Transitions in Concept Acquisition: Using the Hand to Read the Mind

Article *in* Psychological Review · May 1993

DOI: 10.1037//0033-295X.100.2.279 · Source: PubMed

3 authors:

Susan Goldin-Meadow
University of Chicago

Martha W. Alibali
University of Wisconsin–Madison

Ruth Breckinridge Church
Northeastern Illinois University

Psychological Review
1993, Vol. 100, No. 2, 279-297

Transitions in Concept Acquisition: Using the Hand to Read the Mind

Susan Goldin-Meadow, Martha Wagner Alibali, and R. Breckinridge Church

Thoughts conveyed through gesture often differ from thoughts conveyed through speech. In this article, a model of the sources and consequences of such gesture–speech mismatches and their role during transitional periods in the acquisition of concepts is proposed. The model makes 2 major claims: (a) The transitional state is the source of gesture–speech mismatch. In gesture–speech mismatch, 2 beliefs are simultaneously expressed on the same problem—one in gesture and another in speech. This simultaneous activation of multiple beliefs characterizes the transitional knowledge state and creates gesture–speech mismatch. (b) Gesture–speech mismatch signals to the social world that a child is in a transitional state and is ready to learn. The child's spontaneous gestures index the zone of proximal development, thus providing a mechanism by which adults can calibrate their input to that child's level of understanding.

Children typically reveal their understanding of a problem through their speech. However, along with that speech, children often produce gestures, and those gestures have also been shown to provide insight into a speaker's thoughts (Kendon, 1980; McNeill, 1985, 1987, 1992). Surprisingly, the thoughts conveyed through gesture are not always the same as the thoughts conveyed through speech. They may *mismatch.* In this article, we propose a model of the sources and consequences of such mismatches and their role in the acquisition of concepts.

One of developmental psychology's major contributions to the study of concept acquisition is the demonstration that a child's understanding of at least certain concepts is, throughout the period of acquisition, systematic and rule governed. The process of acquisition for these concepts can therefore be characterized as an advance from an inadequate yet systematic understanding of a concept to a more adequate, systematic understanding of a concept. In this context, the transitional knowledge state can be taken to be the bridge between two rule-governed knowledge states. It is this type of transitional state that we explore in children, although we believe that transitional states of this sort may characterize learning at any age. In this regard, it is important to note that the behavior we take to reflect the transitional state—gesture–speech mismatch—is not unique to a particular period in the life span but is found in adults (Goodman, Church, & Schonert, 1991; McNeill, 1992), as well as adolescents (Stone, Webb, & Mahootian, 1991), school-aged children (Church & Goldin-Meadow, 1986; Evans & Rubin, 1979; Perry, Church, & Goldin-Meadow, 1988), and 2-year-olds (Gershkoff-Stowe & Smith, 1991; Morford & Goldin-Meadow, 1992).

The model we propose makes two major claims about the transitional knowledge state and its relation to gesture–speech mismatch. First, the model argues that the transitional knowledge state is the source of gesture–speech mismatch. In previous work, we have found that the mismatch between gesture and speech in a child's explanations of a concept reflects an instability in the child's knowledge of that concept and is, moreover, a good indicator that the child is in a transitional knowledge state with respect to the concept. Here, we suggest that gesture–speech mismatch is not merely an index of transitional knowledge; that is, it is not an epiphenomenon of the transitional state, co-occurring with transition but not related in any essential way to the state itself. Rather, we argue that gesture–speech mismatch is caused by the processes that characterize the transitional state. In gesture–speech mismatch, two beliefs are simultaneously expressed on the same problem—one in gesture and another in speech. We suggest that it is the simultaneous activation of multiple beliefs that characterizes the transitional knowledge state and creates gesture–speech mismatch.

Second, the model argues that gesture–speech mismatch signals to the social world that an individual is in a transitional knowledge state. Occurring as it does in communicative contexts, gesture–speech mismatch may provide a signal to those who interact with a child, announcing to those who can interpret the signal that the child is in a transitional state and thus is ready to learn. In fact, we suggest that gesture–speech mismatch provides a measure of the zone of proximal development, as put forth by Vygotsky (1978). In Vygotsky's account of development, the *zone of proximal development* is defined as the distance between what children can do on their own and what they can do with the guidance of an adult or a more capable peer. Development is powered by the child's internalization of the cognitive processes shared in the zone of proximal devel-

---

Susan Goldin-Meadow and Martha Wagner Alibali, Department of Psychology, University of Chicago; R. Breckinridge Church, Northeastern Illinois University.

This research was supported by Grant R01 HD18617 from the National Institute of Child Health and Human Development to Susan Goldin-Meadow.

We thank Janellen Huttenlocher and David McNeill for inspiring this article, and Lawrence Barsalou, Miriam Bassok, Janellen Huttenlocher, David McNeill, Howard Nusbaum, Jim Stigler, and the anonymous reviewers for their insightful comments on the article itself.

Correspondence concerning this article should be addressed to Susan Goldin-Meadow, Department of Psychology, University of Chicago, 5730 South Woodlawn Avenue, Chicago, Illinois 60637.

opment. However, it is unclear from this account how an adult is able to zero in on a child's zone of proximal development. We suggest that the spontaneous gestures children produce when communicating with adults provide an observable index of the zone of proximal development, and thus they provide a mechanism by which adults can calibrate their input to a child's level of understanding.

In a fundamental sense, the way in which the transitional knowledge state is defined and characterized is of theoretical importance precisely because that characterization constrains the types of mechanisms that can be proposed to account for developmental change—and investigating mechanisms of change is at the heart of research on development and learning (cf. Flavell, 1984). Our purpose here is, first, to review what is currently known about the transitional knowledge state and the processes that define that state, situating gesture–speech mismatch within this field. We then ask how children in a transitional state signal to the social world that they are in such a state, and we argue that gesture and speech taken together provide this signal and, as such, play an important role in the mechanisms responsible for developmental change.

Defining and Characterizing the Transitional State

Theories of learning have been broadly classified as relying on either accumulation or replacement mechanisms to account for change (Mazur & Hastie, 1978). Accumulation theories view learning as the systematic acquisition of increasing amounts of information or skill, whereas replacement theories operate by qualitative reorganizations or changes in mental representations and rules. Particularly if learning is propelled by a replacement mechanism, it is likely that there will be a period of transition (however abrupt) between the time when an old rule is firmly held and the time when it is replaced by a new rule.[1] The difficulty, however, lies in identifying and characterizing that period of transition. Many different operational definitions of the transitional period have been proposed. For example, at various times, the transitional state has been defined as a state of readiness to learn, as a state in which guidance improves performance, as a state of partial knowledge, as a state in which multiple hypotheses are considered, and finally, as a state in which multiple hypotheses are considered simultaneously. We address each of these operational definitions in turn, evaluating how effective each definition is in capturing the transitional knowledge state.

The Transitional State as a State of Readiness to Learn

By definition, the child in a transitional state with respect to a particular concept has not yet mastered that concept. The upper bound of the transitional state is therefore, at least in principle, easy to establish. However, at what point in a child's development of a concept is it sensible to say that the child has entered a transitional state with respect to that concept? For example, children's understanding of balance scale problems develops slowly and, although children can make reliable predictions about balance scale problems at age 5, a complete understanding is often not achieved until age 17 (Siegler, 1976). Do we want to say that the child is in transition from age 5 to 17? To be of theoretical use, the notion of transitional state needs to be more constrained than the period prior to mastery.

Intuitively, one might associate the transitional state with instability prior to mastery. Instability conveys the sense that change is imminent, particularly if appropriate input is provided. Thus, one might argue that a child is in a transitional state with respect to a particular concept when that child is ready to learn or to benefit from input in that concept; that is, given appropriate input, the child with transitional knowledge masters the concept quickly (cf. Beilin, 1965; Brainerd, 1972; Langer & Strauss, 1972; Murray, 1974; Strauss & Langer, 1970; Strauss & Rimalt, 1974).

However, Brainerd (1977) convincingly argued that many of the studies cited as evidence that children in transition are particularly ready to learn are methodologically flawed, primarily because the child's pretest performance is not taken into account when measures of learning are calculated. For example, in many studies, children are divided into two groups on the basis of the knowledge they demonstrate on a conservation pretest: Children who demonstrate an understanding of conservation on some problems but not on all problems are classified as *transitional;* they are contrasted with children who demonstrate no understanding of conservation on the pretest and are therefore classified as *nonconservers.* The transitional children are expected to be particularly likely to benefit from training, and these children are indeed more likely than nonconservers to achieve conservation after training. However, the transitional children also knew more about conservation before training than did the nonconservers. The real question is not whether the transitional children achieved conservation after training but whether they made more progress after training than did the nonconservers. Indeed, using pretest–posttest difference scores to reanalyze data from a large number of conservation studies, Brainerd (1977) found that transitional children made no more progress after training than nonconservers; this suggests that they were no more ready to learn than nonconservers. In addition, Brainerd (1977) pointed out that, to avoid circularity, the measure used to capture learning ought not be the same as the measure used to capture potential to learn. For example, if performance on a conservation task is used to measure how much the child has learned after instruction, that same measure ought not be used to determine whether the child was in a transitional state before instruction.

In general, note that if *readiness to learn* is used as a definition of the transitional state, the transitional state is easy to pinpoint—but only in retrospect. Any group of children can be exposed to instruction, and the subset who benefit from the instruction will have been, by definition, *in transition.* Thus, although intuitively appealing as a definition of the transitional state, *openness to instruction* or *readiness to learn* cannot be used to identify children in the transitional state simply be-

---

[1] It is important to point out that we are not arguing for a domain-general period of transition akin to the transitional stage proposed by Piaget as a bridge between preoperational and concrete operational thought. Rather, our focus is on domain-specific transition, specifically, on the period during which a learner moves from a less adequate to a more adequate understanding of a particular concept.

cause the index is post hoc and lacks predictive value. Nevertheless, readiness to learn can serve as a useful standard against which other prospective characterizations of the transitional state must be evaluated. In other words, any proposed marker of the transitional state ought to be able to divide children into those who are more and less ready to progress to a new level of understanding of a concept.

Thus, in this article, we do not test whether readiness to learn is a good definition of the transitional state. We assume that it is and, on this basis, use it as a criterion for evaluating other operational definitions of the transitional state (definitions that do not suffer from being post hoc). In the rest of this discussion, we follow this heuristic and ask whether each of the definitions of the transitional state we consider successfully identifies children who are ready to profit from instruction and to learn. In evaluating the utility of these definitions, we pay careful attention to the methodological pitfalls pointed out by Brainerd (1977). As described earlier, the measure used to capture learning in studies of readiness must take into account the child's level of knowledge before training; that is, the measure must index improvement in performance rather than absolute level of performance. In addition, to avoid circularity, the measure used to capture learning ought not be the same as the measure used to identify children in the transitional state.

The Transitional State as a State in Which Guidance Improves Performance: The Zone of Proximal Development

Vygotsky (1978) defined the zone of proximal development as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). The zone of proximal development contains those concepts that are currently developing, that is, those concepts that the child has not yet fully acquired but is in the process of acquiring.

If a particular concept is contained in a child's zone of proximal development, the child will have more success solving problems instantiating that concept when assisted by an adult than when working alone. In this sense, the child has transitional knowledge with respect to that concept. Children who show no difference in their performance under the two conditions are either impervious to the opportunities provided by the adult and thus do not yet have that concept in their zone of proximal development or have already incorporated the skills into their everyday activity and thus have mastered the concept (cf. Griffin & Cole, 1985).

Certain authors, most notably, Fischer (e.g., Fischer & Pipp, 1984), have suggested that adjusting the conditions under which a child performs a task can result in better, even optimal, performance. However, Vygotsky (1978) suggested that better performance with adult guidance may itself be a good predictor of future performance alone; if so, it would meet the readiness-to-learn criterion proposed earlier as a crucial test for an operational definition of the transitional state. In fact, the predictive potential of measures of the zone of proximal development has not been extensively tested. There are numerous studies of how adults adjust their input to the level of the child (e.g., Rogoff & Wertsch, 1984) and several studies of the relationship between measures of the zone of proximal development and traditional measures of intelligence (Brown & Ferrara, 1985; Ferrara, Brown, & Campione, 1986). However, few, if any, studies have been specifically designed to test whether performance with adult guidance on a task predicts later performance alone on that task.

Nevertheless, a number of studies of peer interaction contain data that bear on this issue. For example, Azmitia (1988) studied a group of 5-year-olds who, over a series of sessions, built replicas of a model alone, with a peer of the same ability, or with a peer of greater ability. Azmitia found that novices who worked with an expert improved their performance during the interaction and built more accurately than novices who worked alone or with other novices. It is important for our purposes to note that this improved performance was maintained when the children subsequently worked alone on a different model than the one they had worked on in collaboration. Thus, skills children exhibit when working in collaboration with an expert can be maintained when children later work alone.

However, these data do not address the issue of individual differences, which is crucial to determining who is and who is not in a transitional state; that is, are children who work better with an expert more likely to exhibit these skills when working alone than children who do not work as well with an expert? In fact, Forman and Cazden (1985), in a study of 9-year-olds who performed a logical-reasoning task with and without a peer, found that children who excelled when collaborating on the task did not always do better on their own compared with children who did not do as well in collaboration. Thus, some collaborative partners demonstrated a higher level of combinatorial reasoning in a social context than they did individually, suggesting that socially achieved cognitive activities are not always completely internalized by the participants. Given the fact that improved performance with guidance is not necessarily associated with improved performance alone, the zone of proximal development may not be an altogether reliable predictor of children who are ready to learn, although further studies designed to test this point are needed.

The Transitional State as a State of Partial Knowledge

Another definition of the transitional state holds that the child in transition has partial knowledge of the concept in question, where *partial knowledge* is operationally defined as solving some problems instantiating the concept correctly and solving others incorrectly (Wilkinson, 1982a, 1982b). At first glance, one might think that partial knowledge is, in fact, another name for transitional knowledge. However, it is possible to tease apart these two notions, and when one does so, partial knowledge turns out to be neither necessary nor sufficient for transitional knowledge.

Partial knowledge, as just defined, cannot be necessary for a child to be in a transitional state simply because a child can solve all of the problems instantiating a particular concept incorrectly and yet, by the readiness-to-learn criterion, still be in a transitional state with respect to that concept. For example, Perry et al. (1988) examined a group of fourth-grade children, none of whom could correctly solve problems instantiating the concept of mathematical equivalence, and found that some children benefited from instruction in the concept, whereas others did not. These children, by the readiness-to-learn criterion discussed earlier, were in a transitional knowledge state, yet none showed partial knowledge prior to instruction.

In addition, partial knowledge does not appear to be sufficient for a child to be in a transitional state simply because a child who is sometimes right and sometimes wrong on a set of problems instantiating a concept need not be in a transitional state with respect to that concept. For example, Wilkinson (1982a, 1982b) demonstrated that some children who produce both correct and incorrect responses on a series of problems consistently fail the same set of problems each time they solve them; this suggests that their knowledge, although incorrect, is nevertheless consistent. Wilkinson called this state *restricted partial knowledge* and assumed that a child in such a state has a single, incorrect rule that is reliably applied to all problems. Similarly, Siegler (1976, 1981) showed that the majority of children who had not yet grasped the factors underlying the balance scale and had solved only some of the problems correctly arrived at their responses by using a single rule in a consistent fashion. For example, some children solved problems correctly only when the weights on each side of the scale were an equal distance from the fulcrum, and thus they appeared to be using a consistent, albeit incorrect, rule to solve the problems (the weight-only rule: *the side with the most weight always goes down*).

If we apply the readiness-to-learn criterion to children with restricted partial knowledge, we find that some of the children who used a single, incorrect rule on Siegler's tasks benefited from instruction, whereas others did not. In particular, Siegler (1976) found that, although many 5- and 8-year-old children used the weight-only rule to solve the series of balance scale problems prior to instruction, when finally given instruction in the problems, the 8-year-olds were more likely to acquire a more advanced rule (and thus were more likely to have been in transition) than the 5-year-olds. Yet, the partial knowledge criterion identified all of the children as being in transition, even the 5-year-olds, who were far less ready to learn the new rule than the 8-year-olds.

Probing further the children's behavior prior to instruction, Siegler (1976) found that, although both the 5- and 8-year-olds used the rule involving weight to solve the series of problems, only the 8-year-olds seemed to be aware that a second dimension (distance) was also relevant to the problem. The 8-year-olds produced nonverbal cues (in particular, head movements) that suggested that they were aware of the weights' distance from the fulcrum, whereas the 5-year-olds gave no such evidence. In a subsequent encoding study, Siegler confirmed that the 8-year-olds were aware of both of the relevant dimensions on the balance scale—weight and distance—whereas the 5-year-olds were aware only of weight. Thus, although the 8-year-olds used a single rule involving weight to produce answers to the balance scale problems, they also seemed to be entertaining a second hypothesis involving distance. In fact, the consideration of more than one hypothesis has itself been proposed as a defining characteristic of the transitional state, as we describe next.

The Transitional State as a State in Which Multiple Hypotheses Are Considered

When acquiring a concept, it is possible that children abruptly and completely abandon one hypothesis in favor of another. However, it seems more likely that a child will continue to entertain an old hypothesis while beginning to develop a new one. One might therefore expect a period of transition during which there is evidence for more than one hypothesis—an old one and a new one—in the child's behavior. Thus, there is intuitive reason to believe that when children move from one rule-governed state to the next, they pass through a transitional period during which they entertain more than one hypothesis.

Moreover, there is theoretical reason to believe that consideration of more than one hypothesis characterizes the transitional state and, in fact, provides the impetus for such change. For example, Acredolo, O'Connor, and Horobin (1989) suggested that uncertainty serves as the primary force underlying cognitive growth and that this uncertainty stems from the confusion children experience when they consider more than one rule. Similarly, any theory that posits internal conflict as a mechanism of developmental change (cf. Piaget's equilibration theory, 1975/1985) assumes that the impetus for transition comes from discrepancies in the rules a child uses to solve a problem; for these discrepancies to have an impact on the child's development, that child must have at some point considered and compared the rules he or she has available. For example, within the Piagetian tradition, Langer (1969), Snyder and Feldman (1977), and Strauss (1972; Strauss & Rimalt, 1974) argued that the child in transition displays at least two functional structures with respect to a concept; the child's appreciation of the discrepancy between those functional structures leads to disequilibrium, which then acts as an impetus for change (see Turiel, 1969, 1974, for similar arguments within the domain of moral development).

Even traditions that are distinctly non-Piagetian have proposed that multiple solutions to a problem may be a characteristic of the transitional state. For example, in his list of structure-dependent transition mechanisms, Keil (1984) included resolution of internal inconsistencies or contradictions as a mechanism of change. According to Keil, for a child to be internally inconsistent, the child must, at some level, entertain two (incompatible) views of the same problem. In his *skill theory* of cognitive development, Fischer (1980) described five rules that specify how a skill is transformed into a new, more advanced skill; each of these rules involves transforming two or more skills with given structures into one or more skills with a new type of structure, and thus each calls for activation of at least two skills for developmental change to occur. From an information-processing perspective, Klahr (1984) listed conflict-resolution rules—rules that apply when two productions are eligible to be activated on a single problem—as an important mechanism of change in self-modifying systems. Finally, from a Vygotskian perspective, Griffin and Cole (1985) argued that the zone of proximal development embodies multiple levels, both next steps and previous steps. The common thread running through all these characterizations of change is the notion that multiple, potentially incompatible hypotheses are activated or considered in solving a single problem.

Is there empirical evidence that children in transition with respect to a concept activate multiple hypotheses when solving problems instantiating that concept? In his studies, Wilkinson (1982a, 1982b) isolated a second group of children who produced correct and incorrect responses on a series of problems (and in this sense have partial knowledge) but who produced a different set of correct and incorrect responses when asked to solve the same problems a second time; that is, their knowledge was inconsistent across tests. Wilkinson called this a state of *variable partial knowledge* and suggested that one type of conceptual structure that could result in such a variable performance consists of a set of substitutable rules, some of which are correct and others of which are incorrect. On each problem, the child selects a rule from the set and applies it. The child's performance varies from problem to problem depending on which rule was selected for each problem and on whether the rule selected leads to a correct or incorrect solution. For example, a child might have a set of rules for problems of continuous quantity conservation that includes the following rules: (a) *the tallest container has the most,* (b) *the widest container has the most,* and (c) *the amount does not change when you pour it because you could pour it right back.* If the child in this example sampled the first or second rule, he or she would solve the conservation task incorrectly. However, if the third rule was sampled, the child would solve the task correctly.

Strauss (1972) and Strauss and Rimalt (1974) similarly argued that children who display at least two functional structures when solving problems instantiating a concept possess transitional knowledge with respect to that concept. Evidence of more than one functional structure on a series of tasks that instantiate a concept has been called *structural mixture.* Strauss (1972) distinguished between two types of structural mixture. A child who has structural mixture within a concept will be correct on some but not all of a series of tasks that instantiate a single concept, for instance, length conservation. A child who has structural mixture between concepts will be correct on some but not all of a series of tasks that instantiate a set of closely related concepts. For example, a child may succeed on discontinuous quantity (number) conservation but not succeed on length conservation.

Unfortunately, the multiple-hypotheses definition of the transitional state has not been adequately evaluated in terms of the readiness-to-learn criterion. If one uses this definition of the transitional state, one would expect children with variable partial knowledge (who are assumed to have a variety of substitutable rules at their disposal) to benefit from instruction more often than children with restricted partial knowledge (who are assumed to have only one rule available; Wilkinson, 1982a, 1982b). However, Wilkinson did not do the comparative training studies that could be used to address this issue. Moreover, studies purporting to show that structural mixture is a valid predictor of readiness to learn have not met Brainerd's (1977) methodological standards. For example, Strauss and Langer (1970) found that children who displayed structural mixture within a concept (conservation of continuous quantity) were more likely to advance to a sophisticated understanding of conservation than children who did not demonstrate structural mixture within the concept. Similarly, Langer and Strauss (1972) found that children who had structural mixture between concepts (conservation of discontinuous quantity and of length) were more likely to advance to a sophisticated understanding of conservation than children who did not demonstrate structural mixture between the concepts (see also Inhelder & Sinclair, 1969). However, these studies do not demonstrate that the children with structural mixture are more ready to learn than the children without structural mixture. In his reanalysis of the data from these studies, Brainerd (1977) showed that the children with structural mixture were no more likely to improve their understanding of conservation after training than were the children without structural mixture. They achieved a higher level of conservation understanding after training than children without structural mixture simply because they began the study at a higher level of conservation understanding.

The Transitional State as a State in Which Multiple Hypotheses Are Considered Simultaneously

Note that in studies of structural mixture, children are hypothesized to be in a transitional state if they use several different hypotheses when solving a series of problems that instantiate a single concept. However, the models of conceptual change described earlier (e.g., Fischer, 1980; Keil, 1984; Klahr, 1984; Piaget, 1975/1985) imply that what characterizes the transitional state is not merely the availability of more than one hypothesis but rather the simultaneous activation and evaluation of those hypotheses. To show that a child considers multiple hypotheses simultaneously, we must provide evidence that a child considers more than one hypothesis, not just across problems in a series but on a single problem.

Acredolo (Acredolo & O'Connor, 1991; Acredolo et al., 1989) argued that the techniques traditionally used to assess children's knowledge of concepts make it difficult to determine whether a child has considered more than one hypothesis on a single problem simply because the paradigms encourage the child to report a single answer. Acredolo et al. (1989) developed a new procedure to probe the child's understanding of conservation in which children were required to evaluate each of three alternative answers offered by the experimenter (actually offered by three Sesame Street puppets) on a probability scale, as opposed to selecting a single answer. They found that in a group of kindergarten through sixth-grade children, 33% displayed uncertainty (i.e., gave credence to more than one alternative) on the number conservation task and 74% displayed uncertainty on the area conservation task. In fact, many children who were classified as consistent nonconservers with standard conservation tasks appeared to regard conservation as a defensible alternative, whereas many children who were classified as consistent conservers with the standard tasks still seemed to regard nonconservation arguments as possibly correct. Acredolo et al. went on to suggest that children's implicit awareness of competing arguments serves as the primary force underlying cognitive growth.

Thus, if given the opportunity, children at times accept multiple (often incompatible) hypotheses about a single problem.

However, it remains unclear from these data whether the same children who can accept competing hypotheses actually generate more than one hypothesis at a time while solving and reasoning about a problem. We turn to this issue in the next two sections.

Evidence for Simultaneous Activation of Multiple Hypotheses in Explanations: Discordance Between Gesture and the Speech It Accompanies

When asked to explain a task, children often gesture while providing spoken explanations. Several authors (e.g., Evans & Rubin, 1979; Kendon, 1980; McNeill, 1987, 1992) have noted that these spontaneous gestures, taken by themselves, often convey substantive information about the task. Thus, gesture, like speech, can provide insight into the child's reasoning about a problem.

Over the past several years, we have analyzed the spontaneous gestures of children more closely, looking particularly at the congruence between the information contained in gesture and the information found in the speech it accompanies. We have observed that, when asked to explain their judgments of a conceptual task, some children produce gestures that convey the same hypothesis found in their spoken explanations, whereas other children convey one hypothesis in gesture and a different hypothesis in speech. We have termed this latter phenomenon *gesture–speech discordance* and have described its relationship to two different concepts typically acquired at two different stages in development.

Gesture–Speech Discordance in Conservation Tasks

Church and Goldin-Meadow (1986) tested children between the ages of 5 and 8 years on their understanding of a series of Piagetian conservation tasks. When asked to explain their judgments on the conservation tasks, the children gestured spontaneously while speaking and often portrayed specific aspects of the conservation task in their gestures; for instance, in a task probing conservation of liquid quantity, children often used a C-shaped hand to indicate the width of the dish, or they produced a fist and arced the fist from the glass to the dish as though pouring a pitcher to indicate that the water had been poured from the glass into the dish. At times the information conveyed in gesture matched the information conveyed in the speech accompanying the gesture. For example, in the task probing liquid quantity conservation, one child focused on the height of the water in both speech ("there is less water in the dish because the dish is short and the glass is tall") and in gesture (the child demarcated the heights of the two containers of water with his palm). This child had thus conveyed a single hypothesis—expressed in both speech and gesture—to explain his solutions on the task.

In contrast, some of the gestures produced by the children did not convey the same information as did the speech accompanying those gestures. For example, in the liquid quantity conservation task, one child focused on the height of the container in speech ("the dish is lower than the glass") but focused on the width of the container in gesture (the child produced a wide C hand near the dish and a narrower C near the glass). This child had conveyed two distinct hypotheses to explain her solutions on the task—one involving height (in speech) and another involving width (in gesture).

Gesture–Speech Discordance in Mathematical Equivalence

Perry et al. (1988) found a similar phenomenon on a task that is more abstract than conservation and that (unlike conservation) typically requires explicit school-based instruction to be acquired. Perry et al. tested children between the ages of 9 and 10 years on their understanding of equivalence in addition problems (i.e., the understanding that one side of an equation represents the same quantity as the other side of the equation). Children were asked to solve six problems of the form $5 + 3 + 4 = \_ + 4$ and to explain each of their solutions. Most children gestured spontaneously while explaining their solutions, and their gestures typically conveyed specific procedures for solving the problem. As in the conservation study, the procedure conveyed in gesture often matched the procedure conveyed in the speech accompanying that gesture. For example, one child indicated that he had added all of the numbers in the problem to get the answer, both in speech ("I added 5 plus 3 plus 4 plus 4 equals 16") and in gesture (the child pointed at the 5, pointed at the 3, pointed at the left 4, pointed at the right 4, and then pointed at the blank); that is, the child conveyed a single procedure, expressed in speech and in gesture.

However, again as in the conservation study, the gestures produced by the children did not always convey the same procedure as the speech that accompanied that gesture. For example, one child, in speech, indicated that he had added the numbers on the left side of the equation to get the answer ("I added 5 plus 3 plus 4") but, in gesture, indicated that he had considered all of the numbers in the problem (he pointed at the 5, the 3, the left 4, the right 4, and then the blank). This child conveyed two procedures, one that involved adding the numbers up to the equal sign (in speech) and a second one that involved adding all of the numbers in the problem (in gesture).

Gesture–Speech Discordance and Readiness to Learn

Both Church and Goldin-Meadow (1986) and Perry et al. (1988) found that the children who participated in their studies varied in the number of gesture–speech mismatches they produced, some producing no mismatches and some producing as many as six (out of a set of six explanations). In both studies, the children who produced many gesture–speech mismatches in their explanations of either conservation or mathematical equivalence (children whom we have labeled *discordant*) were more likely to benefit from instruction in that concept than the children who produced few gesture–speech mismatches or few gestures overall (children whom we have labeled *concordant*). Thus, the children who produced multiple hypotheses in a single explanation of a concept, one in speech and one in gesture, appeared to be particularly ready to learn that concept compared with the children who produced only one hypothesis in an explanation, either in speech alone or in both gesture and speech. These results suggest that children who simultaneously produce multiple hypotheses in their explanations of a concept are likely to be in a transitional state with respect to that concept.

It is important to stress that these studies of readiness to learn do, in fact, meet the methodological standards established by Brainerd (1977). First, the measure used to tap learning was not the same as the measure used to identify children in a transitional state. In both the math and the conservation studies, children were identified as transitional on the basis of the number of gesture–speech mismatches they produced on the pretest. Progress on the math test was measured in terms of the number of math problems solved correctly on the posttest and the generalization test, and progress on the conservation test was measured in terms of conservation rationales produced on the posttest that had not appeared in that child's pretest.[2] Second, the measure used to tap learning took into account the child's performance on the pretest. In the math studies, none of the children solved any of the problems correctly on the pretest; thus, their posttest scores were a measure of progress after training (Perry et al., 1988). In the conservation studies, learning was measured in terms of improvement from pretest to posttest; specifically, children were considered to have improved after training only if they added a new conservation rationale to their repertoires, one that they had not produced on the pretest (Church & Goldin-Meadow, 1986). Thus, in these studies, children were identified as transitional on the basis of gesture–speech mismatch; children who were transitional were then shown to be particularly ready to learn on the basis of a second measure, one that was independent of the original discordance measure.[3]

Further support for the hypothesis that gesture–speech discordance indexes the transitional state comes from the fact that the discordant state appears to be transitional, not only in the sense that it predicts receptivity to instruction but also in the sense that it is both preceded and followed by a concordant state. Alibali and Goldin-Meadow (in press) gave a different group of fourth-grade children instruction in mathematical equivalence and observed their explanations of the problems they solved over the course of the pretest and training period. The relationship between gesture and speech in each explanation was monitored over the series. Alibali and Goldin-Meadow found that as the children acquired the concept of mathematical equivalence, they progressed through a series of steps. Children began with a single incorrect procedure, reflected in the match between gesture and speech in their explanations of the math problems (a concordant incorrect state). They then proceeded through a transitional period in which more than one procedure was considered, reflected in the mismatch between the procedure expressed in speech and the procedure expressed in gesture (a discordant state). Finally, they closed on a single correct procedure, reflected again in the match between gesture and speech (a concordant correct state). These microgenetic data further suggest that the transitional period in the acquisition of a concept is characterized by the production of more than one hypothesis within a single explanation.

Note that we are putting forth gesture–speech discordance not as a general characteristic of the child but rather as a characteristic of the child's understanding of a particular concept. In other words, we do not expect discordance to be a communicative style that inevitably characterizes a child's explanations regardless of the concept that he or she is explaining. Indeed, we have found that children are likely to produce gesture–speech mismatches for a concept that they are in the process of learning (e.g., mathematical equivalence) but not for a concept that they have already mastered (e.g., conservation; Perry et al., 1988; see also Goldin-Meadow, Nusbaum, Garber, & Church, 1993, for evidence that a child can be discordant on one concept and concordant on another within the domain of mathematics itself). Thus, discordance appears to vary within the individual as a function of that individual's understanding of the concept, as we would expect if discordance is an index of the individual's readiness to learn that concept.

Evidence for Simultaneous Activation of Multiple Hypotheses in Solving Problems

The findings described thus far suggest that children in transition with respect to a concept simultaneously consider more than one hypothesis when explaining their beliefs about the concept. However, the fact that children may exhibit two hypotheses when explaining how they solved a problem does not necessarily mean that the children consider both hypotheses when actually solving the problems. Discordance could reflect post hoc reasoning processes rather than on-line problem solving. To explore this possibility, Goldin-Meadow et al. (1993) conducted a study to determine whether discordant children activate more than one hypothesis not only when they explain their solutions to a problem but also when they solve the problem itself. The approach underlying the study assumes that activating multiple hypotheses when solving a problem takes more cognitive effort than activating a single hypothesis. Thus, a child who activates multiple hypotheses on one task should

---

[2] As we have described, improvement in the conservation study was measured in terms of spoken conservation explanations—a measure that is not completely independent of gesture–speech mismatch simply because these explanations comprise the spoken component of a gesture–speech response. However, even when improvement is calculated in terms of judgment responses (i.e., the number of additional correct conservation judgments produced on the posttest and follow-up test relative to the number produced on the pretest), the results are precisely the same; that is, children who produced many gesture–speech mismatches on the pretest were significantly more likely to improve and maintain that improvement in conservation judgments than children who produced few gesture–speech mismatches (Church, 1987).

[3] To determine whether other possible differences between discordant and concordant children could better account for the differences between the children in success after training, Perry et al. (1988) did a multiple regression analysis using age, grade, math ability level, and discordance as variables to predict success on the posttest and the generalization test. They found that the proportion of variance explained by the combination of all four variables decreased significantly when discordance was removed from the regression but not when each of the other three variables (age, grade, and math level) was removed, suggesting that the discordance variable explained variance in success above and beyond the variance explained by the other three variables. Church (1987) did a comparable analysis for conservation (with age, number of spoken conservation rationales on the pretest, number of correct judgments on the pretest, and discordance as the four variables) and found similar results.

have less capacity left over to simultaneously perform a second task than a child who activates only a single hypothesis.

To test this prediction, Goldin-Meadow et al. (1993) first identified children as concordant or discordant with respect to mathematical equivalence on the basis of their explanations on a pretest. They next compared the concordant and discordant children's performance on two tasks: a math task (which contained problems testing the child's understanding of mathematical equivalence) and a word-recall task. The children were asked to solve each math problem while trying to remember a list of words. Goldin-Meadow et al. (1993) predicted that the discordant and concordant children would perform equally poorly on the math test but that the discordant children—if, in fact, they were activating two hypotheses on each math problem they solved—would expend more effort on the math task overall than the concordant children (who were expected to activate only one hypothesis on each problem). The discordant children would therefore have less capacity left over for the word-recall task than the concordant children and as a result would not perform as well as the concordant children on this task. This prediction was confirmed—the discordant and concordant children produced the same number of correct solutions on the math task (virtually none), yet the discordant children did not recall the word lists as well as the concordant children. The discordant children thus appeared to be working harder to solve the math problems incorrectly than were the concordant children.

These data suggest that the discordant children—the children who are in a transitional state—are working under increased cognitive demands. Thus, there may be a cost to being in a state of transition, a cost that could contribute to the instability of the transitional state. If learners in transition are provided with appropriate input, they might be expected to progress not only to a more stable state (in which only one hypothesis is activated per problem) but also to a more correct one. This is, in fact, what we have found in our training studies with respect to the acquisition of both conservation (Church & Goldin-Meadow, 1986) and mathematical equivalence (Alibali & Goldin-Meadow, in press; Perry et al., 1988). However, if learners in transition are not provided with appropriate environmental input and if the transitional state is indeed an unstable one, then the learners might be expected to regress to a more stable but incorrect state at least as often as, if not more often than, they progress to a stable correct state. Church (1990) charted spontaneous progress in the acquisition of conservation in a group of children over a period of several months and did, in fact, find that, without training, many of the children moved from an unstable incorrect state (in which two hypotheses were produced per problem) to one that was more stable (in which one hypothesis was produced per problem) but that was also incorrect (see also Alibali & Goldin-Meadow, in press).

In general, we take the Goldin-Meadow et al. (1993) findings to support our view that a child in a transitional state with respect to a concept not only expresses multiple hypotheses when explaining that concept but also simultaneously activates multiple hypotheses (demanding extra cognitive capacity) when solving problems instantiating the concept. Simultaneous activation of multiple hypotheses therefore appears to be characteristic of the transitional state, reflected both in the way children communicate about problems and in the amount of effort they expend in solving those problems. Thus, there is converging evidence from two distinct measures—discordance and capacity demand—supporting simultaneous activation of multiple hypotheses as a defining characteristic of the transitional state. Note that, in certain respects, the phenomenon of gesture–speech mismatch is no different from previously described phenomena wherein children give one response when their knowledge is tapped through, for example, judgments in the conservation task and give a different response when their knowledge is tapped through explanations (e.g., Brainerd & Brainerd, 1972; L. S. Siegel, 1978). However, there is one critical difference. A child who produces a gesture–speech mismatch is giving two different responses at the same moment; that is, the responses are activated simultaneously on the same problem. It is this simultaneity that we take to be the hallmark of the transitional state.

Different Types of Learning and Different Types of Transitional States

We have presented evidence suggesting that, at least for the acquisition of certain concepts, children progress through a transitional state in which they entertain multiple hypotheses about the concept (reflected in their gesture–speech mismatches) prior to mastering that concept. However, one might ask whether there must always be a period of transition in the acquisition of a concept. The model of transition we have proposed works well to characterize learning propelled forward by replacement mechanisms. The transitional state in our model is one in which a variety of substitutable procedures coexist, and the path out of this transitional state involves either endorsing one of these procedures and rejecting the others or introducing a completely new procedure. However, there are undoubtedly other concepts that are best described by an accumulation mechanism, and for which our model of transition is less appropriate. In mastering these concepts, children might build on a procedure acquired in an earlier period, experiencing little uncertainty as they do so because the new procedure incorporates the old one. For these concepts, there might be no period during acquisition that is identifiable as a distinct transitional state or, if a transitional state is identifiable, its characteristics might differ from the type of transitional state that we have described.

Our model may also be less appropriate for transitions in which knowledge becomes more efficiently processed but does not change in correctness (comparable to the changes that occur when adults progress from an unpracticed state to a more automatic and skilled state, e.g., Bryan & Harter, 1899; Logan, 1985; Shiffrin & Schneider, 1977). Wilkinson (1982a, 1982b) identified a group of children who appeared to have a correct rule for solving a problem yet applied that rule inconsistently. He considered this to be partial knowledge, in the sense that the children's performance included both correct and incorrect responses, and characterized that knowledge in terms of an underlying cognitive structure consisting of a collection of modular components that together make up a single, correct rule for solving problems. Each component considered separately is adequate for its role in the larger rule; however, the components are not consistently integrated and applied. On some problems the child applies the rule correctly, and on others the child applies the rule incorrectly, resulting in variable performance across a single set of problems and across multiple problem solving assessments. One might imagine that the transition out of such a state would be quite different from the transition out of a state in which the child entertains multiple hypotheses. To progress out of a state in which the child has an unintegrated but basically correct rule, all the child need do is gain expertise in applying that rule as a whole or in applying components of the rule. The child need not entertain alternative rules to improve performance because the rule itself is essentially correct.

Nevertheless, it is possible that the relationship between gesture and speech can provide evidence for this type of transitional state (where the child has an essentially correct procedure that has yet to be integrated) as well as for the transitional state we have described in this article (where the child entertains multiple procedures, some correct and some incorrect). Children who possess a correct procedure that is not yet integrated might produce matching information in their simultaneously gestured and spoken explanations of a problem, but the two modalities might not be synchronized in terms of timing. For example, in response to the problem, $6 + 7 + 4 = \_ + 4$, a child might point to the $6 + 7$ while saying, "What I did was. . ." and then point to the solution while saying "added $6 + 7$." Although the substantive information contained in speech and gesture is identical (and correct), the lack of synchronization between what is said and what is gestured suggests that the procedure may not yet be a smoothly functioning unit.

It is also possible that certain types of gesture–speech mismatches might provide evidence for an overarching rule with components that are correct but not yet integrated into a whole. For example, in our math studies (Alibali & Goldin-Meadow, in press; Perry et al., 1988), the children at times produced explanations containing one correct procedure in speech and a different correct procedure in gesture. We assume that, eventually, the child will come to understand the relationships among the various correct procedures and unite them all into one coherent mathematical system. At this stage, however, the child appears to have only the pieces of the system. Thus, gesture may provide insight into the structure of different types of transitional knowledge, particularly those transitions in which two beliefs are simultaneously activated.

Gesture–Speech Discordance and the Transitional Knowledge State

We have suggested that the mismatch between gesture and speech may be a more sensitive index of transitional knowledge than indices that rely solely on patterns of correct and incorrect solutions to problems. Such indices may overlook children who fail to solve any problems correctly but who have more than one reason for doing so and are nevertheless ready to benefit from instruction (i.e., who have transitional knowledge). For example, none of the children solved any of the math problems correctly in Perry et al.'s (1988) study. Nevertheless, a subset of these children—the subset who produced two hypotheses in their explanations, one in speech and a different one in gesture—were particularly likely to improve on the task after instruction. Moreover, unlike most other indices of transitional knowledge, the mismatch between gesture and speech not only identifies children who are in a transitional state but also provides substantive information about which specific hypotheses children are actually entertaining. Thus, the mismatch between gesture and speech appears to be a robust, sensitive, and informative index of transitional knowledge.

We have furthermore suggested that gesture–speech mismatch can serve as a good index of the transitional knowledge state precisely because it is integral to the state itself. Indeed, we suggest that gesture–speech mismatch is caused by the processes that characterize transitional knowledge. Specifically, we hypothesize that it is the simultaneous activation of multiple hypotheses that characterizes the transitional knowledge state and that creates gesture–speech mismatch. In the next section, we explore in detail the processes that generate the mismatch between gesture and speech.

The Mechanism of Gesture–Speech Mismatch

In our studies, a child is asked to solve a problem and then explain that solution. The child describes a procedure for arriving at a solution and from this procedure we make inferences about that child's understanding of the problem. For example, we assume that a child who says he or she solved the math problem, $4 + 5 + 3 = \_ + 3$, by "adding the 4, the 5, the 3, and the 3" has a representation of the problem that includes all four numbers, with no meaningful subgroupings within the numbers. In contrast, a child who says he or she solved the same problem by "adding the 4, the 5, and the 3" is assumed to have a representation of the problem that includes only those numbers on the left side of the equal sign. Thus, the procedure the child describes provides insight into the way in which that child represents the problem.

Note that by observing both gesture and speech, we have two different access routes to the child's representation, one through the procedure that the child articulates in speech and a second through the procedure that the child describes in gesture. In concordant children, the two access routes provide evidence for the same representation because, by definition, concordant children tend to produce in gesture the same procedures that they produce in speech. In contrast, in discordant children, the two access routes provide evidence for two different representations, one representation accessed by gesture and a second representation accessed by speech. Thus, discordant children appear to be working with two different representations of the same problem.

Do Children Have Representations That Are Accessible to Only One Modality?

The question we now address is whether all of a child's representations are accessible to both modalities (i.e., reflected, at some point in the child's set of responses, in gesture as well as in speech) or whether some are accessible to only one modality (i.e., reflected in gesture but not speech or in speech but not gesture). It is important to recognize that the definition of discordance does not entail that children have representations that can be accessed by one modality and not the other. Discordance is determined on the basis of the number of gesture–

speech mismatches a child produces. Note that a child can produce mismatches by producing, on the first problem, one procedure in gesture (e.g., adding all the numbers in the problem, or add-all) and a different procedure in speech (e.g., adding the numbers on the left side of the equal sign, or add-to-equal) and then, on the second problem, reversing the pattern (i.e., producing add-to-equal in gesture and add-all in speech). If a child achieved a large number of gesture–speech mismatches in this way, that child would not have any procedures that appeared only in one modality. Such a child would be assumed to have two different representations of the problem (one reflected in the add-all procedure and another reflected in the add-to-equal procedure), with both representations accessible to gesture and to speech.

In contrast, a child could produce a large number of gesture–speech mismatches by producing one procedure in gesture on all of his or her responses and a different procedure in speech on those same responses; for example, the child might produce add-all in gesture and add-to-equal in speech on all of his or her responses. Such a child would also have two different representations of the problem, but each of those representations would be accessible to only one modality.

To determine whether children have representations that are accessible to one modality and not the other, we need to examine the entire set of responses that a child produces—what we have called the child's *repertoire* of procedures—and to determine whether a particular procedure appears in one or both modalities across that repertoire. To this end, we examined the repertoire of responses produced before instruction by each of the 58 children in Alibali and Goldin-Meadow's (in press) study.[4] For each set of six responses in a child's pretest, we determined how many procedures were expressed in gesture and never in speech, how many were expressed in speech and never in gesture, and how many were expressed in both gesture and speech (procedures did not have to be expressed in gesture and speech on the same problem to be considered part of the gesture + speech repertoire; the procedure had only to appear sometime in gesture and sometime in speech). We found that the children did indeed have procedures that they produced in one modality and not the other and, in fact, they had slightly more procedures of this type in their repertoires (1.56, *SD* = 1.30) than procedures produced in both modalities (1.31, *SD* = 0.54). Interestingly, the mean number of procedures found in gesture but not speech (1.28, *SD* = 1.17) was larger than the mean number found in speech but not gesture (0.28, *SD* = 0.59).

These data suggest that the children in this study had a relatively large number of representations in their repertoires that were accessible only to one modality, primarily accessible only to gesture. Note that if a representation is accessible only to one modality, whenever a child attempts to articulate a procedure based on that representation, the child will not be able to produce the same procedure in both gesture and in speech; that is, the child will of necessity produce a gesture–speech mismatch. Thus, we would expect that discordant children (who by definition produce a large number of gesture–speech mismatches) would have more representations in their repertoires accessible to only one modality, and therefore more procedures found in only one modality, than concordant children (who produce relatively few mismatches). This, in fact, is what we found. We analyzed the Alibali and Goldin-Meadow (in press) data using analysis of variance with repeated measures, with group (discordant vs. concordant) as the between-subjects factor and modality in which a procedure was produced (in gesture but not speech, in speech but not gesture, or in both gesture and speech) as the within-subjects factor, and found that the crucial interaction between the two factors was indeed significant, $F(2, 56) = 13.811$, $p < .001$. Specifically, we found that the discordant children produced significantly more procedures unique to gesture than did the concordant children: The 35 discordant children produced an average of 1.77 procedures ($SD = 1.19$) in gesture but not speech compared with the 0.52 procedures ($SD = 0.59$) produced by the 23 concordant children, $F(1, 56) = 36.337$, $p < .001$. Interestingly, the two groups of children did not differ in the mean number of procedures produced in both gesture and speech: 1.20 ($SD = 0.53$) for the discordant children compared with 1.48 ($SD = 0.51$) for the concordant children, $F(1, 56) = 1.802$, $p > 0.10$. Moreover, they did not differ in the mean number of procedures produced in speech but not gesture: 0.40 ($SD = 0.70$) for the discordant children compared with 0.09 ($SD = 0.29$) for the concordant children, $F(1, 56) = 2.28$, $p > 0.10$, although the procedures found only in speech accounted for a very small part of the repertoire for both groups.

Note that if the number of procedures found in both gesture and speech is the same for the discordant and concordant children, and if the discordant children have more procedures found only in gesture than the concordant children, then the discordant children will have more procedures in their repertoires overall (3.37, *SD* = 1.24) than the concordant children (2.09, *SD* = 0.85). This means, in effect, that the procedures in a child's repertoire found in both gesture and speech (i.e., the procedures that can, at least in principle, result in a match between the gestural and spoken modalities) account for a greater proportion of the total procedures in a child's repertoire for the concordant children (1.48/2.09 = 0.71) than they do for the discordant children (1.20/3.37 = 0.36). In other words, the concordant children have proportionately more procedures found in both gesture and speech than do the discordant children, suggesting that the concordant children also have proportionately more representations that are accessible to both modalities than do the discordant children.

These data thus provide hints as to how gesture–speech mismatch might come about. We suggest that, when faced with a problem, children sample a representation of how to solve the problem and, on the basis of that representation, attempt to describe a procedure for solution. If a child samples a representation that is accessible to both gesture and speech, our model proposes that the child will express the same procedure in both modalities, thus producing a gesture–speech match. If, however, the child samples a representation that is accessible to gesture but not to speech, the child will be able to describe the

---

[4] We excluded from these analyses 5 of the 63 children in the Alibali and Goldin-Meadow (in press) study because they did not gesture at all on the pretest. We return to the issue of nongesturers in a subsequent section.

procedure in gesture but will be unable to express the same procedure in speech. In that case, the model proposes that the child will then select another representation, one that is accessible to speech, and therefore will produce a gesture–speech mismatch. Thus, this model assumes that when a representation that is accessible to both gesture and speech is the first representation sampled, both gesture and speech will be activated. In this sense, the model is consistent with McNeill's (1992) description of gesture and speech as an integrated system.

If this model is correct, the probability of producing a gesture–speech match on any given problem ought to be equal to the probability that a representation that is accessible to both gesture and speech will be sampled. Our next step is to evaluate this model with respect to the data. Using the pretests for each of the 58 children in the Alibali and Goldin-Meadow (in press) study, we calculated the number of gesture–speech matches each child would be expected to produce assuming this model. We took the number of different procedures that a child produced both in gesture and in speech as an estimate of the number of representations accessible to both modalities for that child. Next, we took the total number of different procedures that a child produced in either gesture or speech as an estimate of the total number of representations for that child. We then calculated the child's probability of producing a gesture–speech match by dividing the number of different procedures found in both gesture and speech by the total number of different procedures the child produced in either gesture or speech. For example, if a child produced two procedures in both gesture and speech and a third procedure in gesture alone, the probability that this child would produce a gesture–speech match would be found by dividing 2 (the number of different procedures in both gesture and speech) by 3 (the total number of different procedures), yielding .67. Thus, .67 of the six responses this child gave (i.e., four) would be expected to be gesture–speech matches and, because this represents over 50% of the child's responses, the child would be predicted to be concordant.

Calculating the number of gesture–speech matches each child was expected to produce in this fashion, we found that 20 of the 58 children were expected to be concordant and 38 were expected to be discordant—a distribution quite close to, and indeed not significantly different from, the 23 concordant and 35 discordant children actually observed, $\chi^2(1) = 0.687$, $p > .40$.[5] The model therefore predicted the observed distribution of concordant and discordant children quite well. Moreover, the model accurately predicted the precise number of matches the children produced. The children varied from 0 to 6 in the number of matches they were observed to produce, and the model predicted a distribution of matches that was not significantly different from this observed distribution ($D = 0.103$, $p > .20$, Kolmogorov-Smirnov One-Sample Test, S. Siegel, 1956). Thus, this model appears to fit the data quite well.

The model we have proposed assumes that the child samples representations that are then encoded into either gesture or speech, or both. According to the model, a child has a single set of representations, some of which are accessible to both gesture and speech and some of which are accessible only to gesture or only to speech. Furthermore, the model assumes that when a representation that is accessible to both gesture and speech is the first representation sampled, both gesture and speech will be activated. In this sense, gesture and speech can be said to form an integrated system. An alternative hypothesis would propose that the child has two distinct sets of representations, one set accessible to gesture and a second set accessible to speech. When asked to explain a problem, this alternative model predicts that a child samples a representation accessible to gesture and independently samples a representation accessible to speech. According to this model, if a child is to produce a gesture–speech match, that child will do so by randomly sampling a representation from the pool of representations accessible to gesture and, by chance, randomly sampling that same representation from the pool accessible to speech. It would, of course, strengthen our argument considerably if we could show that this alternative model, which assumes independence between gesture and speech, does not predict the data as well as our model, which assumes an integrated system between gesture and speech. We therefore evaluated this alternative model with respect to the data.

Using the same data from the Alibali and Goldin-Meadow (in press) study, we recalculated the number of gesture–speech matches a child would be expected to produce if it is assumed that the child samples representations accessible to gesture independent of sampling representations accessible to speech. We took the number of different procedures that a child produced in gesture as an estimate of the number of representations accessible to gesture for that child, and we did the same for speech. We then calculated the child's probability of producing a gesture–speech match as follows. Consider again the child who produces two procedures in gesture and in speech and a third procedure in gesture alone. If add-all and add-to-equal are the two procedures in the child's gesture + speech repertoire, the probability that this child will produce a gesture–speech match is equal to the probability of producing add-all in both gesture and speech, plus the probability of producing add-to-equal in both gesture and speech. The probability that the child will produce the add-all procedure in both modalities is determined by multiplying the probability of producing add-all in gesture (1 out of 3, the total number of procedures found in gesture) times the probability of producing add-all in speech (1 out of 2, the total number of procedures found in speech), yielding .17. The probability of producing the add-to-equal procedure in both gesture and speech is also .17, making the total

---

[5] As described earlier, we determined a child's repertoire of procedures on the basis of the responses that child gave to the six pretest problems. Note that if a child had three procedures in his or her repertoire found only in gesture, that child would produce, at a minimum, three gesture–speech mismatches regardless of the sampling process. To determine whether the number of procedures found only in one modality can, by itself, predict the distribution of concordant and discordant children, we calculated the number of gesture–speech mismatches in the Alibali and Goldin-Meadow (in press) data that would be expected on this basis. We found however that, under this hypothesis, 49 of the 58 children would be expected to be concordant and 9 would be expected to be discordant—a distribution significantly different from the 23 concordant and 35 discordant children actually observed, $\chi^2(1) = 88.907$, $p < .001$. Thus, this model does not adequately describe the data.

probability of producing the same procedure in gesture and in speech .34. Thus, .34 of the six responses this child gives (i.e., two) would be expected to be gesture–speech matches, and the remaining four responses would be expected to be mismatches; the child would therefore be expected to be discordant.

Calculating the number of gesture–speech matches each child was expected to produce in this fashion, we found that the alternative model predicted that 10 of the 58 children would be concordant and that 48 would be discordant—a distribution significantly different from the 23 concordant and 35 discordant children actually observed, $\chi^2(1) = 20.421$, $p < .001$. Moreover, unlike our initial model, this alternative model did not accurately predict the number of matches that the children were observed to produce ($D = 0.328$, $p < .01$, Kolmogorov-Smirnov One-Sample Test). Thus, the alternative model does not describe the data as well as the initial model, suggesting that representations accessible to gesture are not sampled independent of representations accessible to speech. The data therefore lend credence to the initial model we proposed—a model based on the premise that gesture and speech form one integrated system rather than two independent systems.

Moreover, the initial model we proposed also accounts very well for the observed variability in the production of gesture–speech mismatches; that is, it can explain why discordant children produce gesture–speech mismatches on some but not all of their explanations. The notion of sampling from a repertoire is not new to theories of learning. In fact, many models of learning (e.g., Brainerd, 1979; Wilkinson & Haines, 1987) call on the notion of sampling from a repertoire of both correct and incorrect rules to account for variability in performance over a set of problems. What differentiates our model from these other sampling models is that our model involves sampling representations, which are then encoded into gesture or speech; if a particular representation cannot be encoded into both gesture and speech, a second representation is sampled. Consequently, in our model, more than one representation can be sampled on the same problem, resulting in the possibility that more than one procedure will appear in a single response.

Taken in conjunction with our training studies, the analyses we have described thus far suggest that a child who is on the verge of learning a concept is likely to have a relatively large number of representations of that concept that are accessible to gesture and not to speech. Thus, we propose the following description of the steps a learner follows in acquiring a concept. The learner begins the acquisition process with incorrect representations of the concept, most of which are accessible to both the gestural and spoken modalities. The learner then acquires correct representations that are accessible only to gesture and not to speech. It is at this moment that the learner is in the transitional state with respect to this concept and is most open to instruction in that concept. Finally, the learner develops a verbal code for the correct representations that were once accessible only to gesture, thus returning once again to a state in which most of his or her representations are accessible to both gesture and speech (this time, however, the representations are correct).

One obvious implication of this view of the learning process is that when a learner spontaneously acquires correct representations, procedures based on those representations are likely to be found in the learner's gestural repertoire. Indeed, we found that, before instruction, 34 of the 58 children in the Alibali and Goldin-Meadow (in press) study produced some correct procedures and all of those procedures were found in gesture (26 produced correct procedures found only in gesture and 8 produced correct procedures found in both gesture and speech). None of the children produced correct procedures that were found only in speech. Thus, as expected, the correct procedures that the children produced all appeared in the gestural repertoire, and correct procedures appeared in the spoken repertoire only when they also appeared in the gestural repertoire.

Does the Child's Gestural Repertoire Reflect Implicit Knowledge?

The data we have described suggest that for children who are in a transitional state, gesture has access to a larger and more advanced pool of representations than speech. In other words, children in a transitional state have an implicit understanding of a larger and more correct set of representations than they can explicitly articulate in speech. If the knowledge accessible to gesture is indeed implicit knowledge, one might be able to tap that knowledge with techniques traditionally used to tap implicit knowledge, for example, a recognition technique (cf. Broadbent, 1991). A child who has a representation that is accessible to gesture but not to speech will not be able to produce a procedure based on that representation in speech; however, such a child might still be able to recognize a solution generated by that procedure as an acceptable solution to the problem.

To test this possibility, in collaboration with Philip Garber (Garber, Alibali, & Goldin-Meadow, 1992), we gave children the six math problems used in the Perry et al. (1988) and Alibali and Goldin-Meadow (in press) studies and determined, on the basis of their explanations of how they solved those problems, which procedures the children possessed in their spoken and gestural repertoires. We focused on those children who solved all six problems incorrectly and gave each child a judgment task in which the child was asked whether a solution generated by a particular procedure was an acceptable solution to a given math problem. The children were told that more than one solution might be possible. We used the six procedures most commonly found in children's explanations of math problems of this type (cf. Perry et al., 1988) and presented each child with six different solutions, each generated by a different procedure. For example, for the problem, $4 + 3 + 5 = \_ + 5$, on one trial, the child was asked whether 17 was an acceptable response to the problem (17 reflects a procedure in which all of the numbers in the problem are added to produce a solution); on another trial, the child was asked whether 12 was an acceptable response (12 reflects a procedure in which the numbers to the left of the equal sign are added to produce a solution). Each child saw 6 different math problems and was asked to judge the acceptability of six solutions for each problem, resulting in 36 problems presented in blocks of 6 (i.e., the six solutions for a given problem were presented in a block but were randomly presented within that block). For each problem, the child was asked if the solution was "definitely" an acceptable response to the problem, "maybe" an acceptable response, "maybe not" an acceptable response, or "definitely not" an acceptable response.

Looking first at the children's repertoires of procedures as determined by their pretest explanations, we found that 19 of the 20 children in the sample produced at least one procedure in both gesture and speech.[6] Indeed, for 8 of those 19 children, all of the procedures produced were found in both gesture and speech. For the remaining 11 children, in addition to having procedures found in both gesture and speech, at least one of their procedures was found in only one modality (7 had some procedures found only in gesture, 2 had some procedures found only in speech, and 2 had some procedures found only in gesture and others found only in speech).

To determine whether the modality in which a procedure appeared on the pretest was related to the likelihood that the procedure would be considered an acceptable response on the judgment task, we calculated for each child the mean acceptance score for each of the six procedures. If the solution generated by a procedure was judged to be definitely acceptable, the procedure was assigned a score of 4; if the solution was judged maybe acceptable, the procedure was assigned a score of 3; if the solution was judged maybe not acceptable, the procedure was assigned a score of 2; and if the solution was judged definitely not acceptable, the procedure was assigned a score of 1. We then calculated for each child the average acceptance score for groups of procedures: procedures that had not appeared anywhere in a child's pretest explanations, procedures that had appeared in both gesture and speech on the pretest, and (in those children who had them) procedures that had appeared only in gesture on the pretest.

We found that procedures produced in both gesture and speech were accepted more frequently than procedures found in neither gesture nor speech. Focusing on the 8 children who produced all of their procedures in both gesture and speech, we found that the mean acceptance score for the procedures that these children produced on the pretest was 2.79 (SD = 0.90)—significantly higher than the acceptance score for the procedures that the children had not produced on the pretest (1.69, $SD = 0.41$), $t(7) = 2.44$, $p < .05$. This result is unsurprising but suggests, at the least, that the acceptance score has some validity as a measure of the child's knowledge.

We next turned to the question of interest: whether solutions generated by procedures that were found only in gesture on the pretest would be recognized as acceptable responses on the judgment task. To explore this question, we turned to the 9 children who produced some procedures on the pretest that were found only in gesture. We found, in fact, that the children had significantly different acceptance scores for procedures that they had not produced in either gesture or speech on the pretest versus those that they had produced in gesture only versus those that they had produced in both gesture and speech, $F(2, 16) = 21.440$, $p < .001$. In particular, the 9 children were significantly more likely to accept a procedure that they had produced in gesture only on their pretest (2.01, $SD = 0.88$) than they were to accept a procedure that they had not produced in either gesture or speech (1.41, $SD = 0.46$), $F(1, 16) = 5.440$, $p \leq .05$. In addition, the children were significantly less likely to accept a procedure that they produced in gesture only (2.01, $SD = 0.88$) than they were to accept a procedure that they produced in both gesture and speech (3.09, $SD = 0.69$); $F(1, 16) = 17.080$, $p < .005$. In other words, the solutions generated by procedures produced in gesture only were accepted at a level in between procedures produced in neither gesture nor speech and procedures produced in both gesture and speech. These data suggest that having a procedure in one's gestural repertoire does indeed reflect an implicit awareness of that procedure—an awareness that can be tapped by a recognition task.[7]

Thus, the knowledge that appears to be accessible to gesture but not to speech is implicit in the sense that it can be tapped by a recognition technique. Implicit knowledge is often equated with knowing how to perform a task without knowing how to articulate that knowledge (cf. Broadbent, 1991). It is important to note, however, that the knowledge that is accessible to gesture and not to speech differs from knowledge that is traditionally considered implicit in that children can often demonstrate an awareness of a correct procedure in gesture before they can perform correctly on the task. In the Goldin-Meadow et al. (1993) study, all of the discordant children produced some correct explanations and all but one produced those explanations in gesture but not speech. However, the children did not produce correct solutions on the task. Thus, the children solved the task incorrectly (producing an incorrect verbal explanation reflecting that solution) but, at the same time, produced a correct gestural explanation, suggesting that gesture taps a level of knowledge that has not yet been integrated into task performance.

Note that although the procedures reflected only in gesture do not appear to control the actual solution that the child produces, the procedures reflected in gesture do appear to put demands on working memory (Goldin-Meadow, Nusbaum, Garber, & Church, 1993). In addition, the procedures found in gesture also appear to set the agenda for future development. For example, in their training study of conservation, Church and Goldin-Meadow (1986) found that a majority of the children who produced a conservation explanation in speech for the first time on the posttest had produced that same explanation in gesture on the pretest. Thus, what the children said with their hands before training appeared to be what they were most likely to learn during training, suggesting that, although the knowledge expressed in gesture may be implicit, it still has an effect on behavior.

Representation of Information in Gesture and Speech

We have found that, at a certain point in the learning process, the set of representations in a child's repertoire that is accessible to gesture is not identical with the set of representations that is

---

[6] We eliminated 3 children from the sample because they produced no gestures at all (see later for a discussion of nongesturers) and 2 children because they did not appear to understand the judgment task (they judged all of the solutions to be unacceptable).

[7] Note that these data also suggest that if children produced a procedure in both gesture and speech, they were more likely to recognize that procedure than if they produced the procedure only in one modality. The small number of children who produced procedures only in speech prevented us from exploring whether this pattern was true even if the modality was speech, that is, whether children were more likely to recognize a procedure found in both their gestural and spoken repertoires than a procedure found only in their spoken repertoire.

accessible to speech. For such a child, gesture does not encode precisely the same information as speech. Why should there be such a disparity between the two modalities?

McNeill (1992) argued that gesture and speech are two aspects of a single process. The two sides are correlated in meaning but do not always reveal the same meaning. According to McNeill, gesture reflects a global–synthetic image. It is idiosyncratic and is constructed at the moment of speaking—it does not belong to a conventional code. In contrast, speech reflects a linear-segmented, hierarchical linguistic structure, using a grammatical pattern that embodies the language's standards of form and drawing on an agreed-on lexicon of words.[8] A particular syntactic frame may offer no convenient slot for information that is contained in a speaker's image of an event; such information may therefore be left to gesture. For example, consider a speaker whose image of relocating a book includes not only the notion of transfer but also the relative locations of the book before and after the move. Note that if the speaker describes this event with the phrase "place the book on the table," he or she would adequately capture in words the notion of transfer but would not include information about where the book initially came from. Indeed, the verb *place* does not easily allow incorporation of the source (i.e., it is awkward, if not ungrammatical, to say "place the book on the table from the shelf"). However, even though this particular syntactic construction does not easily allow the source to be specified, information about the source could quite easily be incorporated into gesture. For example, if a gesture moving from a spot above the head down toward the waist were to accompany the phrase, one might infer that the book had been moved from a relatively high spot (perhaps a shelf) to a lower spot (the table). Thus, for a given verbal construction, certain types of information may be more easily encoded in gesture than in the speech itself, and vice versa.

The notion that certain information is not easily described in words is not new. Huttenlocher (1973, 1976) argued convincingly that it is not useful, or even possible, to represent all aspects of human experience in natural language. For example, a map can represent the shape of the east coast of the United States accurately and to scale. Not only would a verbal description of this coastline take many pages, but it is not obvious that it could, in principle, provide all the information that the map does. Thus, Huttenlocher argued that there must be representational systems that do not involve words that humans use to encode information. Similarly, Anderson (1983) and Johnson-Laird (1983) each proposed a variety of representational systems (including minor ones that involve imagery) as options for encoding information. Following McNeill (1992), we suggest that gesture is a vehicle for one of these options. Gesture offers children (and adults, for that matter) a vehicle for expressing their understanding of a problem that is distinctly different from speech.

Moreover, our data suggest that, for certain problems and at certain times in the learning process, this vehicle may be better suited to capturing a child's understanding of a problem than is speech. In other words, even though a child's grasp of a problem might, in principle, be encodable into words, the child may be incapable of verbally expressing an understanding of the problem at a moment when the child is capable of expressing that understanding in gesture. Why might this be so? The two domains that we have extensively explored, conservation and mathematical equivalence, both involve some spatial reasoning. Indeed, Hadamard (1945) argued that mathematical thinking, particularly innovative mathematical thinking, is not conducted in words but rather in spatial images—images that might be more easily translated into the global–synthetic representation characteristic of gesture than into the linear-segmented representation characteristic of speech (cf. McNeill, 1992). It is, in fact, quite clear that children have a rich array of quantitative abilities prior to developing conventional verbal methods of operating on quantities (e.g., Levine, Jordan, & Huttenlocher, 1992), and it is possible that those abilities could be better expressed in a modality that more closely maps the phenomenon.

With respect to the two concepts we have studied extensively, gesture may indeed provide a more accessible code for the knowledge a child possesses than does speech. For example, at a certain point in the acquisition of conservation, children may recognize that the tall, skinny beaker in the task differs from the short, wide beaker on two dimensions, height and width. It may be easier to encode those two dimensions in gesture than to explicitly label and articulate the relationship between the two dimensions in speech (e.g., children can, and do, indicate the height of the beakers with two flat palms, each placed at the water level of one of the beakers, and they indicate the width of the beakers with two curved hands, one mirroring the width of the skinny beaker and one mirroring the width of the wide beaker). Similarly, with respect to the acquisition of mathematical equivalence, it may be easier to convey that the two sides of the equation have equal status in gesture (by sweeping a flat palm under the left half of the equation and then making precisely the same movement under the right half of the equation) than in speech.

In general, an image may be more easily encoded in gesture than in speech precisely because gesture allows one to represent the image as a whole, without breaking it into parts. Indeed, children's failure to encode a notion in speech may indicate that they have not yet decomposed their holistic understanding of a concept into parts; until they do so, although they may have an image of how the parts fit together (an image that gesture is adept at conveying), they may not yet have an understanding of the individual parts and the relationships among them. In this view, the translation of an image into speech is a

---

[8] Note that McNeill's (1992) characterization does not imply that the manual modality cannot, under other circumstances, assume a languagelike structure akin to speech. Indeed, conventional sign languages of the deaf such as American Sign Language are structured in the same linguistic ways as speech (e.g., Klima & Bellugi, 1979) and, in fact, are quite distinct from the gestures that accompany the speech of hearing individuals (McNeill, 1992). Moreover, when gestures are used as the sole means of communication (i.e., without speech) by deaf children who know no conventional sign language (Goldin-Meadow & Mylander, 1983, 1984, 1990) or by hearing individuals in an experimental situation (Singleton, Goldin-Meadow, & McNeill, in press), these gestures take on aspects of languagelike structure and thus no longer resemble the spontaneous gestures that accompany speech (see Goldin-Meadow, 1993, for further discussion).

matter not merely of learning the appropriate verbal labels but rather of analyzing and resynthesizing the notion itself. This translation process is akin to the notion of redescription proposed by Karmiloff-Smith (1986) to account for progression from a state in which knowledge is implicitly grasped to one in which it is grasped consciously and explicitly.

Note that, although we have found that correct explanations are produced in gesture before speech in children acquiring both conservation and mathematical equivalence, our results leave open the possibility that speech could anticipate gesture in other (perhaps less spatial) domains. For example, Goodman et al. (1991) found that gesture and speech do not always match in children's and adult's responses to Kohlberg's moral reasoning tasks (Colby, Kohlberg, Gibbs, & Lieberman, 1983). It is certainly possible that, because moral reasoning is more culturally and socially bound than mathematical reasoning, talk might be essential to acquiring the concept. In this case, we might not expect gesture to have privileged access to initial insights into the domain, and advances in reasoning might well appear first in speech rather than gesture.

The Function of Gesture–Speech Mismatch in Transition

We have shown that gesture–speech mismatch in a child's explanations of a concept reflects the fact that the child is in a transitional state with respect to that concept. Thus, the mismatch between gesture and speech can serve as an index that experimenters may use to identify and characterize children in transition. One might also ask whether gesture–speech mismatch has significance, not only for the experimenter but also for learners themselves. In other words, what role (if any) does gesture–speech mismatch play in the mechanism of cognitive change?

Does Gesture–Speech Mismatch Facilitate Transition?

Some investigators (see Karmiloff-Smith, 1985) have suggested that externalizing beliefs in two modalities, for example, gesture as well as speech, may help the child pull back and simultaneously evaluate those beliefs, resulting in reorganization of the beliefs into a single system. For example, if a child recognizes (either explicitly or implicitly) that the information contained in gesture conflicts with the information in the accompanying speech, or that the notions conveyed in gesture and speech could be consolidated to create a more efficient system of representation, that child may be compelled to reorganize his or her thinking. In this way, gesture–speech mismatch may serve as both a signal of uncertainty and a vehicle for its resolution. In a similar vein, McNeill (1992) argued that the act of gesturing itself can affect thought. According to McNeill, some dimensions of thought are presented in gesture and others are presented in linguistic form. There is a synthesis at the moment of speaking when language and gesture are combined into one unified presentation of meaning. This is an act of communication, but it is also an act of thought in which speakers themselves are affected.

If, in fact, gesture plays a role in shaping thought at the moment of speaking, the act of producing a gesture–speech mismatch might itself play a role in cognitive change. On the other hand, it seems naive to argue that one must gesture to learn. Indeed, in screening the subjects for their study, Alibali and Goldin-Meadow (in press) found that many of the children tested did not gesture during either the pretest or the training session. Nevertheless, some of these nongesturers did acquire the concept after training. These data lead us to ask how necessary gesture–speech mismatch is to the learning process.

This is a difficult question to answer primarily because a child who fails to gesture will obviously not produce gesture–speech mismatches. What we need is a technique that identifies children who would produce mismatches if they were to gesture. One possibility is to construct a situation that encourages gesture. Another possibility is to bypass gesture altogether and find a measure that identifies potential mismatchers without relying on gesture. Recall that Goldin-Meadow et al. (1993) showed that children who frequently produce gesture–speech mismatches in their explanations of a concept (thereby displaying two procedures in their explanations) are precisely the children who show an effect of cognitive load when solving problems instantiating the concept; that is, they expend a considerable amount of effort solving just those problems on which they had previously produced mismatched explanations, leading us to infer that they activate two procedures on these problems when they solve them as well as when they explain them. Thus, the cognitive-load technique could be used to identify children who activate two procedures per problem, and it is these children whom we would expect to produce mismatches if they were to gesture.

Using the cognitive load technique, we can first determine whether any nongesturers display an effect of cognitive load (i.e., expend a considerable amount of effort when solving the relevant problems). If so, we then assume that these children activate two procedures when solving problems of this type and thus, in this sense, are in a transitional state. If gesture–speech mismatch is itself necessary for a child to benefit from instruction, we would expect these nongesturing children to make little progress after instruction—a result that we think is unlikely. In contrast, if gesture–speech mismatch is not necessary for a child to benefit from instruction, we would expect these children, despite their lack of gesture and gesture–speech mismatch, to make a significant amount of progress after instruction. In fact, we expect that instruction will have some impact on these children but perhaps less impact than on children who demonstrate an effect of cognitive load and produce gesture–speech mismatches as well. Thus, this technique would allow us to determine whether the act of producing mismatch itself facilitates cognitive change.

Even if it turns out that gesture–speech mismatch has little role to play in facilitating cognitive change by affecting the learner directly, it is still possible that mismatch can play an indirect role in cognitive change by exerting an influence on the learning environment. More specifically, the match or mismatch between gesture and speech may serve as a signal to others that the child is in a transitional state and may encourage the child's communication partner to adjust his or her interactions with the child accordingly. In a sense, gesture–speech mismatch, if appropriately interpreted, can signal to the communication partner the area in which the child is currently developing—the zone of proximal development.

Gesture–Speech Mismatch as a Window Into the Zone of Proximal Development

Brown and Ferrara (1985) described two important educational implications of Vygotsky's (1978) notion of the zone of proximal development. The first is diagnostic and focuses on the potential application of the notion to the design of dynamic measures of learning potential (i.e., to consider a child's ability to make use of instruction as one measure of intelligence). The second educational implication focuses on the importance of aiming instruction at the upper bound of a child's zone of proximal development. However, if the zone of proximal development is defined in terms of the child's ability to make use of input from a partner, one is, in effect, recommending that input be directed toward a child who is capable of receiving it—a recommendation that is of limited use unless the partner has some independent way of determining when the child is capable of receiving input and what input would be best for that child to receive. Indeed, this has been the frustration researchers have had with the notion of the zone of proximal development—it is circular unless there is a way of assessing a child's zone of proximal development that is independent of that child's success on the task. We propose that gesture–speech mismatch may, in fact, provide one way of independently assessing the child's zone of proximal development.

We have shown that gesture–speech mismatch identifies children in a transitional state. By this we mean that such children are ready to learn the particular concept on which they display mismatch. Of course, whether children actually learn the concept depends on many factors, not the least of which is the type of input they encounter. If, by using gesture–speech mismatch, children alert their communication partners to the areas in which they are developing, those partners may be able to provide input tailored to the children's needs. The first step in making this argument is to show that adults, not trained in observing gesture, can detect and interpret gesture–speech mismatch.

Goldin-Meadow, Wein, and Chang (1992) investigated whether adults who had not been trained to observe and code gesture could detect and interpret gesture in relation to the speech it accompanies and could use that information in appraising a child's knowledge. In particular, they asked whether untrained observers (elementary school teachers, whose profession necessitates that they become expert in interpreting the cues children transmit about their knowledge, and college students who had no experience teaching young children) could detect the match or mismatch between gesture and speech in a child's explanations of Piagetian conservation tasks and, if so, whether specific information about the child's knowledge of conservation could be gleaned from these gesture–speech matches and mismatches. These investigators found that the adults displayed more uncertainty in their appraisals of children who produced gesture–speech mismatches than in their appraisals of children who produced gesture–speech matches. Moreover, the untrained observers often incorporated the information conveyed in the children's gestures into their own spoken appraisals of the children's reasoning. These data suggest that, even without training, adults form impressions of a child's knowledge—and therefore derive some idea of the kind of input the child needs next—based not only on what children say with their mouths but also on what they say with their hands.

It is worth noting that the adults in the Goldin-Meadow et al. (1992) study were not necessarily aware of the fact that they were noticing—and interpreting—the children's gestures. Few adults mentioned the children's gestures when assessing the children's knowledge. However, it may not be necessary for adults to be explicitly aware of a child's gestures for them to change their input to the child on the basis of those gestures. For example, one of the children whose knowledge of number conservation was assessed by the adult subjects in the Goldin-Meadow et al. study indicated, in speech, that the rows had different numbers of checkers after one had been spread out "because you moved 'em" but, in gesture, indicated that the checkers in one row could be matched in a one-to-one fashion with the checkers in the other row (he pointed to a checker in one row and then to the corresponding checker in the other row and repeated this gesture with another pair of checkers). In her verbal assessment of this child's knowledge, one adult attributed to the child reasoning based on one-to-one correspondence (which appeared only in the child's gesture) as well as reasoning based on the fact that the checkers had been moved (which appeared in the child's speech). If this adult were to interact with this child, the adult might be expected to act as though the child understood one-to-one correspondence, as indeed, at some level, the child had. Being treated as though he understood one-to-one correspondence might be sufficient to force the child to realize that one-to-one correspondence was one of the hypotheses he activated on the conservation task and to encourage the child to integrate this hypothesis with his other, more explicitly stated, views of the concept.

Further evidence that untrained adults can detect and interpret gesture–speech mismatch comes from a study conducted by McNeill, in collaboration with Cassell, McCullough, and Tuite (McNeill, 1992). The experimenters presented a videotaped narration of a cartoon story to adult subjects and asked them to retell the story to a listener. The subjects did not see the cartoon itself, and they retold the story only from this videotaped narrative. Unknown to the subject, the videotaped narration was staged and included a number of mismatching gesture–speech combinations. For example, one mismatch involved a verb, *comes out,* that does not convey any special manner of motion; along with this verb, the accompanying gesture showed a manner, that is, bouncing up and down. If subjects notice and interpret the gesture in the stimulus videotape, they might be expected to alter their own narrations, adjusting them to consolidate the information conveyed in gesture and speech. This is, in fact, what the experimenters found. In response to the above mismatch, subjects tended to incorporate the manner information conveyed only in gesture in the stimulus into speech in their own renditions; for instance, 1 subject said "goes downstairs" to describe this scene, although no stairs had been mentioned in the stimulus at all. Thus, the up-and-down movement originally seen in gesture resurfaced in the retelling as a verbal lexical choice.

In general, McNeill (1992) argued that for speakers, gesture and speech are aspects of a single process. Each modality contributes its own unique level of representation, and the total representation is a synthesis of the imagistic and linear-segmented modes. On the basis of the mismatch experiment, he argued that this same synthesis occurs in listeners. When a listener understands someone, that listener also forms a single unified combination of imagery and speech. The imagery is an integral part of the comprehension. If the speaker provides gestures, they are taken in by the listener—not necessarily consciously—and combined with the verbal stream to recover the speaker's intended meaning. This line of reasoning suggests that listeners cannot avoid noticing and interpreting the gestures that accompany a speaker's words. Indeed, in a study where adults viewed a videotape in which spoken and gestural forms of reference were independently manipulated, Thompson and Massaro (1986) found that the adults used both gestural and spoken information to make their decisions about the referent. Moreover, gesture influenced the adults' judgments to a greater extent when the speech information was ambiguous. In this regard, it is important to note that children's speech has been found to be particularly vague and ambiguous when they are on the verge of making a transition (Graham & Perry, in press; Siegler & Jenkins, 1989). Thus, it is quite likely that adults do notice, and process, the gestures children produce, particularly at the time of transition.

The data from the Goldin-Meadow et al. (1992) study suggest that adults are capable of using information they extract from a child's gestures to influence their assessment of the child's conceptual knowledge. The study did not address the issue of whether adults actually use the information found in gesture to modify the way in which they interact with those children. However, in a study of four mothers' responses to the object-related gestures produced by their first-born infants, Masur (1982) found that the mothers responded to their children's pointing gestures, providing labels for the indicated objects. These observations suggest that adults do make use of children's gestures in determining what their input to a child will be, at least in this relatively straightforward setting.

In our own work, we are attempting to address the question under somewhat more complex communicative conditions. We have begun by asking adults (who were not trained in coding of gesture) to instruct children on an individual basis in conservation. We have then observed the various teaching procedures adopted by these adults and have examined the relationship between these procedures and the gesture-speech discordance status of the individual child. We have found that adults offered different types of instruction to children whose gestures matched their speech than to children whose gestures did not match their speech (Church, Momeni, Williams, Garber, & Goldin-Meadow, 1992). Moreover, the instruction that the adults offered to children with gesture-speech mismatches stressed the principles underlying conservation—precisely the type of instruction that has been found to facilitate learning in discordant children (Perry, Church, & Goldin-Meadow, 1992). These observations suggest that in a relatively naturalistic setting, adults not only interpret the information conveyed in a child's gestures but also use that information in deciding how best to instruct the child (i.e., how to optimize their behavior within that child's zone of proximal development).

In summary, we suggest that gesture-speech mismatch, in addition to reflecting a learner's transitional status, may itself play a role in cognitive change. Gesture-speech mismatch may lead to the resolution of multiple hypotheses by its direct effect on the child learner (a hypothesis we will investigate in our future work, as we have described). Even if it does not have a direct effect, however, gesture-speech mismatch may have an indirect effect by shaping the child's learning environment. The match or mismatch between gesture and speech may serve as a signal alerting children's communication partners (including teachers, parents, more advanced peers, etc.) to calibrate input to children in such a way that they receive the most useful information for their own conceptual reorganization. Thus, children may play a role in shaping their own learning environments by providing signals to their teachers through their production (or lack of production) of gesture-speech mismatches.

Summary: Gesture as a Window Into the Mind of the Child in Transition

McNeill (1985, 1987, 1992) argued that gestures, like speech, can serve as a channel of observation into mental processes and representations, that is, as a window into the mind. Moreover, because gesture is less codified than speech and is dictated by different constraints, it tends to reflect different kinds of knowledge than does speech. Gesture may, for example, reflect knowledge that is imagistic and more implicit than the knowledge conveyed in speech.

We agree with McNeill's formulation and have argued here that when taken in relation to speech, gesture can serve as a window into the mind of the child in a transitional state. We suggest that the relationship between gesture and speech not only serves as an index of the transitional state but, more important, also provides insight into the internal processes that characterize the mind of the child in transition.

We have characterized the transitional knowledge state as one in which multiple hypotheses are simultaneously activated. This simultaneous activation is directly reflected in gesture-speech mismatch, which entails the juxtaposition of two hypotheses, one in speech and one in gesture, within a single explanation. Gesture-speech mismatch itself appears to be an outgrowth of the fact that at a certain point in the learning process, the set of hypotheses accessible to gesture is different from (and in the cases we have studied, larger than) the set accessible to speech; that is, the child possesses knowledge that can be encoded in gesture but not in speech.

Note that the characterization of the transitional state as one in which multiple hypotheses are simultaneously activated constrains the types of learning mechanisms that can be posited. Any mechanism of change purported to account for this type of transition must involve two different processes. One process serves to introduce a new hypothesis into the learner's repertoire (in many instances, into the learner's gestural repertoire rather than into his or her spoken repertoire); this process consequently creates a transitional state characterized by multiple hypotheses. A second process serves to sort out the multiple hypotheses in the learner's repertoire, perhaps by allowing the learner to recode the imagistic knowledge encoded in gesture into the linear and segmented code characteristic of speech; this process thus results in a single, correct hypothesis that is encoded in both modalities, or perhaps even in a set of interrelated hypotheses, all of which are encoded in both modalities.

We argue finally that the relationship between gesture and speech serves an important role in the mechanism of developmental change itself, signaling to those who interact with the child that the child is in a transitional state and providing insight into the hypotheses the child is entertaining. We suggest that gesture, taken in relation to speech, reflects those concepts currently in the child's zone of proximal development—the areas in which the child is able to profit from instruction and learn. Gesture and speech taken together provide an observable, and interpretable, index of the child's conceptual knowledge, and thus they provide a mechanism by which adults can calibrate their input to a child's current level of understanding.

References

Acredolo, C., & O'Connor, J. (1991). On the difficulty of detecting cognitive uncertainty. *Human Development, 34*(4), 204–223.

Acredolo, C., O'Connor, J., & Horobin, K. (1989, April). *Children's understanding of conservation: From possibility to probability to necessity.* Paper presented at the biennial meeting of the Society for Research in Child Development, Kansas City, MO.

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

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59,* 87–96.

Beilin, H. (1965). Learning and operational convergence in logical thought development. *Journal of Experimental Child Psychology, 2,* 317–339.

Brainerd, C. J. (1972). The age–stage issue in conservation acquisition. *Psychonomic Science, 29,* 115–117.

Brainerd, C. J. (1977). Cognitive development and concept learning: An interpretative review. *Psychological Bulletin, 84,* 919–939.

Brainerd, C. J. (1979). Markovian interpretations of conservation learning. *Psychological Review, 86,* 181–213.

Brainerd, C. J., & Brainerd, S. H. (1972). Order of acquisition of number and quantity conservation. *Child Development, 43,* 1401–1406.

Broadbent, D. (1991). Recall, recognition, and implicit knowledge. In W. Kessen, A. Ortony, & F. Craik (Eds.), *Memories, thoughts, and emotions: Essays in honor of George Mandler* (pp. 125–134). Hillsdale, NJ: Erlbaum.

Brown, A. L., & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 273–305). Cambridge, England: Cambridge University Press.

Bryan, W. L., & Harter, N. (1899). Studies of the telegraphic language. The acquisition of a hierarchy of habits. *Psychological Review, 6,* 345–376.

Church, R. B. (1987). *The mismatch between gesture and speech as an index of transitional knowledge.* Unpublished doctoral dissertation, University of Chicago.

Church, R. B. (1990, May). *Equilibration: Using gesture and speech to monitor cognitive change.* Paper presented at the Twentieth Anniversary Symposium of the Jean Piaget Society, Philadelphia, PA.

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

Church, R. B., Momeni, C., Williams, T., Garber, P., & Goldin-Meadow, S. (1992, May). *Detection of unstable knowledge: The function of gesture and speech in natural teaching interactions.* Paper presented at the annual meeting of the Jean Piaget Society, Montreal, Canada.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development, 48,* (Serial No. 200), 1–2.

Evans, M. A., & Rubin, K. H. (1979). Hand gestures as a communicative mode in school-aged children. *Journal of Genetic Psychology, 135,* 189–196.

Ferrara, R. A., Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child Development, 57,* 1087–1099.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87,* 477–531.

Fischer, K. W., & Pipp, S. L. (1984). Processes of cognitive development: Optimal level and skill acquisition. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 45–80). New York: Freeman.

Flavell, J. H. (1984). Discussion. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 187–210). New York: Freeman.

Forman, E. A., & Cazden, C. B. (1985). Exploring Vygotskian perspectives in education: The cognitive value of peer interaction. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 323–347). Cambridge, England: Cambridge University Press.

Garber, P., Alibali, M. W., & Goldin-Meadow, S. (1992). *Knowledge conveyed in gesture as implicit knowledge.* Unpublished manuscript.

Gershkoff-Stowe, L., & Smith, L. B. (1991, April). *Changes in pointing and labeling during the naming explosion.* Poster presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.

Goldin-Meadow, S. (1993). When does gesture become language? A study of gesture used as a primary communication system by deaf children of hearing parents. In K. R. Gibson & T. Ingold (Eds.), *Tools, language and cognition in human evolution* (pp. 63–85). Cambridge, England: Cambridge University Press.

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Non-effect of parental input on language development. *Science, 221,* 372–374.

Goldin-Meadow, S., & Mylander, C. (1984). Gestural communication in deaf children: The effects and non-effects of parental input on early language development. *Monographs of the Society for Research in Child Development, 49* (3, Serial No. 207).

Goldin-Meadow, S., & Mylander, C. (1990). Beyond the input given: The child's role in the acquisition of language. *Language, 66,* 323–355.

Goldin-Meadow, S., Nusbaum, H., Garber, P., & Church, R. B. (1993). Transitions in learning: Evidence for simultaneously activated hypotheses. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1–16.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction, 9*(3), 201–219.

Goodman, N., Church, R. B., & Schonert, K. (1991, May). *Moral development and gesture: What can the hands reveal about moral reasoning?* Paper presented at the annual meeting of the Jean Piaget Society, Philadelphia, PA.

Graham, T., & Perry, M. (in press). Indexing transitional knowledge. *Developmental Psychology.*

Griffin, P., & Cole, M. (1985). Current activity for the future: The zoped. In B. Rogoff & J. V. Wertsch (Eds.), *Children's learning in the "zone of proximal development": New directions for child development* (Vol. 23, pp. 45–64). San Francisco: Jossey-Bass Inc.

Hadamard, J. (1945). *The psychology of invention in the mathematical field.* New York: Dover.

Huttenlocher, J. (1973). Language and thought. In G. A. Miller (Ed.), *Communication, language, and meaning: Psychological perspectives* (pp. 172–184). New York: Basic Books.

Huttenlocher, J. (1976). Language and intelligence. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 261–281). Hillsdale, NJ: Erlbaum.

Inhelder, B., & Sinclair, H. (1969). Learning cognitive structures. In P. Mussen, J. Langer, & M. Covington (Eds.), *Trends and issues in developmental psychology* (pp. 2–21). New York: Holt, Rinehart & Winston.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes, 1,* 61–85.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23,* 95–147.

Keil, F. C. (1984). Mechanisms of cognitive development and the structure of knowledge. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 81–100). New York: Freeman.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *Relationship of the verbal and nonverbal communication* (pp. 207–228). The Hague, The Netherlands: Mouton.

Klahr, D. (1984). Transition processes in quantitative development. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 101–140). New York: Freeman.

Klima, E., & Bellugi, U. (1979). *The signs of language.* Cambridge, MA: Harvard University Press.

Langer, J. (1969). Disequilibrium as a source of development. In P. Mussen, J. Langer, & M. Covington (Eds.), *Trends and issues in developmental psychology* (pp. 22–37). New York: Holt, Rinehart & Winston.

Langer, J., & Strauss, S. (1972). Appearance, reality and identity. *Cognition, 1,* 105–128.

Levine, S. C., Jordan, N. C., & Huttenlocher, J. (1992). Development of calculation abilities in young children. *Journal of Experimental Child Psychology, 53,* 72–103.

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology, 39,* 367–386.

Masur, E. F. (1982). Mothers' responses to infants' object-related gestures: Influences on lexical development. *Journal of Child Language, 9,* 23–30.

Mazur, J. E., & Hastie, R. (1978). Learning as accumulation: A reexamination of the learning curve. *Psychological Bulletin, 85,* 1256–1274.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92,* 350–371.

McNeill, D. (1987). *Psycholinguistics: A new approach.* New York: Harper & Row.

McNeill, D. (1992). *Hand and mind.* Chicago: University of Chicago Press.

Morford, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language, 19,* 559–580.

Murray, J. P. (1974). Social learning and cognitive development: Modelling effects on children's understanding of conservation. *British Journal of Psychology, 65,* 151–160.

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

Perry, M., Church, R. B., & Goldin-Meadow, S. (1992). Is gesture/speech mismatch a general index of transitional knowledge? *Cognitive Development, 7,* 109–122.

Piaget, J. (1985). *The equilibration of cognitive structures.* Chicago: University of Chicago Press. (Original work published 1975)

Rogoff, B., & Wertsch, J. V. (Eds.). (1984). *Children's learning in the "zone of proximal development": New directions for child development* (Vol. 23). San Francisco: Jossey-Bass.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Siegel, L. S. (1978). The relationship of language and thought in the preoperational child: A reconsideration of nonverbal alternatives to Piagetian tasks. In L. S. Siegel & C. J. Brainerd (Eds.), *Alternatives to Piaget: Critical essays on the theory* (pp. 43–68). San Diego, CA: Academic Press.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

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

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development, 46* (2, Serial No. 189).

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies.* Hillsdale, NJ: Erlbaum.

Singleton, J., Goldin-Meadow, S., & McNeill, D. (in press). The cataclysmic break between gesticulation and language. In K. Emmorey & J. Reilly (Eds.), *Sign, gesture, and space.* Hillsdale, NJ: Erlbaum.

Snyder, S. S., & Feldman, D. H. (1977). Internal and external influences on cognitive developmental change. *Child Development, 48,* 937–943.

Stone, A., Webb, R., & Mahootian, S. (1991). The generality of gesture-speech mismatch as an index of transitional knowledge: Evidence from a control-of-variables task. *Cognitive Development, 6,* 301–313.

Strauss, S. (1972). Inducing cognitive development and learning: A review of short-term training experiments: I. The organismic developmental approach. *Cognition, 1,* 329–357.

Strauss, S., & Langer, J. (1970). Operational thought inducement. *Child Development, 40,* 163–175.

Strauss, S., & Rimalt, I. (1974). Effects of organizational disequilibrium training on structural elaboration. *Developmental Psychology, 10,* 526–533.

Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology, 42,* 144–168.

Turiel, E. (1969). Developmental processes in the child's moral thinking. In P. Mussen, J. Langer, & M. Covington (Eds.), *Trends and issues in developmental psychology* (pp. 92–133). New York: Holt, Rinehart & Winston.

Turiel, E. (1974). Conflict and transition in adolescent moral development. *Child Development, 45,* 14–29.

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

Wilkinson, A. C. (1982a). Partial knowledge and self-correction: Developmental studies of a quantitative concept. *Developmental Psychology, 18,* 876–893.

Wilkinson, A. C. (1982b). Theoretical and methodological analysis of partial knowledge. *Developmental Review, 2,* 274–304.

Wilkinson, A. C., & Haines, B. A. (1987). Learning a cognitive skill and its components. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology: Progress in cognitive development research* (pp. 86–114). New York: Springer-Verlag.

Received October 1, 1991
Revision received April 30, 1992
Accepted July 30, 1992 ■