech and gestures matched were classified as concordant. Just over one third of the sample of children were classified as discordant. When the improvement from pre- to posttest was analyzed, both in terms of improved competence in balancing asymmetrical beams and improved understanding in terms of the representational level, it was found that there was a reliable association between being discordant at pretest and subsequent posttest improvement. This supports the hypothesis that children in a discordant state are ready to learn and that their gesture–speech mismatches predict later learning. It was also found that discordant and concordant children differed in their ability to make use of the intervention experience. We had previously shown that the condition in which children observed a model and explained what they saw (OE) was more facilitative than a condition in which children merely observed but did not explain (OO; Pine & Messer, 2000). By reanalyzing these data to see whether concordance–discordance was a mediating factor in the learning process, we found that those children who did improve despite being in the less helpful condition were more likely to have been discordant at pretest. Thus, discordant children fared better, even when the intervention condition was less than optimal (i.e., the OO condition). Therefore, the worst outcome was for many of the children who were concordant at pretest and who experienced the OO condition. Because the concordance of many of these children indicated they were not open to instruction, only the most optimal learning conditions would have had a chance of inducing change. When this was not provided, these children were likely to remain at the same state of learning they were in when the study began, as indeed the majority of concordant children in the OO condition did.

These findings are by no means conclusive and can only reflect the general trend found in these data. There were still a few children who were discordant and failed to improve or who were concordant and improved, but the fact that these were the least likely associations suggests that these findings are of importance and that assessing children's gestures can be informative about their readiness to learn. This study is also an example of how failing to take account of children's gestures means that an important source of information is overlooked. Looking at the concordance or discordance of their gestures has been found to explain why some children improve and others do not when given the same learning conditions. For example, in our previous study, we found that the OE condition produced learning gains in most of the children but that there were still 14 children who failed to improve. By looking at these children's gestures, we now see that more than three times as many of those nonimprovers were concordant as opposed to discordant. Learning intervention studies rarely produce 100% success, and measures of effectiveness are frequently based on whether the majority of children benefit. Looking at children's gestures as a means of identifying children who are ready to learn offers insight into the differential outcomes of these studies and extends knowledge about why some children respond to intervention and others do not.

During this study, it was noted that sometimes children appeared to first display an idea in a gesture before they were able to express it verbally. This has important practical and educational implications. An adult who is interacting with a child may use these gestural signals as a way of interpreting the child's knowledge state. If an adult can detect when a child is ready to learn, as indicated by gesture–speech mismatch, then they may modify the way that they interact with the child. There is evidence to suggest that most adults, and not just those involved with teaching children, attend to gestures when assessing what children know (Alibali et al., 1997; Goldin-Meadow & Momeni-Sandhofer, 1999; Pine, 2003). Furthermore, adults are not just sensitive to the gestures that children produce; they may even modify the information they provide to the child accordingly. Goldin-Meadow and Singer (2003) have shown that adults teach a wider range of problem-solving strategies to children who are producing mismatches than to those who are not.

Table 3
Number (and Percentage) of Children Classified as Discordant-Concordant at Pretest According to Representational Level

Representational level at pretest	Concordant	Discordant
Implicit	9 (64%)	5 (36%)
Implicit transition	6 (73%)	2 (25%)
Abstraction nonverbal	17 (68%)	8 (32%)
Abstraction verbal	25 (62%)	15 (38%)
Explicit transition	4 (33%)	8 (67%)

It is therefore possible that the first point of exit for a child's newly emerging knowledge may be via his or her hand gestures. These send a reliable signal to an interactive partner about the child's cognitive state, about whether the child is on the verge of a new insight, or about the child's zone of proximal development, as first described by Vygotsky (1978). Adults can then respond to those cues with the most appropriate instruction, and, as a result, the child has successfully shaped his or her own learning environment.

Finally, we sought to locate gestural knowledge within a theoretical framework of knowledge representation on the basis of Karmiloff-Smith's (1992) RR model. This model was invoked because we had already established a means of coding children at most of the levels of the model on the balance beam task based on extensive previous work (Pine & Messer, 1999, 2000, 2003; Pine, Messer, & Godfrey, 1999; Pine, Messer, & St. John, 2002). However, more importantly, it is one model that takes account of children's nonverbal knowledge, and we hoped to establish whether this knowledge was detectable in the children's gestures. The results at each level of representation were interesting but somewhat equivocal. One hypothesis was that discordance would be more likely to appear at transition levels, that is, when the child was in transition between the implicit and E1 (abstraction) levels or between E1 (abstraction) levels and E3. This was not found to be the case because at every level, a number of children exhibited discordance between their speech and gestures. This casts some doubt about the conceptual stability of these representations and suggests that each level may not be a unitary representation and, in fact, encompasses more than just a single idea. In terms of the nature of cognitive representations, it tells us that cognitive variability, as manifested by gesture–speech mismatches, occurs at all stages of the knowledge-acquisition process. Even when knowledge appears to be in a rigid theory-driven representational state (such as Level E1 or, as depicted here, the abstraction levels), there will be times when the child is also entertaining an alternative hypothesis and this may leak out in gesture.

Thus, although it has been hard to find empirical evidence for Level E2 as Karmiloff-Smith (1992) described it, her notion that knowledge can be in the system at a conscious, though nonverbalizable level, has considerable credence. Observations of the nature of children's discordance in this study suggested that children may gesture knowledge that they were unable to verbalize. In other words, their gestures may be in advance of their speech or information may be conveyed uniquely in gesture (i.e., never appear in their speech). We are now conducting microgenetic analyses of such mismatches to determine more accurately the degree of temporal and semantic synchrony between children's speech and gestures (Pine, Lufkin, & Messer, 2004).

This leads us to conclude that the elusiveness of Level E2 is due to Karmiloff-Smith's (1992) conceptualization of it as a level of representation per se. The notion of knowledge that is in the cognitive system at a conscious but nonverbalizable level may be more plausibly viewed as a pervasive characteristic of the redescription process rather than a level per se. Our data revealed no reliable association between representational level and concordance–discordance. In other words, although children are at any of the other levels of representation, they may have knowledge that is consciously accessible and that leaks out in gesture before it is verbalizable. Although failing to shed further light on Level E2, this finding does speak to the dynamic nature of knowledge development and the crucial role that variability plays continuously in the learning process (see Siegler, 1996). It also extends Karmiloff-Smith's conjecture regarding the RR model being based on a multirepresentational system and provides empirical support for this notion. Gestures may even play a role in triggering RR from one level to the next, an issue about which the model has hitherto been relatively silent.

In summary, an important contribution of the research presented here has been to extend the range of domains within which children's gestures have been shown to play an important role in learning. McNeill (1992) described gestures as "microgenetically evolving representations" (p. 250), and these findings endorse this notion of gestures as an integral part of the child's thinking processes. The children in this study were found to express meaningful information in gesture, and over one third of them conveyed different information than that expressed in their speech. Moreover, as well as reflecting the children's emerging knowledge, gesture–speech mismatches were also found to be a reliable indicator of children's receptivity to instruction and their ability to make use of different learning conditions. We conclude that gestures are an important research tool for assessing children's knowledge and for illuminating the process of learning.

References

Alibali, M. W. (1995). Processes of cognitive change revealed in gesture and speech. *Dissertation Abstracts International, 55*(12), 5580B.

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science, 10,* 327–333.

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than just keeping track. *Cognitive Development, 14,* 37–56.

Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology, 89,* 183–193.

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture–speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology, 25,* 468–523.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15,* 593–613.

Breckinridge-Church, R., Kelly, S., & Lynch, K. (2000). Immediate memory for mismatched speech and representational gesture across development. *Journal of Nonverbal Behaviour, 24,* 151–174.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23,* 43–71.

Evans, J. L., Alibali, M. W., & McNeil, N. M. (2001). Divergence of verbal expression and embodied knowledge: Evidence from speech and gesture in children with specific language impairment. *Language and Cognitive Processes, 16,* 309–331.

Garber, P. R. (1998). Using gesture and speech to index planning a problem solving task: A comparative study of adults and children explaining the Tower of Hanoi puzzle. *Dissertation Abstracts International, 58*(10), 5670B.

Garber, P. R., Alibali, M. W., & Goldin-Meadow, S. (1998). Knowledge conveyed in gesture is not tied to the hands. *Child Development, 69,* 75–84.

Garber, P. R., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science, 26,* 817–831.

Goldin-Meadow, S. (2002). Constructing communication by hand. *Cognitive Development, 17,* 1385–1405.

Goldin-Meadow, S., & Alibali, M. W. (1994). Do you have to be right to redescribe? *Behavioural and Brain Sciences, 17,* 718.

Goldin-Meadow, S., & Alibali, M. W. (2002). Looking at the hands through time: A microgenetic perspective on learning and instruction. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 80–105). Cambridge, England: Cambridge University Press.

Goldin-Meadow, S., & Momeni-Sandhofer, C. (1999). Gestures convey substantive information about a child's thoughts to ordinary listeners. *Developmental Science, 2,* 67–74.

Goldin-Meadow, S., Nusbaum, H., Garber, P., & Church, B. (1993). Transitions in learning: Evidence for simultaneously activated strategies. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 97–107.

Goldin-Meadow, S., & Singer, M. (2003). From children's hands to adults' ears: Gesture's role in the learning process. *Developmental Psychology, 39,* 509–520.

Graham, T. (1999). The role of gesture in children's learning to count. *Journal of Experimental Child Psychology, 74,* 333–355.

Halford, G. S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Erlbaum.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Norton.

Iverson, J. M., & Goldin-Meadow, S. (1997). What's communication got to do with it? Gestures in children blind from birth. *Developmental Psychology, 33,* 453–467.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

Karmiloff-Smith, A., & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition, 3,* 195–212.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago Press.

Perry, M., Breckinridge-Church, R., & Goldin-Meadow, S. (1992). Is gesture–speech mismatch a general index of transitional knowledge? *Cognitive Development, 7,* 109–192.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3,* 359–400.

Pine, K. J. (2003, July). *Assessing children's understanding of a balance task: Do untrained observers attend to gestures?* Poster presented at the meeting of the Experimental Psychology Society, Reading, England .

Pine, K. J., Lufkin, N., & Messer, D. (2004, April). *Assessing children's non-verbal knowledge through their hand gestures.* Paper presented at the annual conference of the British Psychological Society, London.

Pine, K. J., & Messer, D. J. (1998). Group collaboration effects and the explicitness of children's knowledge. *Cognitive Development, 13,* 109–126.

Pine, K. J., & Messer, D. J. (1999). What children do and what children know: Looking beyond success using Karmiloff-Smith's RR framework. *New Ideas in Psychology, 17,* 17–30.

Pine, K. J., & Messer, D. J. (2000). The effects of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction, 18,* 37–54.

Pine, K. J., & Messer, D. J. (2003). The development of representations as children learn about balancing. *British Journal of Developmental Psychology, 21,* 285–301.

Pine, K. J., Messer, D. J., & Godfrey, K. (1999). The teachability of children with naive theories: An exploration of the effects of two teaching methods. *British Journal of Educational Psychology, 69,* 201–211.

Pine, K. J., Messer, D. J., & St. John, K. (2002). Children's learning from contrast modelling. *Cognitive Development, 17,* 1249–1263.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481–520.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Vygotsky, L. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.

Received June 16, 2003
Revision received May 11, 2004
Accepted May 18, 2004 ■