COGNITIVE PSYCHOLOGY 25, 468–523 (1993)

Gesture–Speech Mismatch and Mechanisms of Learning: What the Hands Reveal about a Child's State of Mind

MARTHA WAGNER ALIBALI AND SUSAN GOLDIN-MEADOW

The University of Chicago

Previous work has shown that, when asked to explain a concept they are acquiring, children often convey one procedure in speech and a different procedure in gesture. Such children, whom we label "discordant," have been shown to be in a transitional state in the sense that they are particularly receptive to instruction—indeed more receptive to instruction than "concordant" children, who convey the same procedure in speech and gesture. This study asks whether the discordant state is 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. To address this question, children were asked to solve and explain a series of problems instantiating the concept of mathematical equivalence. The relationship between gesture and speech in each explanation was monitored over the series. We found that the majority of children who learned to correctly solve equivalence problems did so by adhering to the hypothesized path: They first produced a single, incorrect procedure. They then entered a discordant state in which they produced different procedures—one in speech and another in gesture. Finally, they again produced a single procedure, but this time a correct one. These data support the notion that the transitional state is characterized by the concurrent activation of more than one procedure, and provide further evidence that gesture can be a powerful source of insight into the processes involved in cognitive development.

This research was supported by Grant RO1 HD18617 from the National Institute of Child Health and Human Development to Goldin-Meadow. We thank Martha Scott for her help at every stage of this project. We also thank Sharon Syc for help in collecting data; Cecilia Chang, Jennifer Walton, and Philip Garber for help in coding data; Steve Shevell for statistical advice; Miriam Bassok, Susan Levine, David McNeill, Bill Meadow, Jim Stigler, and Tom Trabasso for insightful comments on earlier drafts of the manuscript; and Rochel Gelman, Annette Karmiloff-Smith, and two anonymous reviewers for their helpful comments and suggestions. We are indepted to many others for their help in developing the ideas presented in this paper, including Cynthia Butcher, Breckie Church, Theresa Graham, Jill Morford, Carolyn Mylander, Michelle Perry, Jenny Singleton, and Petrit Alibali. Finally, we are most grateful to the principals, teachers, and students who made this research possible. Address correspondence and reprint requests to M. W. Alibali, University of Chicago, 5848 S. University Avenue, Chicago, IL 60637, or to S. Goldin-Meadow, University of Chicago, 5730 S. Woodlawn Avenue, Chicago, IL 60637.

0010-0285/93 $9.00

One of the central goals of cognitive and developmental psychology is to explain the mechanisms and processes which underlie learning. Learning, in both adults and children, involves moving from a less adequate to a more adequate understanding of a task. Although it is important to be able to describe the learner's state before and after the task has been mastered, characterizing the process that bridges these two states is the key to understanding learning. Unfortunately, in many studies of learning in both adults and children, performance, procedures, and mental representations are described before and after learning, while little attention is paid to the transition between these states (see Glaser & Bassok, 1989, for discussion).

For example, while much research has sought to describe and characterize the mental representations of novice and expert problem solvers (see, e.g., Chase & Simon, 1973; Chase, & Ericsson, 1981; Glaser & Chi, 1988), this research rarely examines the nature of the cognitive path between these states. Similarly, although research in developmental psychology often acknowledges the importance of understanding transitions in learning (e.g., Flavell, 1984), in practice, much of the research also falls short of exploring the period of transition itself. Developmental studies tend to document the fact that children progress from one state to another but pay little attention to the changes in cognitive processing that take place during the transition between states.

One of Piaget's major contributions to the study of children's acquisition of concepts has been the demonstration that a child's understanding of many tasks is, throughout the period of acquisition, systematic and rule-governed. In studies of over a dozen Piagetian tasks, Siegler (1981, 1983) has shown that, although the particular rules that children use to solve a task change substantially with age, the percentage of children classified as using a single rule on that task is high and changes little from ages 5 to 17. Findings of this sort suggest that acquisition in certain cognitive domains is best characterized as a progression from one (presumably inadequate) rule to another (more adequate) rule. Nevertheless, little work has been designed to explore how children make the *transition* from one rule-governed state to the next (see, however, Siegler & Jenkins, 1989). Thus, with respect both to children and adults, the study of learning could benefit from a greater focus on the mechanisms by which new rules supplant old ones.

The absence of focus on transition per se in studies of learning may derive from difficulties inherent in studying the transient, and therefore relatively short-lived, state. It is considerably less onerous to undertake a description of the steady states that characterize the endpoints of learning than to describe the fleeting state that occurs between them. In fact, what is needed to study the transient state is a technique for isolating learners when they are in such a state and for characterizing the nature of their instability. In our previous work, we have shown that the mismatch between the thoughts a learner expresses in speech and those expressed in gesture serves as a signal that the learner is in a state of transition and ready to benefit from instruction in a particular task (Goldin-Meadow, Alibali, & Church, 1993). For example, a child who said, "the glass is tall" while producing a gesture indicating the *width* of a glass on a conservation task, was more likely to benefit from instruction in the task than a child who said, "the glass is tall" while producing a gesture indicating the *height* of the glass (Church & Goldin-Meadow, 1986). Thus, gesture–speech mismatch is a technique which can be used to distinguish those who are on the threshold of learning a concept from those who are not. Moreover, and most importantly, the technique also offers access to the cognitive processes that characterize transition.

The purpose of the present study is to use gesture–speech mismatch to probe the cognitive processes that characterize the transition from an incorrect, yet rule-governed, understanding of a problem to a correct, rule-governed understanding of the problem. We have chosen to explore this phenomenon in children primarily because, in contrast to adults, it is relatively easy to find a domain where a child is not just inefficient at a task but has an incorrect understanding of the task. We begin by reviewing the relevant work on transitions in children; we will consider the implications of our findings for general theories of cognitive change under Discussion.

Transitions in Learning: Characterizing Children on the Verge of Learning

By definition, the child in transition with respect to a particular task has not yet mastered that task. The upper bound of the transitional state is therefore, at least in principle, easy to establish. However, at what point in a child's mastery of a task is it sensible to say that the child has *entered* a transitional state with respect to that task? For example, children's understanding of balance scale problems develops slowly and, although children can make reliable predictions about problems at age 5, a complete understanding is often not achieved until 17 years (Siegler, 1976). Do we want to say that the child is "in transition" from age 5 to 17? In order 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 at the point of mastery. Instability implies that change is imminent, particularly if the appropriate input is provided. Thus, one might argue that a child is in a transitional state with respect to a particular problem when that child is ready to learn or to benefit from input in that task; that is, given appropriate input, the child with transitional knowledge masters the task quickly (cf. Beilin, 1965; Brainerd, 1972; Langer & Strauss, 1972; Murray, 1974; Strauss & Langer, 1970; Strauss & Rimalt, 1974).

If we assume that children in a transitional state with respect to a task are ready to learn that task, we are then faced with the problem of determining what characterizes learners on the verge of change. One characterization that has been proposed in the developmental literature—one that has intuitive appeal—is that children who are on the verge of learning a task have at their disposal a variety of hypotheses for dealing with that task. For example, Strauss (1972) and Strauss and Rimalt (1974) have argued that children who display at least two functional structures when solving problems instantiating a concept are in a transitional state with respect to that concept. Evidence of more than one functional structure on a series of tasks which instantiate a concept has been called "structural mixture." Strauss (1972) has distinguished between two types of structural mixture—mixture within vs. mixture between concepts. A child who has structural mixture *within* a concept will be correct on some but not all of a series of tasks which instantiate a single concept, e.g., length conservation. A child who had structural mixture *between* concepts will be correct on some but not all of a series of tasks which instantiate a set of closely related concepts. For example, a child may succeed on discontinuous quantity (number) conservation, but not succeed on length conservation.

Although the notion of structural mixture has intuitive appeal, the evidence which suggests that children who display structural mixture on a task are more likely to improve their understanding of that task than children who do not display structural mixture is not compelling. Strauss and Langer (1970) found that, after training, children who displayed structural mixture *within* a concept (conservation of continuous quantity) advanced to a more sophisticated understanding of conservation than children who did not demonstrate structural mixture within the concept. Similarly, Langer and Strauss (1972) found that, after training, children who had structural mixture *between* concepts (conservation of discontinuous quantity and of length) advanced to a more sophisticated understanding of conservation than children who did not demonstrate structural mixture between the concepts (see also Inhelder & Sinclair, 1969). However, these studies did *not* convincingly demonstrate that children with structural mixture are more ready to learn than children without structural mixture. In a reanalysis of the data from these studies taking the children's initial level of understanding into account, 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. Instead, Brainerd argued, children with structural mixture 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.

In general, the difficulty in interpreting these studies highlights two important methodological caveats that Brainerd (1977) has proposed for studies purporting to identify children on the verge of learning. First, Brainerd argued that the measure used to capture learning in studies of readiness to learn 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. Second, Brainerd argued that, in order 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; 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.

As a group, children who have two hypotheses for dealing with a problem appear no more likely to improve their understanding of that problem than children who have only one. In fact, one might expect that having two hypotheses for solving a problem would not be sufficient to motivate change—unless those hypotheses had, at least at some point, been considered on the same problem. For example, if a child were to activate one hypothesis on certain types of problems and the second hypothesis on a different set of problems, that child would not necessarily ever consider those two hypotheses concurrently. It may be that what characterizes children on the verge of learning is not the mere availability of more than one hypothesis, but rather the *concurrent* activation and evaluation of those hypotheses. We consider this possibility in the next section.

The Concurrent Activation of Multiple Hypotheses as a Characterization of Children on the Verge of Learning

There is theoretical reason to believe that the *concurrent* consideration of more than one hypothesis leads to uncertainty, which then provides the impetus for transition in learning. 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 hypotheses a child uses to solve a problem. In order for these discrepancies to have an impact on the child's development, that child must have at some point considered and compared the hypotheses he or she has available. Similarly, Acredolo, O'Connor, and Horobin (1989) suggest that it is uncertainty which serves as the primary force underlying cognitive growth, and that this uncertainty stems from the confusion children experience when they consider more than one hypothesis on a single problem.

Even traditions that are distinctly non-Piagetian have proposed that multiple solutions to a single problem provide the motivating force for transition. For example, in his discussion of structure-dependent transition mechanisms, Keil (1984) includes resolution of internal inconsistencies or contradictions as a mechanism of transition; in order to be internally inconsistent, a child must entertain two incompatible views of the same problem. In his theory of cognitive development, called skill theory, Fischer (1980) describes five procedures that specify how a skill is transformed into a new, more advanced skill; each of these procedures involves transforming two or more skills with given structures into one or more skills with a new type of structure and thus calls for activation of at least two skills in order for developmental change to occur. From an information processing perspective, Klahr (1984) lists conflict-resolution procedures, which apply when two productions are eligible to be activated on a single problem, as an important mechanism of change in self-modifying systems. The common thread running through all these theoretical characterizations of change is the notion that multiple, potentially incompatible hypotheses are activated or considered in solving a single problem.

Is there *empirical evidence* for the claim that it is the concurrent consideration of multiple hypotheses that characterizes children on the verge of learning? In order to provide evidence for this claim, we must show, at a minimum, that the child considers more than one hypothesis on the *same* problem. As Acredolo and colleagues (Acredolo & O'Connor, 1991; Acredolo, O'Connor & Horobin, 1989) have pointed out, evidence of this sort is difficult to obtain simply because the procedures typically used to tap children's knowledge of a concept encourage the child's natural inclination to close on one solution (see also Miller, Brownell, & Zukier, 1977).[1] In a study designed to overcome this difficulty, Acredolo et al. (1989) provided children with the opportunity to assign probabilities to a

---

[1] Note that, even though an experimenter designs a study to elicit a single solution, children may provide evidence of additional knowledge through other aspects of their behavior. However, in most cases, experimenters do not attend to these other behaviors and instead focus exclusively on the response the study was designed to elicit. One exception is Siegler (1976) who found that, while both 5- and 8-year-olds used a weight-only rule in solving a series of balance scale problems (the behavior the experiment was designed to elicit), the 8-year-olds produced nonverbal cues (in particular, head movements) which suggested that they were aware of the weights' distance from the fulcrum, while the 5-year-olds gave no such evidence. Thus, although the 8-year-olds used a single rule involving weight to produce answers to the balance scale problems, their nonverbal behavior suggested that they were also entertaining a second hypothesis involving distance.

variety of alternative solutions to a problem. Using this paradigm, Acredolo et al. found that children (particularly children who had not yet acquired conservation according to traditional measures) frequently did consider more than one solution to be possible on a single conservation problem when given a variety of solutions or hypotheses from which to choose. However, it is important to note that data of this sort do not (and, indeed, cannot) demonstrate that a child *spontaneously* entertains more than one hypothesis in solving a problem.

Children rarely cite more than one hypothesis when asked to explain how they solved a particular problem; thus, their verbal explanations will not necessarily reveal whether they have considered multiple hypotheses on a single problem. However, previous work has shown that, when asked to explain their performance on a task, children frequently gesture along with their spoken explanations, and these gestures often convey substantive information about the task itself (e.g., Evans & Rubin, 1979). Our previous work has shown that, while gesture may convey the same information as conveyed in speech (and thus match speech), this is not always the case. At times, a child's gestures may convey a hypothesis different from the one expressed in the accompanying speech (and thus mismatch speech), suggesting that the child has concurrently considered more than one hypothesis. Note that we are not making a specific claim as to whether the hypotheses displayed in a gesture–speech mismatch are processed in parallel or in sequence. Rather, we suggest simply that gesture–speech mismatch provides evidence that multiple hypotheses are considered on a single problem, and that these hypotheses are expressed within the same response. We have found this phenomenon of gesture-speech mismatch in studies of the acquisition of two different concepts at two different ages: conservation in 5- to 8-year-olds (Church & Goldin-Meadow, 1986) and mathematical equivalence in 9- to 10-year-olds (Perry, Church & Goldin-Meadow, 1988, 1992). In the next section, we review our findings on gesture-speech mismatch in the acquisition of mathematical equivalence and consider the implications of those findings for understanding learning.

Multiple Hypotheses in a Single Explanation: Gesture–Speech Mismatch

Perry et al. (1988) tested children between the ages of 9 and 10 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. When asked to explain their solutions, the children usually gestured spontaneously while speaking and often used those gestures to convey specific procedures that described how to solve the problem. At times, the procedure conveyed in gesture 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).

However, as mentioned above, at other times, the gestures produced by the children did *not* convey the same procedure as the speech which 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).

Perry et al. (1988) found that the children in their study varied in the number of gesture-speech mismatches they produced, some producing none and some producing as many as 6 (out of a possible 6). Thus, some children routinely produced one procedure in their spoken responses and a different procedure in the accompanying gesture, suggesting not only that they had two procedures in their repertoire but also that they considered those procedures concurrently, while *explaining* a single problem.

Crucial to our exploration of the role that multiple hypotheses play in transition is the fact that Perry et al. (1988) found that the children who produced many gesture–speech mismatches in their explanations (labeled "discordant" children by Perry et al.) were more likely to benefit from instruction in equivalence than the children who produced few gesture-speech mismatches (labeled "concordant" children). The relative ease with which the discordant children learned to correctly solve equivalence problems was particularly striking given that none of the children (either concordant or discordant) produced correct solutions on any of the problems before training. Moreover, the explanations produced by the children before training always contained procedures which, if followed, led to incorrect solutions. Thus, the discordant children, who by definition gave explicit evidence of considering two procedures on a single problem, were in transition with respect to acquiring mathematical equivalence (i.e., were ready to learn to correctly solve the problems), while the concordant children, who by definition gave evidence of considering only a single procedure on a problem, were not (see Church & Goldin-Meadow, 1986, for comparable results with respect to conservation).

It is important to stress that the Perry et al. (1988) study does satisfy the methodological standards established by Brainerd (1977) for studies of learning readiness. First, the measure used to identify children in a transitional state was not the same as the measure used to assess learning.

Children were identified as transitional on the basis of the number of gesture–speech mismatches they produced on the pretest. Learning was measured in terms of the number of math problems solved correctly on the posttest and the generalization test. Second, the measure used to assess learning did take into account the child's performance on the pretest. None of the children solved any of the math problems correctly on the pretest; thus, their posttest scores were a measure of improvement after training.

These findings suggest that children on the verge of learning a task consider more than one hypothesis when *explaining* their solutions to a single task. However, the fact that children may exhibit two hypotheses when explaining how they solved a task does not necessarily mean that the children consider both hypotheses when actually solving the task. We address this issue in the next section.

Do Children on the Verge of Learning Activate More than One Hypothesis When Solving Problems?

It is conceivable that gesture-speech mismatch could reflect post hoc reasoning processes rather than on-line problem solving. To explore this possibility, Goldin-Meadow, Nusbaum, Garber and Church (1993) conducted a study to determine whether discordant children not only activate more than one hypothesis when they *explain* their solutions to a problem, but also activate those hypotheses 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 have less capacity left over to simultaneously perform a second task, compared to a child who activates only a single hypothesis.

To test this prediction, Goldin-Meadow et al. first identified children as concordant or discordant with respect to mathematical equivalence based on their explanations on a pretest. They then 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. predicted that the discordant and concordant children would both perform poorly on the math test, but that the discordant children—if, in fact, they were activating two hypotheses or procedures 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 perform less well 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 recalled the word lists significantly less well than the concordant children. The discordant children thus appeared to be working harder to solve the math problems incorrectly than were the concordant children.

We take the Goldin-Meadow et al. (1993) findings to support our view that a child on the verge of learning a task not only *expresses* multiple hypotheses when explaining solutions to the task, but also concurrently *activates* multiple hypotheses (demanding extra cognitive capacity) when solving the task itself. We argue that the concurrent activation of multiple hypotheses is characteristic of the transitional state and is reflected both in the way children communicate about problems and in the amount of effort they expend in solving those problems.

In a sense, the phenomenon of gesture-speech mismatch is similar to previously described phenomena wherein children give one response when their knowledge is tapped through, for example, judgments in a conservation task, and a different response when their knowledge is tapped through explanations (e.g., Siegel, 1978; Brainerd & Brainerd, 1972). In the same way, gesture–speech mismatch can be viewed as similar to instances in which children oscillate between two or more *verbal* explanations for a single problem, as described by Piaget (1967, p. 156) and others (e.g., Church, 1993). In this regard, however, it is worth noting that Church (1993) found that the use of multiple verbal explanations on conservation tasks was both rare and a poorer predictor of readiness to learn than gesture-speech mismatch. Although both judgment–explanation discrepancies and multiple verbal explanations indicate that a child may entertain multiple hypotheses about a problem, we believe that gesture–speech mismatch differs from these phenomena in an important way. A child who produces a gesture–speech mismatch produces two different hypotheses *at the same moment.* It is this concurrence which we take to be the hallmark of the transitional state.

Rationale for the Study: A Microgenetic Study of Learning

To summarize thus far, we have shown in previous work that the mismatch between gesture and speech in a child's explanations of a problem is a good indicator that the child is in a transitional knowledge state with respect to the problem (i.e., is ready to learn the problem). We have further shown that the concurrent expression of two hypotheses in an explanation of a problem reflects concurrent activation of multiple hypotheses when solving the problem. These observations 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, gesture–speech mismatch appears to reflect processes which are central to the transitional state. In gesture–speech mismatch, two hypotheses are expressed in a single explanation of a problem—one in gesture and another in speech. We suggest that it is the concurrent activation of multiple hypotheses that characterizes the transitional knowledge state and creates gesture–speech mismatch.

One shortcoming of our previous work in assessing transition has been its use of the traditional design for studies of learning: The child's initial abilities are assessed, the child is then trained, and the child's abilities are reassessed. If the child's performance changes after training (that is, if there is a difference between the initial assessment and the reassessment), we infer that the child has undergone a transition. Although adequate for showing *that* learning has taken place, this design does not shed much light on the processes which characterize that learning. In order to explore the process of transition itself, we require repeated and detailed observations of individual children as they learn a task, with particular focus on the period of change—what has been termed the "microgenetic" method (cf., Inhelder, Ackerman-Vallado, Blanchet, Karmiloff-Smith, Kilcher-Hagedorn, Montagero, & Robert, 1976; Siegler & Crowley, 1991).

The study presented here was designed to follow the course of learning in just this way. Specifically, this study was designed to monitor the changes in the number and kinds of procedures children consider on a particular task over the course of their acquisition of that task. To accomplish this goal, we monitored the relationship between gesture and speech in the explanations children gave to math problems as they acquired an understanding of mathematical equivalence. This particular task was well suited to such a microgenetic study, since previous work has shown that children can learn to solve mathematical equivalence problems correctly in a short period of time, given appropriate instructional input (Perry et al., 1988), and because there are a variety of different procedures (both correct and incorrect) that children use when solving problems instantiating the concept (see below).

If we are correct in suggesting that transitional knowledge is characterized by the concurrent activation of multiple hypotheses—and if gesture–speech mismatch provides a good index of concurrently activated multiple hypotheses—then the process of learning in general, and learning mathematical equivalence in particular, might be expected to proceed as follows: The child ought to begin with a single incorrect hypothesis about a task, characterized by gesture–speech *match* in explanations of the task which describe an incorrect procedure. The child should then move through a transitional period in which he or she entertains multiple hypotheses, characterized by *mismatch* between the hypothesis expressed in speech and the hypothesis expressed in gesture. Finally, the child should close on a single correct hypothesis, characterized again by gesture–speech *match* in explanations which describe a correct procedure. The purpose of this study was to validate this hypothesized path of acquisition using a microgenetic design.

METHOD

Training Conditions and Overview of Procedure

Our goal was to observe the steps children take in learning to solve mathematical equivalence problems. To achieve this goal, we asked each child to solve and explain a series of such problems. We then examined the explanation each child gave to each of the problems in the series, and charted the changes in the relationship between gesture and speech that occurred over the series of explanations.

In order to explore the process of learning to solve mathematical equivalence problems, we needed first to ensure that children would learn to solve the problems correctly. We therefore gave two groups of children instruction in the concept of mathematical equivalence—instruction previously shown to be effective in inducing an understanding of mathematical equivalence (Perry et al., 1988). In addition, we included a third group of children who participated in a training session but received no explicit instruction during that session. We expected that few of these children would learn to solve the problems correctly, and that this group would therefore serve as a control for the two instructional groups.

Consequently, the children were divided into three groups, each of which received a different type of training experience: (1) *No Instruction.* The children in this group were asked to solve and explain a series of 12 addition problems during the training session but received no instruction or feedback on any of the problems. (2) *Addition.* The children in this group were asked to solve and explain the same 12 addition problems during the training session, and received feedback and instruction after each of the 12 problems. (3) *Addition-plus-Multiplication.* The children in this group were asked to solve and explain the 12 addition problems and, to give the children experience in generalizing their knowledge to a new operation, 6 multiplication problems as well (i.e., 18 problems in all). Like the children in the Addition group, these children received feedback and instruction after each of the problems.

The overall procedure for the study was as follows (see Fig. 1): Each child was initially given a pretest assessing his or her understanding of mathematical equivalence. The child then participated in one of the three types of training sessions, as described above. Immediately following the training session, the child was given a posttest comparable to the pretest. Finally, the child was retested after a period of approximately 2 weeks on a follow-up test also comparable to the pretest.

Subjects

The subjects were drawn from the fourth grade classes of six parochial schools in Chicago. We chose to train fourth grade children in mathematical equivalence because previous work has shown that most fourth grade students do not fully understand the concept, as indicated by their incorrect solutions to problems of the form $3 + 4 + 5 = \_\_ + 5$ (Perry, 1985).

PRETEST

I. Paper and Pencil Test
six problems of the form 3+4+5=__+5
four problems of the form 4+7+8=__+9
four problems of the form 3x2x4=__x4
II. Problem Explanations
six problems of the form 3+4+5=__+5
(problems from paper and pencil test)

INSTRUCTION

No Instruction	Addition	Addition-plus-Multiplication
twelve problems	twelve problems	eighteen problems
six 5+9+3=__+3	six 5+9+3=__+3	six 5+9+3=__+3
six 4+6+8=__+7	six 4+6+8=__+7	six 4+6+8=__+7
		six 4x2x3=__x3

POSTTEST

Paper and Pencil Test
six problems of the form 8+4+3=__+3
four problems of the form 3+6+8=__+7
four problems of the form 4x5x3=__x3

FOLLOW-UP TEST

Paper and Pencil Test
six problems of the form 7+6+5=__+5
four problems of the form 5+9+7=__+8
four problems of the form 2x3x5=__x5

FIG. 1. Design of the study. Children in the No Instruction condition received no explicit instruction with any of the problems, while children in both the Addition and the Addition-plus-Multiplication groups received instruction with each problem.

The goal of this study was to use the relationship between gesture and speech in a child's explanations of mathematical equivalence to explore that child's transition from an incorrect understanding of equivalence problems to a correct understanding. In order to determine whether the relationship between gesture and speech provides a window through which the process of change can be observed and investigated, it was necessary to include in the study a sufficient number of children who gestured as they spoke. We therefore continued to test children until we had obtained a minimum of 20 children who gestured in each of the three training conditions. In the process of collecting data on these Gesturers, we also collected data on 27 Non-Gesturers. Children were considered Non-Gesturers if they did not gesture at all or if, on each set of 6 problems, they gestured on fewer than 2 of those problems (in contrast, children were considered Gesturers if they gestured at least three times in one or more of the sets of 6 problems which comprised the study). We chose this criterion simply because children must produce gesture–speech mismatches on at least 3 of their 6 explanations in order to be considered "discordant" (see below). Of the 27 subjects who were classified as Non-Gesturers, 16 gestured very rarely and 11 did not gesture at all. The Non-Gesturers provided no opportunity for us to observe patterns of change in the relationship between gesture and speech. Thus, their data have no bearing on, and will not be included in, analyses of the changing relationship between gesture and speech during transition. However, data from the Non-Gesturers will be analyzed later when we explore whether gesturing itself plays a role in facilitating change.

One hundred and seven children were screened for participation in the study; 5 were eliminated because they were already able to solve equivalence problems correctly (all solved 9 of the 10 pretest addition problems correctly), and another 12 were eliminated due to technical difficulties (6 due to equipment malfunction and 6 due to an inability to either see or hear the child's responses on the videotape). The remaining 90 children, including 63 Gesturers (32 females and 31 males) and 27 Non-Gesturers (14 females and 13 males), comprised the sample for the study. These children were randomly assigned to one of the three training conditions. The mean ages for the Gesturers in each of the three groups were 10;0 for the 20 children in the No Instruction group (range 9;3 to 11;1), 10;1 for the 22 children in the Addition group (range 9;3 to 11;3), and 10;0 for the 21 children in the Addition-plus-Multiplication group (range 9;5 to 11;4). The mean ages for the Non-Gesturers in each of the three groups were 10;1 for the 12 children in the No Instruction group (range 9;5 to 10;10), 9;10 for the 10 children in the Addition group (range 9;5 to 10;11), and 9;10 for the 5 children in the Addition-plus-Multiplication group (range 9;4 to 10;5). Note that there were fewer Non-Gesturers in the Addition and Addition-plus-Multiplication conditions (the two conditions in which instruction was given) than in the No-Instruction condition. This uneven distribution of Non-Gesturers reflects the fact that children who did not gesture on the pretest were more likely to begin gesturing on the training problems when given instruction than when not given instruction. This observation suggests that challenging children with explicit and (for them) novel information encourages them to begin gesturing. We suggest that the onset of gesturing reflects the fact the children have become actively engaged in thinking about the problem, and that active thinking about the problem is more likely when instruction is given.

Procedure

Each subject was tested individually by one of the authors (MWA) and an assistant in a quiet room in the school. Except for the paper and pencil tests taken before and after training and at follow-up, all components of the study were videotaped. The experimenter wrote each problem on the blackboard sufficiently high so that the child's body did not obscure his or her gestures, yet low enough so that the child could reach the problem and write an answer in the blank. The camera was placed at right angles to the blackboard, focused so that the child and the problem that he or she was working on were visible. Using this camera angle, the child's hand gestures to the problem, which were the focus of the study, were visible on videotape, and easily coded.

Pretest. The children in each of the three training groups were given a pretest consisting of two parts: (1) a paper and pencil test taken independently of the experimenter, and (2) a session at the blackboard with an experimenter. The paper and pencil test contained 14 problems: 6 addition problems of the form $4 + 3 + 8 = \_\_ + 8$; 4 addition problems of the form $3 + 6 + 5 = \_\_ + 4$ (i.e., without equivalent addends on the two sides of the equation); and 4 multiplication problems of the form $2 \times 3 \times 4 = \_\_ \times 4$. In every group of problems used in the study, the blank was adjacent to the equal sign in half of the problems, as in the preceding examples, and the blank was the last symbol in the problem in half of the problems (e.g., $3 + 6 + 4 = 3 + \_\_$ or $5 \times 3 \times 4 = 5 \times \_\_$). After the paper and pencil test, each child was asked to explain to an experimenter at the blackboard how he or she arrived at the solutions to the first six addition problems on the paper and pencil test. The experimenter wrote the first addition problem on the blackboard, including the child's solution to the problem, and asked, "How did you get the answer?" After the child responded, the experimenter erased that problem, wrote the second problem on the blackboard, and repeated the question. The pretest continued in this way until the child had explained his or her solutions to the first six addition problems.

Training session. After completing the pretest, each child was asked to solve a series of training problems and to explain his or her solutions to each problem. The training problems were administered at the blackboard by a second experimenter who was not present during the pretest.

The *training problems* were divided into three sets of six problems. The first two sets, consisting of 12 addition problems, were given to all three groups of children. The first six problems of the set were addition problems with equivalent addends on both sides of the problem (5 + 9 + 3 = __ + 3), and the second six problems were addition problems without equivalent addends (9 + 6 + 8 = __ + 7). The final set of six problems was given only to the children in the Addition-plus-Multiplication group and consisted of multiplication problems with equivalent multiplicands on both sides of the problem (5 × 3 × 2 = __ × 2).

The problem set was designed in this way to encourage children who used a specific procedure during the first part of the training to reevaluate that procedure. For example, if a child used a grouping procedure ("add the two numbers on the left side that are different from the number on the right side and put that sum in the blank"), the child would be able to correctly solve the first six problems in the training set (4 + 3 + 5 = __ + 5) but might have difficulty correctly extending this procedure to the next six problems (3 + 5 + 7 = __ + 6); the child might thus be forced to reevaluate this procedure and perhaps introduce a more general procedure during the next set of six problems. Similarly, a child could use the procedure "add all the numbers on the left side and subtract the number on the right from that sum" to arrive at correct solutions to both the first and second sets of addition problems, but could not easily extend this procedure to the multiplication problems. Note that only the procedure "make both sides of the problem equal" can be used to arrive at correct solutions to all 12 addition problems and 6 multiplication problems.

Instruction was given only to the children in the Addition and Addition-plus-Multiplication groups and emphasized the principle of equivalence. At the beginning of the training session, the experimenter wrote the first training problem on the blackboard and then told each child in these two groups that "the goal in solving these problems is to make both sides of the problem equal." She next provided each child with a principle-based procedure for solving the problem. She told the child that one way to make both sides equal is to find out how much there is on the left side, then look at how much there is already on the right side, and find a number to put in the blank which makes both sides the same. The experimenter then asked the child to solve the problem and to explain his or her solution. After the child's explanation, the experimenter gave the child feedback. If the child gave a correct solution and a correct explanation (or, in the infrequent event that the child gave a correct solution and an incorrect explanation), the experimenter said, "Very good! That's the right answer because it makes both sides of the problem equal."

If the child gave an incorrect solution and an incorrect explanation, the experimenter said, "That's a good try, but it's not the right answer because it doesn't make both sides equal." The experimenter then recapped the child's reasoning (basing the recap on the child's incorrect verbal explanation of the problem), and explained once again the definition of the equal sign. For example, if the child responded that he or she added all the numbers in the problem, the experimenter said, "It seems to me that you were thinking of the equal sign as an instruction to add up all the numbers in the problem, but that isn't what the equal sign means. Really the equal sign means 'is the same as.' It tells you to make both sides of the problem the same." In the infrequent event that the child solved the problem incorrectly but gave a correct verbal explanation, the experimenter said, "That's not the right answer but I think you're starting to think about the problem in the right way. Remember that the equal sign means 'is the same as,' and it tells you to make both sides of the problem the same." Finally, for all of the children who solved the problem incorrectly, the experimenter provided the correct answer to the problem and explained that "the reason it's the right answer is because it makes both sides of the problem equal." The experimenter then erased the problem, wrote the next problem on the blackboard, and asked the child to solve that problem and to explain his or her solution. The training session then continued in this way for the remainder of the training problems.

During the training session, the experimenter was careful to gesture only in a highly stylized fashion so that it would be readily apparent if the children were imitating her gestures in their verbal responses. Inspection of a subset of the videotapes showed that gestural responses which could possibly be interpreted as imitations of the experimenter's gestures accounted for fewer than 6% of the children's gestural responses, and no response was an exact copy of the experimenter's stylized gesture.

Posttest. Immediately following the training session, subjects in each of the three groups were asked to complete a 14-problem paper and pencil posttest comparable to the pretest.

Follow-up test. After an interval of approximately 2 weeks (mean 14 days; range, 11 to 21 days), subjects in each of the three groups were asked to complete a 14-problem paper and pencil follow-up test comparable to the pretest and posttest.

Coding Solutions

We calculated the number of addition and multiplication problems each child solved correctly on each paper and pencil test. In assessing children's knowledge of mathematical equivalence, we were interested in the procedures children followed in solving the problems and not in their skill at calculation. Since many of the children who knew a correct procedure for solving a problem nevertheless made calculation errors when carrying out that procedure (e.g., they said, "I added 3 plus 4 and got 8"), we counted as correct any solution which was within plus or minus 2 of the correct answer. It is important to note, however, that if we use a more stringent criterion for success, the number of children considered successful on the paper and pencil tests decreases but the pattern of results described in subsequent sections remains unchanged.

Coding Explanations

All verbal and gestural explanations of problem solutions were evaluated in terms of the procedure each conveyed. For each subject, verbal and gestural explanations were coded separately by two independent coders. Verbal explanations were coded by listening to the audio portion of the videotape only, without reference to the video portion (i.e., with the picture turned off). Gestural explanations were coded by viewing the video portion of the videotape only, without reference to the audio portion (i.e., with the sound turned off). Finally, the relationship between gesture and speech was evaluated by comparing the codes for the verbal and gestural components of a given explanation.

Coding types of explanations in speech alone. Each verbal explanation was coded according to the system initially described in Perry et al. (1988). Like the children in Perry et al.'s studies, the children in this study produced six basic types of spoken explanations (see Table 1): three which described procedures yielding incorrect solutions to problems of the form $4 + 3 + 5 = \_\_ + 5$, and three which described procedures yielding correct solutions to these problems. The incorrect explanations to addition problems were exemplified by the following three procedures: add all the numbers in the problem (*Add All*), add the numbers which appeared to the left of the equal sign (*Add to Equal*), or take one number from the left side of the equation and place it in the blank (*Carry*). The correct explanations were exemplified by the following three procedures: add the numbers which did not appear on both sides of the equation (*Grouping*), add the numbers on the left side of the equation and subtract the number on the right (*Add-Subtract*) or make both sides of the equation sum to the same total (*Equalizer*).

The same basic types of explanations were also used for multiplication problems. The incorrect explanations to multiplication problems were exemplified by three procedures

TABLE 1
Examples of Types of Procedures Expressed in Spoken Explanations and the Matching Gesture Accompanying Those Explanations
The math problem eliciting these explanations is: $4 + 6 + 9 = \_\_ + 9$

Type of procedure	Speech	Matching gesture
	Correct procedures	
Grouping	"The 9 was there so I added the 4 and 6"	Hand grabs below the 4 and 6, pause, point at solution
Add–Subtract	"I added 4 plus 6 plus 9 and that equals 19; to make both sides equal, I had to subtract the 9 so the answer is 10"	Point at 4, point at 6, point at 9 on the left side of the equation, pause, hand pulls down under the 9 on the right side of the equation, point at solution
Equalizer	"4 plus 6 plus 9 equals 19, so to make the other side equal 19, you need 10 more"	Sweep across the 4, 6, and 9 on the left side of the equation, point at the equal sign, sweep across the solution and 9 on the right side of the equation
	Incorrect procedures	
Add All	"I added 4 plus 6 plus 9 plus 9 equals 28"	Point at 4, point at 6, point at left 9, point at right 9, point at solution
Add to Equal	"I added 4 plus 6 plus 9 equals 19"	Point at 4, point at 6, point at left 9, point at solution
Carry	"They don't have another 4 like that so I put the 4 over there"	Point at the 4 on the left side of the equation, point at solution

comparable to the incorrect procedures used for addition problems (coded as *Multiply All*, *Multiply to Equal*, and *Carry*), and a fourth, infrequently used procedure (multiply the numbers on the left side of the equation, and subtract the number on the right, coded as *Multiply–Subtract*). The correct explanations to multiplication problems were exemplified by three procedures comparable to the correct procedures used for addition problems (coded as *Grouping*, *Multiply–Divide*, and *Equalizer*).

In addition, the children produced a small number of spoken explanations in which they described adding (or multiplying) particular subsets of numbers in the problem which did not conform to any of the above patterns. These explanations were coded as references to the particular numbers indicated and were considered incorrect. Spoken explanations which could not be assigned to a procedure category, or which did not indicate specific subsets of numbers in the problem, were classified as ambiguous (4% of the 1254 spoken explanations produced during the study by the Gesturers were classified as ambiguous).

Coding types of explanations in gesture alone. Gestures were transcribed by a second coder, using the lexicon of gestures established by Perry et al. (1988). Each of the verbal procedures described above had a counterpart in gesture (see Table 1). In addition, points to specific subsets of numbers which did not conform to any of the patterns in the procedures were coded as references to the particular numbers indicated, and were considered incorrect. Gestural explanations which could not be assigned to a procedure category, or which did not indicate specific subsets of numbers in the problem, were classified as ambiguous (9% of the 1013 gestured explanations produced during the study by the Gesturers were classified as ambiguous).

Coding the relationship between speech and gesture. In the final stage of coding, the verbal explanation and the gestural explanation given for each problem were compared. If the procedures given in gesture and in speech were different, the explanation was coded as a gesture–speech mismatch. If the procedures given in gesture and in speech were identical, the explanation was coded as a gesture–speech match. If no procedure was given in gesture (i.e., in speech alone responses), the explanation was also coded as a gesture-speech match. We classified no-gesture responses as gesture–speech matches simply because there is only one procedure expressed in a no-gesture response (the procedure expressed in speech), just as there is only one procedure expressed in a gesture–speech match (the procedure expressed in speech and in gesture). Support for this decision comes from the fact that, when children produced a procedure in speech in a no-gesture response, they typically produced that same procedure in gesture in some other response. In other words, that particular procedure could have been expressed by that child in gesture as well as in speech (i.e., it had the potential to be expressed by that child in a gesture–speech match). This result is not an obvious one given that, as we will show in later sections, many of the procedures a child produced in one modality were never produced by that child in the other modality.

As an example, in response to the problem, 4 + 7 + 5 = 4 + __, if a child said, "I added the 4, the 7, and the 5" (Add to Equal) while pointing to the left 4, the 7, and the 5 (Add to Equal), that response would be coded as a gesture-speech *match* since both gesture and speech conveyed the same procedure. Such a response suggests that the child was, in fact, entertaining only one (incorrect) hypothesis about how to solve the problem. Similarly, in response to the same problem, if a child said, "I added up 4 plus 7 plus 5 and it made 16, so I figured out what had to go with the 4 over here to make 16, and the answer was 12" (Equalizer) while first sweeping under the left side of the problem and then sweeping back and forth underneath the right 4 and the solution (Equalizer), that response would be coded as a gesture–speech match and would be considered evidence that the child was entertaining only one hypothesis (a correct one).

In contrast, in response to the problem, 4 + 7 + 5 = 4 + __, if a child said, "I added the 4, the 7, and the 5" (Add to Equal) while pointing to the left 4, the 7, the 5, and the right 4 (Add All), that response would be coded as a gesture-speech *mismatch* since the procedure conveyed in speech was not identical to the procedure conveyed in gesture. Such a response suggests that the child was entertaining two hypotheses about how to solve the problem. Similarly, in response to the same problem, if a child said, "I added the 4, the 7, and the 5" (Add to Equal) while first sweeping under the left side of the problem and then sweeping back and forth underneath the right 4 and the solution (Equalizer), that response would be coded as a gesture–speech mismatch and would be considered evidence that the child was entertaining two different hypotheses.

Coding a Child's State over Six Explanations

Explanations were grouped into sets of six: the six pretest problems, the first six training problems, the second six training problems, and (for the Addition-plus-Multiplication group) the third six training problems. Each set of six explanations was first evaluated in terms of the predominant type of gesture–speech response: match or mismatch. A child who produced 3 or more (i.e., 50% or more) explanations characterized by gesture–speech mismatch within a set of six explanations was considered to be in a "discordant" state with respect to that set. A child who produced fewer than 3 mismatches within a set of six explanations was considered to be in a "concordant" state with respect to that set. Note that we use the terms "discordant" and "concordant" to characterize the child's state with respect to a set of 6 explanations, and the terms "mismatch" and "match" to characterize the child's explanation of a single problem. The decision to treat discordance as categorical was, to a certain extent, arbitrary. However, it is important to note that, in previous work, we have shown that if discordance is considered to be a continuous variable (i.e., if children are classified according to the number of mismatches they produce), the phenomenon remains unchanged; in particular, the more mismatches a child produces on the pretest, the more likely that child is to be successful after instruction (Perry et al., 1988).

Sets of concordant responses were also categorized for correctness; sets in which the procedures expressed were predominantly incorrect were classified as "concordant incorrect," and sets in which the procedures expressed were predominantly correct were classified as "concordant correct." In most of the concordant sets of responses, it was quite easy to determine whether the responses were predominantly correct or incorrect since there was a clear majority of one response or the other. For the few concordant sets of responses which included an equal number of correct and incorrect responses ($N = 3$), we arbitrarily assigned these sets to a Concordant Incorrect state. It is important to note, however, that if these sets of responses are reclassified as Concordant Correct, the pattern of results described in subsequent section remains unchanged.

Reliability

Interrater reliability was established by having additional trained coders transcribe and code a subset of the videotapes. There was 93% ($N = 123$) agreement between coders on coding procedures in speech alone, 84% ($N = 121$) agreement on coding procedures in gesture alone, 92% ($N = 108$) agreement on coding type of explanation (the relationship between gesture and speech and the correctness of each explanation), and 94% agreement ($N = 18$) on coding child's state (Concordant Incorrect, Discordant, or Concordant Correct).

RESULTS

Success in Solving the Posttest and Follow-up Problems

Our first step was to determine whether the children who produced gestures had, in fact, learned to solve equivalence problems correctly. To assess success after instruction, we calculated the number of addition and multiplication problems each of the 63 Gesturers solved correctly on the posttest and follow-up test. Any child who solved 9 of the 10 addition problems correctly was considered to be successful on addition, and any child who solved 3 of the 4 multiplication problems correctly was considered to be successful on multiplication. We acknowledge that these are arbitrary definitions of "success." However, when we analyzed the data using alternative criteria for success (e.g., 10 out of 10 addition problems; 4 out of 4 multiplication problems), the patterns seen in Fig. 2 (as well as those seen in Figs. 4, 5, and 6) were unchanged. None of the children in any of the three training groups was successful on the addition problems on the *pretest*. Thus, if success was achieved after training, it reflected improvement in performance for each child.

We looked first at success on the addition problems. None of the children in the No Instruction group was successful on the addition problems

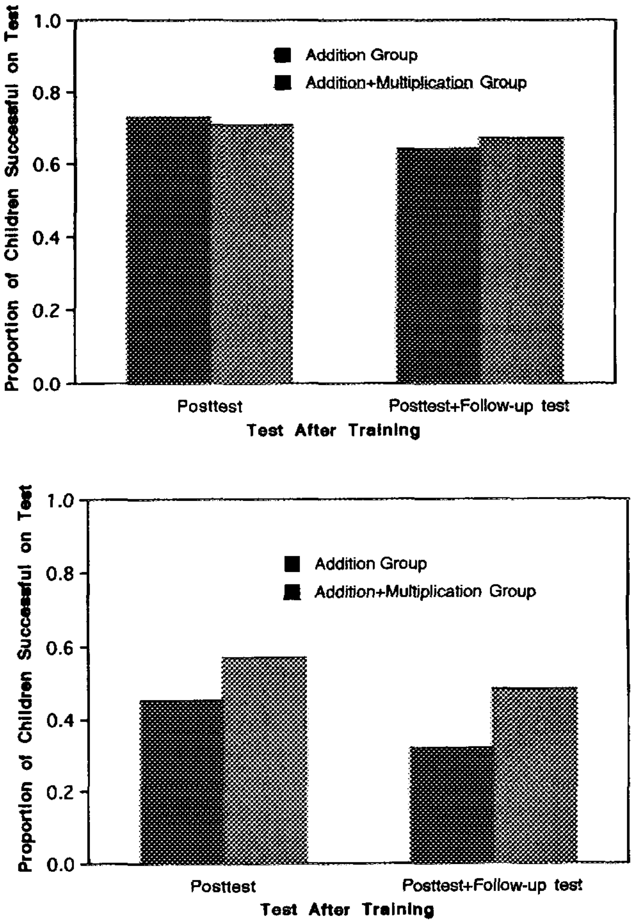

FIG. 2. The proportion of children in the Addition group and the Addition-plus-Multiplication group who were successful on the paper and pencil tests after training. The top panel displays the proportion of children successful on the addition problems on the posttest and the proportion successful on the posttest *and* the follow-up test. The bottom panel displays the proportion of children successful on *both* the addition and the multiplication problems on the posttest and the proportion successful on the posttest *and* the follow-up test.

after the training period. In contrast, a sizable proportion of the children in the two trained groups were successful on the addition problems on the posttest, and virtually all maintained that success on the follow-up test (see the top panel of Fig. 2). To investigate differences in success after the training period between the training groups, we partitioned the test of association on the 3 (groups) by 2 (success vs. no success) contingency table into two separate tests (see Bresnahan & Shapiro, 1966). First, we compared the No Instruction group to the two trained groups combined. As expected, fewer children in the No Instruction group than in the two groups who received instruction were successful on the addition problems after the training period ($\chi^2 = 28.4$, $df = 1$, $p < .001$ for success on the addition problems on the posttest; $\chi^2 = 23.4$, $df = 1$, $p < .001$ for success on addition problems on the posttest *and* the follow-up test). Second, we compared the two trained groups to one another, and found that the two groups did *not* differ in success on the addition problems after instruction ($\chi^2 = .007$, $df = 1$, n.s., for success on addition problems on the posttest; $\chi^2 = .04$, $df = 1$, n.s., for success on addition problems on the posttest *and* the follow-up test).

We next determined whether the children were able to generalize the knowledge exhibited on the addition problems to a set of multiplication problems. To do so, we assessed the proportion of children who succeeded not only on the addition problems, but also on the multiplication problems. Since none of the children in the No Instruction group succeeded on the addition problems, there were obviously none who succeeded on both the addition and multiplication problems. As a result, it is not surprising that the No Instruction group differed significantly from the two instructed groups in terms of success on both addition and multiplication after training ($\chi^2 = 15.7$, $df = 1$, $p < .001$ for success on addition and multiplication problems on the posttest; $\chi^2 = 10.8$, $df = 1$, $p < .001$ for success on addition and multiplication problems on the posttest *and* the follow-up test). We also compared the two trained groups to one another (see the bottom panel of Fig. 2), and found that the two groups did not differ in terms of success on addition and multiplication after instruction ($\chi^2 = .66$, $df = 1$, n.s., for success on addition and multiplication problems on the posttest; $\chi^2 = 1.4$, $df = 1$, n.s., for success on addition and multiplication problems on the posttest *and* the follow-up test).

Thus, as expected, instruction was indeed effective in fostering the acquisition of mathematical equivalence in the two groups which received instruction. A large proportion of the children in both groups succeeded on the posttest and follow-up addition problems, and many succeeded on the posttest and follow-up multiplication problems as well.

The Hypothesized Path of Development: Individual Children's Paths

Given that many of the children did learn to solve mathematical equivalence problems, we next examined the children's performance during the period of acquisition. We looked at the state which characterized a child at each of the observation periods, and evaluated whether the series of states through which the child passed was consistent with our hypothesized path of acquisition. As described above, we expected the children first to apply a single, incorrect procedure to a problem, and thus to be in a Concordant Incorrect state. We next expected the children to consider more than one procedure in solving a problem, and thus to enter a Discordant state. Finally, we expected the children again to apply a single, correct procedure to a problem, and thus to enter a Concordant Correct state.

To investigate whether individual children learned in a manner consistent with this hypothesized path, we analyzed the explanations they produced over the course of the pretest and training session. Each set of six consecutive explanations was assigned a state, as described above. For children in the No Instruction group and in the Addition group, this procedure yielded three assessments of states (one for the six pretest problems, and one for the first six addition problems in the training set, and one for the second six addition problems in the training set). For children in the Addition-plus-Multiplication group, this procedure yielded four assessments of states (one for the pretest problems, two for the addition training problems, and one for the six multiplication training problems). Finally, for each child, the progression of states, or the path, beginning from the pretest and continuing through each part of the training session was determined.

We begin by examining the paths followed by individual children in learning to solve addition equivalence problems. We therefore restrict this analysis to the explanations produced by each child on the 6 addition problems on the pretest and the 12 addition problems in the training set. We will examine the multiplication problems in the training set in a later section.

The Acquisition of Equivalence as it Applies to Addition Problems

There are many potential paths which an individual child might follow as he or she moves from the pretest through the training session. For the purposes of this study, we distinguished four different types of paths.

(Path 1) Progression through a discordant state. Any child who moved from a Concordant Incorrect state to a Discordant state, or from a Discordant state to a Concordant Correct state, without moving backward, was considered to have progressed in a manner consistent with our hypothesized path of acquisition. If a child progressed from a Concordant Incorrect state to a Concordant Correct state, the child was required to have passed through a Discordant state at the appropriate moment (i.e., after the Concordant Incorrect state and before the Concordant Correct state), again without moving backward, in order to be counted as following the hypothesized path.

(Path 2) Progression without moving through a discordant state. Any child who progressed from a Concordant Incorrect state to a Concordant

Correct state without moving through a Discordant state was considered to have progressed in a manner inconsistent with the hypothesized path.

(Path 3) Regression during training. Any child who moved from a Concordant Correct state to either a Discordant state or a Concordant Incorrect State, or from a Discordant state to a Concordant Incorrect state, was classified as having regressed.

(Path 4) Staying in the same state. Any child who remained in the same state throughout the pretest and both training assessments was considered to have stayed in place.

The paths taken by children with and without instruction. Table 2 presents the proportion of children who followed each of the four types of paths. The children are divided into those who did not receive instruction during the training session (i.e., the No Instruction group) and those who did (i.e., the Addition and Addition-plus-Multiplication groups). The Appendix displays all of the paths taken by children, with and without instruction, and the proportion of children in each group who followed each path.

As can be seen in Table 2, the children without instruction, who had not learned to solve the problems correctly during training, differed in several respects from the children with instruction, many of whom had learned to solve the problems correctly. Not surprisingly, the children who did not receive instruction were more likely to *stay* in the same state than were the children who did receive instruction: 50% (10) of the 20 children without instruction vs. 19% (8) of the 43 children with instruction remained in the same state ($\chi^2 = 6.6$, $df = 1$, $p \leq .01$). Again, as might be expected, the children who did not receive instruction were also more likely to *regress* than were the children who did receive instruction: 30% (6) of the 20 children without instruction vs. 9% (4) of the 43 children with instruction regressed ($\chi^2 = 4.4$, $df = 1$, $p < .05$).

Finally, the children who did not receive instruction were less likely to *progress* through a discordant state than the children who did receive

TABLE 2
Proportion of Children with and without Instruction Classified According to the Type of Path Taken from the Pretest through the Addition Training Problems

Path taken from pretest through addition training problems	Children who received no instruction ($N = 20$)	Children who received instruction ($N = 43$)
Stay in the same state	.50	.19
Regress during training	.30	.09
Progress through a Discordant state	.15	.60
Progress without moving through a Discordant state	.05	.12

instruction: 15% (3) of the 20 children without instruction vs. 60% (26) of the 43 children with instruction progressed through a discordant state ($\chi^2 = 11.4$, $df = 1$, $p < .001$). Few children in either group made progress which involved skipping a discordant state: 5% (1) of the 20 children without instruction vs. 12% (5) of the 43 children with instruction progressed without passing through a discordant state ($\chi^2 = .70$, $df = 1$, n.s.). In fact, of the 35 children (in both the instructed and the uninstructed groups) who made progress over the observation sessions, 83% (29) did so via a discordant state—a number significantly higher than that expected by chance ($p < .001$, Binomial Test). The 29 children who progressed via a discordant state reflected all segments of the hypothesized path of acquisition: 11 children progressed from a Concordant Incorrect state to a Discordant state (2 in the No Instruction group, and 9 in the Instruction group), 15 progressed from a Discordant state to a Concordant Correct state (1 in the No Instruction group, and 14 in the Instruction group), and 3 progressed from a Concordant Incorrect state through a Discordant state to a Concordant Correct state (all in the Instruction group).

Transitional probabilities between states. We turn next to an analysis of transitions between states. Given that a child has arrived at a certain state, how likely is that child to stay in that state, move forward, or move backward? Figure 3 presents transitional probabilities for different moves given that a child began the transition in a Concordant Incorrect state (top graph), a Discordant state (middle graph), or Concordant Correct state (bottom graph), and presents those probabilities separately for children who did not receive instruction and for those who did. These figures summarize data from both transitions made by each child (i.e., the transition from the pretest to the first six problems of the training set, and the transition from the first six to the second six problems of the training set).

Focusing first on the children who did *not* receive instruction during the training period, note that, if a child began in a *Concordant Incorrect* state and did not receive instruction, the child was very likely to stay in that state. The one child in the uninstructed group who began a transition in a *Concordant Correct* state also remained in that state. In contrast, if a child began in a *Discordant* state and received no instruction, the child was relatively likely to move out of that state. Moreover, when children who began a transition in a Discordant state moved, they tended to regress back to a Concordant Incorrect state. Thus, if no instruction is given, the discordant state appears to be relatively unstable, particularly in comparison to the concordant state.

We turn next to the children who *did* receive instruction during the training period. Not surprisingly, children who began a transition in a *Concordant Correct* state were likely to stay in that state. In contrast, children who began in a *Concordant Incorrect* state and were given instruction were likely to move out of that state. The majority of these children progressed forward to a Discordant state, as predicted. A smaller number of children progressed directly to a Concordant Correct state without passing through discordance; that is, they violated the hypothesized path. It is worth noting, however, that the children who progressed to a correct state without passing through discordance did quite poorly on the post test and follow-up test, particularly on the multiplication problems which required generalizing the knowledge gained during training (see below); thus, these children may have acquired a more superficial understanding of equivalence than children who passed through discordance.

Children who began in a *Discordant* state and were given instruction were also relatively likely to move out of that state. In fact, children who began in a Discordant state moved out of that state about equally often when given instruction as when *not* given instruction. However, the Discordant children who received instruction progressed *forward* to a Concordant Correct state, while the Discordant children who did not receive instruction regressed *backward* to a Concordant Incorrect state. Thus, with or without instruction, a child in a Discordant state is relatively likely to move out of that state. If provided with instruction, the child will move to a more stable, correct state; if not provided with instruction, the child will also move to a more stable state, but it is very likely to be an incorrect state.

A somewhat surprising result seen in Fig. 3 is that, when given instruction, Discordant children were *more* likely to remain in the same state than Concordant Incorrect children (compare the "Stay" bars in the graphs on the right side of the top and middle panels). However, if we look more closely at the types of responses the Discordant children produced over the course of the training, we find that many of them did indeed show progress, although that progress was *within* the Discordant state. There were 30 transitions in which children who received instruction stayed in a Discordant state (i.e., the children produced a large proportion of gesture–speech mismatches both before and after the transition). However, in 16 of these transitions (53%), the children's predominant response was a different type of gesture–speech mismatch before and after the transition. In 13 of the 16 transitions, the children produced a more advanced type of mismatch after the transition than before (more advanced in the sense that the new mismatch contained more correct procedures than the previously produced mismatch, see examples in Table 3). In only 3 of the 16 transitions did the children produce a less advanced type of mismatch after the transition—significantly fewer than expected by chance ($p < .02$, Binomial Test).

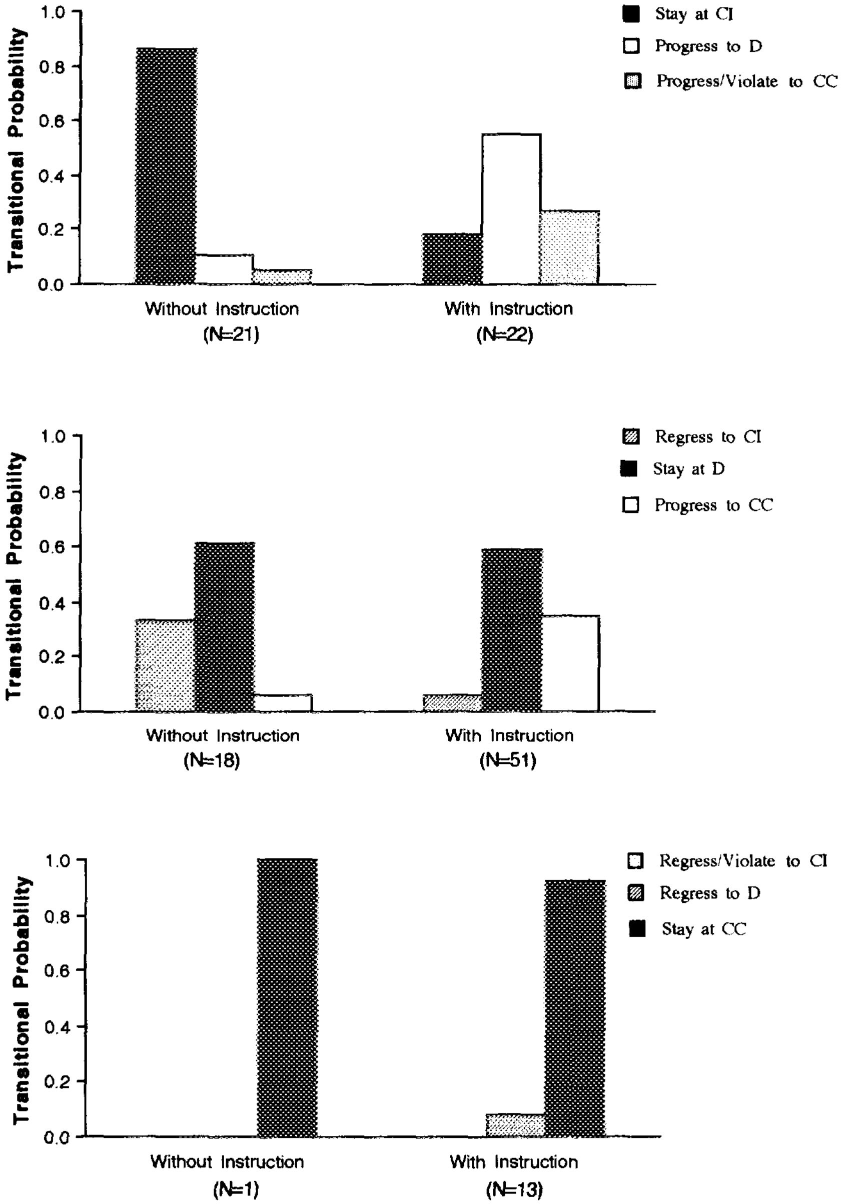

FIG. 3. The probability that a child will move to a Concordant Incorrect, Discordant, or Concordant Correct state on the addition problems, given that the child began the transition in a Concordant Incorrect state (top panel), in a Discordant state (middle panel), or in a Concordant Correct state (bottom panel). Probabilities are presented separately for children who received no instruction during training and for children who did receive instruction. The N's in each graph represent the number of transitions beginning at that particular state. CI, Concordant Incorrect; D, Discordant; CC, Concordant Correct.

TABLE 3
Examples of Gesture–Speech Mismatches[a]
The math problem eliciting these explanations is: $3 + 4 + 5 = \_\_ + 5$

Type of gesture–speech mismatch	Gesture	Speech
Two incorrect procedures	Point to 3, to 4, to left 5, to right 5 [Add All]	"I added 3 + 4 + 5 and came up with 12" [Add to Equal]
One incorrect and one correct procedure	Sweep under the left side of the problem, sweep under the right side of the problem [Equalizer]	"I added 3 + 4 + 5 and I got 12" [Add to Equal]
One incorrect and one correct procedure	Point to 3, to 4, to left 5 to right 5 [Add All]	"I just added 3 + 4 so I put 7" [Grouping]
Two correct procedures	Sweep under the left side of the problem, sweep under the right side of the problem [Equalizer]	"I just added 3 + 4 so I put 7" [Grouping]

[a] The brackets contain the name of the procedure that each string of gestures and that each string of words exemplifies. In the first example, the procedure in gesture (Add All) leads to an incorrect solution and the procedure in speech (Add to Equal) leads to a different, but also incorrect solution. In the second example, the procedure in gesture (Equalizer) leads to a correct solution but the procedure in speech (Add to Equal) leads to an incorrect solution. In the third example, the procedure in gesture (Add All) leads to an incorrect solution but the procedure in speech (Grouping) leads to a correct solution. In the final example, the procedure in gesture (Equalizer) leads to a correct solution and the procedure in speech (Grouping) leads to a different, but also correct solution.

These data stand in contrast to the 11 transitions in which children who did *not* receive instruction remained in a Discordant state. In none of these 11 transitions did a child produce more advanced gesture–speech mismatches after the transition ($\chi^2 = 7.3$, $df = 1$, $p < .01$). Thus, when children who did not receive instruction remained in a Discordant state, they typically continued to produce the same type of mismatch, or produced a less advanced type of mismatch. In contrast, when children who did receive instruction remained in a Discordant state, they often changed the type of mismatch they produced to a more advanced type.

To summarize thus far, the transitional probabilities in Fig. 3 suggest that two factors contribute to whether or not children change their state: the type of input the child receives (in this case, the presence or absence of instruction), and the child's state at the time the input is administered. The data also suggest that, when a child receives instruction and changes state, that child tends to move in accordance with the hypothesized path. If children received instruction and moved between a *Concordant Incor-*

rect state and a *Discordant* state, they were more likely to move *from* the Concordant Incorrect state *to* the Discordant state than the reverse (12 of the 15 moves between these two states were in the predicted direction, $p < .02$, Binomial Test). In contrast, if children received instruction and moved between a *Concordant Correct* state and a *Discordant* state, they were more likely to move *from* the Discordant state *to* the Concordant Correct state than the reverse (17 of the 18 moves between these two states were in the predicted direction, $p < .001$, Binomial Test). Thus, the Discordant state appears to *follow* the Concordant Incorrect state but to *precede* the Concordant Correct state.

These results suggest that the children in this study who learned to correctly solve addition equivalence problems did so by adhering to the path which we have suggested characterizes acquisition in general: They began with a single incorrect hypothesis, moved through a period in which they concurrently considered multiple hypotheses, and concluded with a single correct hypothesis. We next examine what happens when children first learn about equivalence as it applies to addition problems, and then are required to generalize their understanding to multiplication problems. We ask whether children regress at this point in the acquisition process and, if so, how far back they move.

The Acquisition of Equivalence as It Applies First to Addition Problems and Subsequently to Multiplication Problems: Regression in the Face of a New Problem Type

Recall that the children in the Addition-plus-Multiplication group, in addition to being asked to solve the 12 addition problems during training, were also asked to solve 6 multiplication problems. This group thus allows us to explore how the children dealt with a new problem type (i.e., multiplication), as a function of the state the child had achieved at the end of the addition problems. We looked first at the 10 children in the Addition-plus-Multiplication group who had reached a Discordant state at the end of the 12 addition problems on the training set. We found that 8 of these 10 children (80%) remained in a Discordant state on the multiplication problems, and 2 (20%) progressed forward to a Concordant Correct state. Thus, children who entertained multiple hypotheses on the addition problems either continued to entertain multiple hypotheses on the multiplication problems, or narrowed in on a single correct procedure. *None* of the children entertained a single, incorrect procedure (i.e., regressed back to a Concordant Incorrect state) in the face of a new problem type.

We next turned to the 9 children in the Addition-plus-Multiplication group who had reached a Concordant Correct state at the end of the addition problems. Unlike the children in a Discordant state, these children had narrowed in on a single, correct procedure for solving mathematical equivalence problems at the end of the addition problems. The question then was whether they would continue to entertain a single, correct procedure on the multiplication problems, or regress back to a state in which they entertained a variety of procedures. We found that 44% (4) of the 9 children in a Concordant Correct state at the end of the addition problems remained in that state on the multiplication problems, while 56% (5) regressed. Importantly, we also found that *none* of the 5 children who regressed entertained a single incorrect procedure on the multiplication problems (i.e., none moved back to a Concordant Incorrect state); instead, all 5 began to entertain multiple hypotheses on the multiplication problems. That is, they regressed only as far back as the Discordant state. These data provide some of the strongest evidence that children tend to pass through a discordant state when moving out of a concordant state.

Transitions in Progression and Regression: Passing through a Discordant State

We have shown that if a child begins a transition in a Discordant state and is given instruction, the child is likely to remain in that state or to progress to a Concordant Correct state; if a child begins in a Discordant state and is not given instruction, the child is likely to stay in that state or regress back to a Concordant Incorrect state. Note that these findings, although consistent with the hypothesized path, do not address whether or not the child *must* pass through a discordant state in moving from one concordant state to another. To explore this question, we need to examine the moves children make when beginning a transition in either a Concordant Incorrect or a Concordant Correct state.

We found that, across all three training groups and across the entire training period, children began a transition in either a Concordant Incorrect or a Concordant Correct state 68 times. In 40 of these 68 transitions, the child remained in the concordant state in which he or she began; these 40 transitions do not bear directly on our question. We turn therefore to the remaining 28 transitions in which children began in a Concordant state and moved out of that state. The question we ask is whether, when moving out of a Concordant state, the child is likely to move to a Discordant state (as we predict), or to the other Concordant state. We found that on 20 occasions (out of 28) the children moved from a Concordant state to a Discordant state. Thus, the children violated our predictions on only 8 occasions (out of 28)—a significantly smaller number than expected by chance ($p < .02$, Binomial Test). These data suggest that children are very likely to pass through a discordant state when moving out of a concordant state, and thus provide support for the hypothesized path.

Gesture as a Window into Transitional States: Post-training Performance as a Function of Final State on the Path

We have shown that, when evaluated in relation to a child's speech, gesture provides a window into the transitional state a child passes through in acquiring mathematical equivalence. However, it is important to note that we did not need to look at gesture in order to learn that the children had benefited from instruction. Most of the children in the groups receiving instruction gave incorrect verbal explanations before instruction, and correct verbal explanations after instruction; that is, most of the children would have been classified as having learned the task if traditional measures which rely on speech alone were used to assess learning. What then have we gained by looking at gesture in the children's explanations?

If we consider verbal performance alone, note that there is no reason to expect children to differ on their post-training assessments; all of the children who produced correct verbal explanations after instruction would be expected to perform alike on the posttest and follow-up test assessments. Our claim, however, is that, despite that fact that children appear to show the same understanding when their speech alone is examined, there are differences among the children in their understanding of mathematical equivalence—differences which are reflected in the way in which gesture and speech are related in the children's explanations. Thus, we predict that, among the group of children (both concordant and discordant) who produced correct *verbal* explanations at the end of training, the children who had reached a Discordant state (i.e., children whose gestures did not match their *correct* spoken explanations) ought to have a less developed understanding of mathematical equivalence than the children who had reached a Concordant Correct state (i.e., children whose gestures did match their correct spoken explanations).

To test this prediction, we examined the post-training performance of two groups of children, all of whom had received instruction during the training period and all of whom had produced correct procedures in *speech* after instruction: (1) Children who progressed from a Concordant Incorrect state to a Discordant state during training ($N = 7$) and (2) children who progressed through a Discordant state to a Concordant Correct state ($N = 15$).[2] We calculated the number of children in each group

---

[2] Since the posttest was given after the entire training period, it was necessary to determine progress along the path not just on the 12 addition problems but (for the Addition-plus-Multiplication group) on the 6 multiplication problems as well. As a result, the number of progressors included in Fig. 3 (which describes children who progressed along the path from the pretest through the *entire* training set, including the 6 multiplication problems for the Addition-plus-Multiplication group) is different from the number of progressors listed in

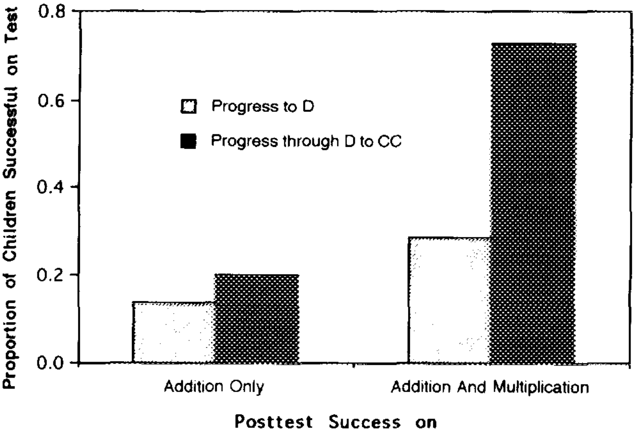

FIG. 4. The proportion of children who were successful on the paper and pencil posttest, classified according to the type of progress they exhibited during training: (1) Children who progressed to a Discordant state during training and (2) children who progressed through a Discordant state to a Concordant Correct state during training. All of the children in both groups produced correct verbal explanations after instruction. The figure presents the proportion of children who were successful *only* on the addition problems on the posttest and the proportion of children who were successful on *both* the addition and multiplication problems on the posttest. D, Discordant; CC, Concordant Correct.

who were successful on the addition and multiplication problems on the posttest. Figure 4 presents the proportions of children in each group who succeeded on *only* the addition problems (and not the multiplication problems) on the posttest and the proportion of children in each group who succeeded on *both* the addition and multiplication problems on the posttest. We found that an equal proportion of children in the two groups was successful on the addition problems only. However, the two groups differed significantly in terms of success on *both* the addition and multiplication problems on the posttest. Children who had attained the Concordant Correct state succeeded on both the addition and multiplication problems on the posttest significantly more often than children who had reached the Discordant state ($\chi^2 = 4.0$, $df = 1$, $p < .05$). Thus, although all of the children in these two groups produced correct procedures in their speech after training, children who produced correct, matching procedures in gesture at the end of training were significantly more successful than children who produced incorrect, mismatching procedures in gesture. This difference between the groups was maintained through the

---

Table 2 (which describes progress through the 12 addition problems for all subjects). In particular, four children progressed through the addition problems and then regressed on the multiplication problems, and these children are not included in Fig. 3.

follow-up test, although it was no longer significant when data from both the posttest and the follow-up were considered ($\chi^2 = 1.2$, $df = 1$, n.s.).[3]

Thus, there are differences among the children who made progress during training in terms of the depth of the children's understanding of mathematical equivalence at the end of the study. Moreover, since all of the children whose data are displayed in Fig. 4 produced correct verbal explanations after instruction, these differences *cannot* be predicted on the basis of the children's speech. The differences can, however, be predicted if the children's speech is assessed in relation to the gesture produced along with it.

The Effect of Skipping the Discordant State on Post-training Performance

Recall that, although few, there were some children who progressed from a Concordant Incorrect state directly to a Concordant Correct state without passing through a Discordant state. We suggest that these children, in skipping the discordant state, may have developed a relatively superficial understanding of mathematical equivalence. Children who skip the discordant state switch directly from entertaining a single, incorrect procedure to entertaining a single, correct procedure. In adopting a new procedure and discarding an old procedure so quickly, they may not take time to consider the implications of the new procedure or to consider how it fits together with the other procedures in their repertoire. In this respect, they may acquire only a superficial algorithm for solving the specific problem type at hand, rather than a deeper understanding which they could then apply to other problem types. They may therefore perform less well on post-training assessments than children who reach a Concordant Correct state via a state in which they do entertain multiple procedures. To explore this question, we examined the post-training performances of the 5 children who progressed to a Concordant Correct state without passing through a Discordant state, and compared them to children who progressed to a Concordant Correct state via discordance.

Figure 5 presents the proportion of children who succeeded on *only* the addition problems on the posttest and the proportion of children who succeeded on *both* the addition and multiplication problems on the post-

---

[3] Recall that the children in the Addition-plus-Multiplication condition were slightly (although not significantly) more likely to do better on the post-training assessments (particularly on the multiplication problems) than the children in the Addition condition (see Figure 2). It is therefore important to point out that the difference between the groups in success on both the addition and multiplication problems seen in Figure 4 (and in Figures 5 and 6 as well) remains the same when the children are divided by condition (i.e., when the comparisons are done separately for children in the Addition condition and for children in the Addition-plus-Multiplication condition).

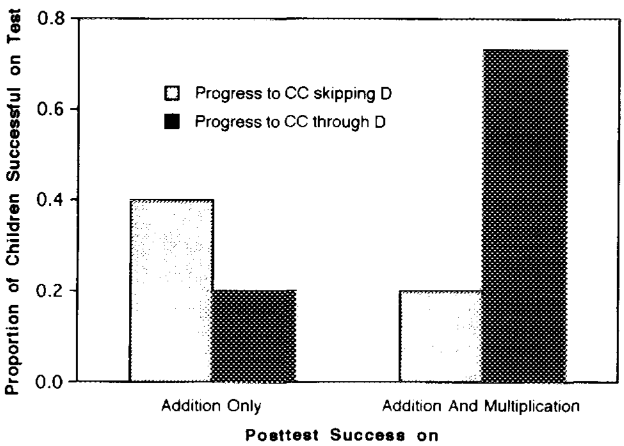

FIG. 5. The proportion of children who were successful on the paper and pencil posttest, classified according to whether, during training, the child progressed to a Concordant Correct state by passing through a Discordant state or by skipping a Discordant state. The figure presents the proportion of children who were successful *only* on the addition problems on the posttest and the proportion of children who were successful on *both* the addition and multiplication problems on the posttest. D, Discordant; CC, Concordant Correct.

test, as a function of the type of progress the child had made in reaching a Concordant Correct State: (1) Children who progressed to a Concordant Correct state by skipping a Discordant state ($N = 5$) and (2) children who progressed to a Concordant Correct state by passing through a Discordant state ($N = 15$). We found that many of the children who progressed to a Concordant Correct state by skipping discordance were successful on the addition problems on the posttest. However, these children were significantly less likely than the children who progressed to a Concordant Correct state via discordance to succeed on *both* the addition and multiplication problems on the posttest ($\chi^2 = 4.4$, $df = 1$, $p < .05$). Moreover, this difference between the groups was maintained through the follow-up test ($\chi^2 = 4.4$, $df = 1$, $p < .05$).

As predicted, the children who achieved a Concordant Correct state by passing through discordance were able to generalize the knowledge gained on addition to multiplication and to maintain that knowledge over a 2-week period, while children who achieved a Concordant Correct state without passing through discordance were unable to do so. Children who learned to correctly solve equivalence problems by moving through a period in which they entertained multiple hypotheses thus appear to have developed a deeper understanding of mathematical equivalence than children who learned by switching from a single incorrect procedure to a single correct procedure.

The Repertoire of Procedures Children Produce as They Progress along the Path

In this study, children were asked to explain their solutions to mathematical equivalence problems. For each problem, each child described 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 solved the math problem $4 + 5 + 3 = \_\_ + 3$ by "adding the four, the five, the three, and the three" has a representation of the problem which includes all four numbers, with no meaningful subgroupings within the numbers. In contrast, a child who says she solved the same problem by "adding the four, the five, and the three" is assumed to have a representation of the problem which 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 since, 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.

Although the discordant child, by definition, is working with two different representations of a single problem, it is important to point out that the definition of discordance does *not* require that the discordant child possess a larger number of different representations than the concordant child. For example, consider a discordant child who, on the first problem, produces one procedure in gesture (e.g., Add All) and a different procedure in speech (e.g., Add to Equal) and then, on the second problem, reverses the pattern (i.e., he produces Add to Equal in gesture and Add All in speech). Such a child would have produced two gesture-speech mismatches, and those mismatches provide evidence for two different procedures (reflecting two different representations of the problem) in the child's repertoire of procedures. Now consider a concordant child who, on the first problem, produces Add All in both gesture and speech and then, on the second problem, produces a different procedure, Add to Equal, in both gesture and speech. Unlike the discordant child, this child would have produced two gesture–speech matches; however, these matches also provide evidence for two different procedures in the child's repertoire. In other words, despite the fact that the discordant child produced mismatches and the concordant child produced matches, both children may possess the same number of different procedures in their repertoires. Thus, even though the definition of discordance entails that discordant children produce two different procedures on a single problem, the definition does *not* require that discordant children produce more different types of procedures in their repertoires overall than concordant children.

However, in previous analyses (Goldin-Meadow, Alibali & Church, 1993), we have found that, prior to instruction, discordant children do indeed display significantly more types of procedures in their repertoires than concordant children. Moreover, the additional procedures that discordant children possess are expressed predominantly in gesture and not speech. The question we address here is whether a child's repertoire of procedures changes over the course of instruction. In particular, does a child's repertoire of procedures expand or decrease as that child progresses along the path toward mastery of a concept?

Given that, prior to instruction, concordant children tend to display smaller repertoires than discordant children, we would predict based on our previous work that, as a child progresses from a Concordant Incorrect state to a Discordant state, that child's repertoire ought to expand (with expansion most likely in procedures found only in gesture). Conversely, as a child progresses from a Discordant state to a Concordant Correct state, we predict that the child's repertoire would contract (with contraction most evident in procedures found only in gesture).

We tested these predictions by examining the repertoires of procedures produced by the children who received instruction and who progressed along the path. Recall that 26 children progressed through discordance on the addition problems: 9 progressed from a Concordant Incorrect state to a Discordant state, 14 progressed from a Discordant state to a Concordant Correct state, and 3 progressed from a Concordant Incorrect state through a Discordant state to a Concordant Correct state. Thus, there were 12 children who made the transition from a Concordant Incorrect state to a Discordant state, and 17 children who made the transition from a Discordant state to a Concordant Correct state. We calculated the total number of different procedures each child produced before and after the transition. The data are displayed in Table 4.

As predicted, the mean number of *total procedures* produced by children when they were in a concordant state (either Concordant Incorrect or Concordant Correct) was smaller than the mean number produced when they were in a discordant state. With respect to the first transition, the 12 children who progressed from a Concordant Incorrect state to a

TABLE 4
Mean Number of Different Procedures Produced by Children Who Progressed through Discordance

	Children who progressed from a Concordant Incorrect state to a Discordant state		Children who progressed from a Discordant state to a Concordant Correct state	
	(N = 12)		(N = 17)	
	Concordant Incorrect state	Discordant state	Discordant state	Concordant Correct state
Total number of different procedures produced	1.9 (.7)[a]	3.2 (1.0)***	3.2 (1.1)	2.0 (.8)***
Number of different procedures produced in both gesture and speech	.9 (.8)	1.5 (1.1)*	1.4 (1.0)	1.1 (.7)
Number of different procedures produced in speech but not gesture	.7 (.9)	.7 (.8)	.7 (.7)	.5 (.6)
Number of different procedures produced in gesture but not speech	.3 (.5)	1.0 (.8)**	1.1 (.8)	.4 (.5)**

[a] The numbers in parentheses are standard deviations.
* $p \leq .05$, Paired t test, comparing children in a concordant and discordant state.
** $p \leq .01$.
*** $p \leq .001$.

Discordant state produced significantly *fewer* different procedures when Concordant Incorrect than when Discordant (t = 4.8, df = 11, $p < .001$). Moreover, 11 of the 12 children followed the predicted pattern, and increased their repertoires as they progressed to a Discordant state. With respect to the second transition, the 17 children who progressed from a Discordant state to a Concordant Correct state produced significantly *more* different procedures when Discordant than when Concordant Correct (t = 4.0, df = 16, $p \leq .001$). Moreover, 13 of the 17 children followed the predicted pattern and decreased their repertoires as they progressed to a Concordant Correct state. Thus, as predicted, children tended to *increase* the size of their repertoires as they moved from a concordant state in which they had an incorrect understanding to a discordant state. They then *decreased* the size of their repertoires as they moved from this discordant state to a concordant state in which they had a correct understanding.

In addition to the total number of procedures produced, we also examined the modality in which each procedure was expressed. We calculated the number of different procedures that were expressed in gesture and never in speech, the number of expressed in speech and never in gesture, and the number expressed in both gesture and speech. Procedures did not have to be expressed in gesture and speech on the same problem in order to be considered part of the gesture-speech repertoire; the procedure had only to appear at some time in gesture and at some time in speech. As before, we calculated the number of different procedures children produced when in a Concordant Incorrect state, when in a Discordant state, and when in a Concordant Correct state. The data are displayed in Table 4.

We found that children in a discordant state had more different types of procedures in *both gesture and speech* than children in a concordant state, although this difference was significant only when the children moved from Concordant Incorrect to Discordant (t = 2.2, df = 11, $p <$ .05), and not when they moved from Discordant to Concordant Correct (t = 1.9, df = 16, n.s.). We found no reliable differences in the number of procedures produced in *speech but not gesture* for either transition (t = 0, df = 11, n.s., for the transition from Concordant Incorrect to Discordant; t = 1.2, df = 16, n.s., for the transition from Discordant to Concordant Correct). In contrast, we did find significant differences in the number of procedures produced in *gesture but not speech* during both transitions. With respect to the first transition, the 12 children who progressed from a Concordant Incorrect state to a Discordant state produced significantly *fewer* different procedures in gesture alone when Concordant Incorrect than when Discordant (t = 3.1, df = 11, $p <$ .01). Nine of the 12 children followed this pattern and increased the number of procedures found exclusively in gesture as they progressed to a Discordant state. With respect to the second transition, the 17 children who progressed from a Discordant state to a Concordant Correct state produced significantly *more* different procedures in gesture alone when Discordant than when Concordant Correct (t = 3.0, df = 16, $p <$ .01). Eleven of the 17 children followed this pattern and decreased the number of procedures found exclusively in gesture as they progressed to a Concordant Correct state. Thus, the increase in repertoire size associated with entry into a discordant state appears to be attributable primarily to an increase in the number of procedures which are produced in gesture and not in speech.

The Repertoire of Procedures Children Produce as They Regress along the Path

We have found that, when children *progress* in their acquisition of mathematical equivalence, their repertoires increase when they enter a discordant state and decrease when they exit a discordant state. We next ask what happens to children's repertoires when they *regress* in their understanding. If regression is the same process at all points in development, we might expect that children's repertoires would change in the same way—independent of the type of regression. However, if our model of transition is correct, we would predict that, as children regress from a Concordant Correct state to a Discordant state, their repertoires ought to *increase* in size. In contrast, as children regress from a Discordant state to a Concordant Incorrect state, their repertoires ought to *decrease* in size. We tested these predictions against the available data on regression.

Recall that there were five children who, when given the multiplication problems during training, regressed from a Concordant Correct state to a Discordant state. In addition, there were six children in the No Instruction condition who regressed from a Discordant state to a Concordant Incorrect state during the training period. We calculated the total number of different procedures these children produced when in a Concordant Correct state, a Discordant state, or a Concordant Incorrect state. The means for the *total number of procedures* produced are presented in Table 5. As predicted, the five children who regressed from a Concordant Correct state to a Discordant state produced significantly *fewer* proce-

TABLE 5
Mean Number of Different Procedures Produced by Children Who Regressed through Discordance

	Children who regressed from a Concordant Correct state to a Discordant state		Children who regressed from a Discordant state to a Concordant Incorrect state	
	($N = 5$)		($N = 6$)	
	Concordant Correct state	Discordant state	Discordant state	Concordant Incorrect state
Total number of different procedures produced	2.0 (1.0)[a]	3.8 (.5)**	3.0 (.9)	1.6 (.5)*
Number of different procedures produced in both gesture and speech	1.0 (.7)	1.2 (.5)	.8 (.8)	.2 (.4)
Number of different procedures produced in speech but not gesture	.4 (.6)	.6 (.9)	.5 (.6)	1.2 (.8)
Number of different procedures produced in gesture but not speech	.6 (.6)	2.0 (1.2)*	1.7 (.9)	.3 (.4)*

[a] The numbers in parentheses are standard deviations.

* $p \leq .05$, Paired t test, comparing children in a concordant and discordant state.

** $p \leq .01$.

dures when Concordant Correct than when Discordant ($t = 4.8$, $df = 4$, $p < .01$). Moreover, all five of the children showed the predicted pattern and increased their repertoires as they regressed to a Discordant state. Also as predicted, the six children who regressed from a Discordant state to a Concordant Incorrect state produced significantly *more* procedures when Discordant than when Concordant Incorrect ($t = 3.8$, $df = 5$, $p \leq .01$). Moreover, five of the six children showed the predicted pattern and decreased their repertoires as they regressed to a Concordant Incorrect state.

As in our analyses of change in repertoire size during progress, in addition to the total number of procedures produced, we also examined the modality in which each procedure was expressed. The data are displayed in Table 5. Note that there were no differences in the number of different procedures produced in *both gesture and speech* when children regressed from Concordant Correct to Discordant ($t = .41$, $df = 4$, n.s.), nor were there differences when children regressed from Discordant to Concordant Incorrect ($t = 1.9$, $df = 5$, n.s.). Moreover, there were no differences in the number of different procedures produced in *speech but not gesture* when children regressed from Concordant Correct to Discordant ($t = .54$, $df = 4$, n.s.), nor were there differences when children regressed from Discordant to Concordant Incorrect ($t = 2.0$, $df = 5$, n.s.).

In contrast, for both types of regression, there were differences in the number of different procedures the children produced in *gesture but not speech* when in each of the two states. The children who regressed from a Concordant Correct state to a Discordant state produced significantly *fewer* different procedures uniquely in gesture when Concordant Correct than when Discordant ($t = 3.5$, $df = 4$, $p < .05$). Moreover, four of the five children showed the predicted pattern and increased the number of different procedures they produced uniquely in gesture when they regressed to a Discordant state. Similarly, the children who regressed from a Discordant state to a Concordant Incorrect state produced significantly *more* different procedures uniquely in gesture when Discordant than when Concordant Incorrect ($t = 3.6$, $df = 5$, $p < .02$). Moreover, five of six children showed the predicted pattern and decreased the number of different procedures they produced uniquely in gesture when they regressed to a Concordant Incorrect state. Thus, as in progression, the change in repertoire size associated with entry into, and exit from, a discordant state during regression appears to be attributable primarily to a change in the number of procedures which are produced in gesture and not in speech.

The data on regression in Table 5, taken in conjunction with the data on progression in Table 4, suggest that children increase the size of their repertoires (in particular, the number of procedures encoded exclusively in gesture) as they enter the discordant state. Moreover, children decrease the size of their repertoires (primarily the number of procedures encoded exclusively in gesture) as they exit the discordant state. These changes in repertoire size appear to occur regardless of the direction of change; that is, the discordant state appears to be associated with a relatively large number of different procedures whether children enter that state as part of a path of progression or as part of a path of regression.

Does Gesturing Facilitate Transition?

We have shown that, as they learn to solve mathematical equivalence problems, children who gesture in their problem explanations pass through a period during which they produce a substantial number of gesture–speech mismatches. However, recall that in screening subjects for the study, we found a number of children who produced no or very few gestures during the pretest and training session. Were these Non-Gesturers any less likely to acquire mathematical equivalence than the children who did produce gestures? A priori, it seems naive to argue that a child must gesture in order to learn. However, it is possible that the act of gesturing could make cognitive change more likely. In other words, a child who produces gesture in explaining a task might be more likely to succeed on that task than a child who does not gesture. We turn to data from the Non-Gesturers to address this question.

We examined the post-training performance of the Non-Gesturers and Gesturers, and calculated the number of children in each group who were successful on the addition and multiplication problems on the posttest. Focusing first on the children who did *not* receive instruction, we found, not surprisingly, that none of the Gesturers ($N = 20$) and none of the Non-Gesturers ($N = 12$) were successful on the posttest problems. We turn next to the two groups who did receive instruction. Figure 6 presents the proportion of Gesturers ($N = 43$) and Non-Gesturers ($N = 15$) receiving instruction who succeeded on *only* the addition problems and the proportion who succeeded on *both* the addition and multiplication problems. We found that many of the Non-Gesturers were successful on the addition problems on the posttest. However, the Non-Gesturers were somewhat less likely than the Gesturers to succeed on *both* the addition and the multiplication problems on the posttest, although the difference between groups did not reach significance ($\chi^2 = 2.7$, $df = 1$, $p \leq .10$). Moreover, this difference between the groups was maintained through the follow-up test and, in fact, approached significance when data from both the posttest and the follow-up test were considered ($\chi^2 = 3.5$, $df = 1$, $p \leq .06$). Thus, the difference between Gesturers and Non-Gesturers was most robust when the children were required to generalize their under-

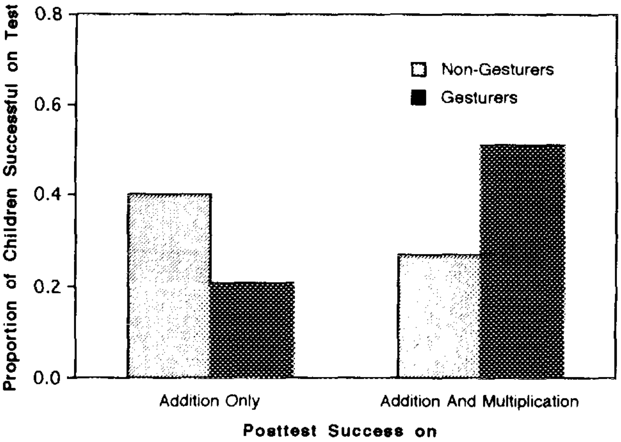

FIG. 6. The proportion of Gesturers and Non-Gesturers who were successful on the paper and pencil posttest. The figure presents the proportion of children who were successful *only* on the addition problems on the posttest, and the proportion of children who were successful on *both* the addition and multiplication problems on the posttest.

standing of mathematical equivalence and retain that understanding over a 2-week period.

Although not conclusive, these data are suggestive with respect to the role gesture plays in cognitive change. It is clear that a child can learn even if that child does not gesture. Nevertheless, the children who gestured during the task performed better on the posttest than the children who did not gesture. It is possible that the act of gesturing itself helped the children to think about the task. However, it is also possible that gesture does not play a causal role, but instead reflects a permanent or transient characteristic of the child. For example, the tendency to gesture might be an attribute which is correlated with other attributes that make a child likely to succeed on the posttest. Or, gesturing might reflect the child's temporary involvement in the task. Indeed, we found that a few of the children who did not gesture on the pretest began to gesture when they were given instruction, that is, when their incorrect understanding of equivalence was challenged and when they became actively engaged in thinking about the task (cf. the section on Subjects). Whatever the cause, learning appears to be enhanced when a child gestures during the period of instruction, particularly when progress is assessed in terms of the child's ability to generalize an understanding of the task and to retain that understanding over a period of weeks.

DISCUSSION

The results of this study provide strong support for a model of learning wherein a child begins with a single incorrect hypothesis, moves through a transitional period in which he or she entertains multiple hypotheses, and finally closes on a single correct hypothesis. Furthermore, our data demonstrate that the relationship between gesture and speech can be a valuable tool for assessing the number and kinds of hypotheses which a child entertains. *Match* between the hypothesis expressed in speech and the hypothesis expressed in gesture provides evidence for the activation of a single hypothesis, while *mismatch* between the hypothesis expressed in speech and the hypothesis expressed in gesture provides evidence for the activation of multiple hypotheses. Our results suggest that the concurrent activation of multiple hypotheses is a feature of the internal conceptual organization which characterizes the transitional knowledge state. Note, however, that children (and adults as well) can hold contradictory hypotheses over long periods of time without necessarily having transitional knowledge. The critical feature of transitional knowledge is that multiple hypotheses are considered on the same problem at the same time. It is this temporal concurrence which we take to be the distinguishing characteristic of the transitional state.

The Mechanism of Gesture–Speech Mismatch

We have 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 does not merely reflect, but rather is caused by the processes that characterize transitional knowledge. Specifically, we hypothesize that it is the concurrent activation of multiple hypotheses that not only characterizes the transitional knowledge state but actually creates gesture–speech mismatch. How then might the mismatch between gesture and speech be generated?

We found in the present study that children had approximately the same number of procedures in their repertoires in *both gesture and speech* when in a concordant state and when in a discordant state. However, because concordant children had fewer *total procedures* than the discordant children, the number of procedures in both gesture and speech accounted for a greater proportion of the total procedures in a concordant child's repertoire than in a discordant child's repertoire. In other words, children in a concordant state had proportionately more procedures found in both gesture and speech (i.e., proportionately more procedures that can, at least in principle, result in a match between the gestural and spoken modalities) than did children in a discordant state. We thus infer that children in a concordant state had proportionately more *representations* that were accessible to both modalities than did children in a discordant state.

These findings suggest a mechanism by which gesture–speech mismatch might be generated. 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 which 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 which is accessible to gesture but *not* to speech, the child will be able to describe the 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 which *is* accessible to speech) to express in speech and therefore will produce a gesture–speech mismatch.

The model we have proposed assumes that the child samples representations which are then encoded into gesture and/or speech. 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. Thus, according to this model, a child will produce a gesture–speech match whenever the child initially selects a representation which is accessible to both gesture and speech. In this sense, gesture and speech can be said to form an *integrated* system (cf. McNeill, 1992). 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. This alternative model predicts that, when asked to explain a problem, a child samples a representation accessible to gesture and *independently* samples a representation accessible to speech. According to this second 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. Goldin-Meadow, Alibali, and Church (1993) assessed how well each of these models predicted the frequency of gesture-speech matches and mismatches produced by a group of children. They found that the model we have suggested here (which assumes that gesture and speech form an integrated system) fits the data significantly better than the alternative model (which assumes independence between gesture and speech).

Implications for Mechanisms of Developmental Change

The data described in this study suggest that, at least for certain types of concepts, a child who is on the verge of learning is likely to have a relatively large number of representations which are accessible to gesture and not to speech. Thus, we propose the following description of the steps a learner follows in acquiring such concepts. 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 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, and returns once again to a state in which most of his or her representations are accessible to both gesture and speech.

One obvious implication of this view of the learning process is that when a learner spontaneously (i.e., without instruction) acquires correct representations, procedures based on those representations are likely to be found in the learner's gestural repertoire. Indeed, we found that, prior to instruction, 25 of the 43 Gesturers in this study who received instruction produced some correct procedures and *all* of these were produced in gesture (20 produced correct procedures only in gesture and 5 produced correct procedures in both gesture and speech). None of the children produced correct procedures only in speech. Thus, as expected, the children who produced correct procedures did so in gesture, and they produced correct procedures in speech only when they also produced them in gesture.

Another implication of this model is that children who produce correct procedures in gesture prior to instruction ought to be more likely to learn to solve the problems correctly than children who produce no correct procedures in gesture. In fact, we found that 65% (13) of the 20 children who produced correct procedures in gesture (and not speech) prior to instruction were successful on both the addition and multiplication problems on the posttest after instruction, compared to 22% (4) of the 18 children who produced no correct procedures on the pretest ($\chi^2 = 7.0$, $df = 1$, $p < .01$). Thus, children who had a correct strategy which was accessible to gesture (and not speech) on the pretest were more likely to succeed on the posttest than children who did not yet have a correct strategy in their repertoire on the pretest. In this way, the first step a learner takes in acquiring a concept appears to be reflected in gesture.

Note that the characterization of the transitional state as one in which multiple hypotheses are concurrently 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 most instances, into the learner's gestural repertoire rather than his or her spoken repertoire), thereby creating a transitional state characterized by the activation of 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 knowledge encoded in gesture into the more conventionalized code characteristic of speech (see below for further discussion of this point); this process thus results in the expression of a single, correct hypothesis encoded in both modalities, or perhaps even in a set of interrelated hypotheses, all of which are encoded in both modalities.[4]

Our data suggest that the same type of instruction can foster both of these two processes of change. The particular process which occurs in response to instruction depends, at least to some extent, on the child's state when the instruction is given. When provided with instruction, many children in our study progressed from a concordant state that was incorrect to a discordant state, thus entering a state characterized by multiple hypotheses. However, when provided with precisely the same instruction, an even larger number of children progressed from a discordant state to a concordant state that was correct, thus leaving a state characterized by multiple hypotheses and entering a state characterized by the activation of a single, correct hypothesis. Finally, and again with the same input, a small number of children experienced both processes, and moved from a concordant state that was incorrect, through a discordant state, to a concordant state that was correct. Thus, as Gelman (1991) points out, instruction fosters the type of change that a child is ready to experience and, as a result, is likely to have differential effects on children who enter the task in different states. The way in which a child interprets input on a particular problem depends, to a large extent, on the conception of that problem that the child brings to the task (cf. Gelman, 1991).

How then does instruction work to effect change? In our study, instruction appeared to serve three different functions. First, instruction provided the children with an explicit procedure for solving the problem. By the end of training, all of the 26 children who received instruction and

---

[4] Indeed, as Siegler and Shrager (1984) point out, for certain types of problems, it may be desirable to have a variety of hypotheses or strategies available to solve the problem. For such problems, activating multiple hypotheses on a single problem is a final state rather than a transitional state. If so, we would expect that, when in a stable state with respect to a problem of this sort, a child would have a systematic process for deploying the set of strategies. In fact, Siegler and Shrager show that young children do have a variety of strategies available for negotiating simple addition tasks (e.g., $3 + 4 =$ __), and that they proceed from one strategy to the next only if the solution attained via the first strategy does not reach the confidence criterion for that problem, i.e., there is a systematic process by which the child selects among the strategies. Thus, even though the child has several strategies in his or her repertoire, those strategies appear to be unified within one framework and, in this sense, they function as a single system.

progressed along the hypothesized path on the addition problems produced the Equalizer procedure (which was the procedure the experimenter used in her training) in speech. For those children who had produced the Equalizer procedure in gesture (and not speech) before training ($N = 6$), the instruction appeared to be helping the child make explicit in speech a procedure which, prior to instruction, the child displayed only in gesture. For those children who did not produce the Equalizer procedure at all before training ($N = 17$), the instruction provided the child with a completely new, and explicitly articulated, correct procedure (the remaining 3 of the 26 children produced the Equalizer procedure in speech on the pretest). In this regard, it is important to point out that the children who acquired the Equalizer procedure in speech by the end of training were not merely parroting an undigested and ill-understood procedure: 22 of the 26 children who progressed along the hypothesized path on the addition problems produced the Equalizer procedure in gesture as well as speech, and 19 of the children produced at least one correct procedure (in addition to Equalizer) that had not been presented by the experimenter. In fact, only one child produced Equalizer in speech and not in gesture, and produced no other correct procedures; this then was the only child in the sample of progressors who might have been parroting the procedure modeled by the experimenter.

Second, instruction encouraged the children to generate their own correct procedures that had *not* been modeled by the experimenter. As noted above, 19 of the 26 children who progressed along the path generated a new, correct procedure (Add–Subtract, Grouping, or both). Note that the first two functions of instruction help to increase the number of procedures children have in their repertoires, thus facilitating transition from a single, incorrect hypothesis to a state of multiple hypotheses.

Finally, instruction appeared to help children weed out and discard incorrect procedures. By the end of instruction, 25 of the 26 children who progressed along the path had discarded one or more incorrect procedures from their repertoires. Consequently, the third function of instruction helps to decrease the number of procedures children have in their repertoires, thus facilitating transition from a state of multiple hypotheses to a state of a single, correct hypothesis.

It is important to point out that, although few, some children in our study did make progress in acquiring mathematical equivalence even without instruction (two moved from a concordant incorrect state to a discordant state, and one moved from a discordant state to a concordant correct state). For these children, solving the training problems was itself sufficient input to foster change. Thus, although the instruction we provided was effective in fostering both types of transitional processes, for at least some children, it was not necessary to foster either process.

The 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 to the set of representations 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?

Huttenlocher (1973, 1976) argued convincingly that it is not useful, nor 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 argues that there must be representational systems not involving words which humans use to encode information. Similarly, Anderson (1983) and Johnson-Laird (1983) have each proposed a variety of representational systems (including minor ones which involve imagery) as options for encoding information. We suggest that gesture can be a vehicle for one of these options. Gesture offers children (and adults, for that matter) a vehicle that is distinctly different from speech for expressing their understanding of a problem.

Indeed, McNeill (1992) has argued that, even though gesture and speech are two aspects of a single process and are correlated in meaning, the two do not always reveal the same meaning. According to McNeill, gesture reflects a global-synthetic image. It is idiosyncratic and constructed at the moment of speaking—it does not belong to a conventional code. In contrast, speech reflects a linear-segmented, hierarchical linguistic structure, utilizing a grammatical pattern that embodies the language's standards of form and drawing on an agreed-upon lexicon of words.[5] 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. Thus, certain types of information may be more easily encoded in gesture than in the speech itself.

---

[5] Note that McNeill's (1992) characterization does not imply that the manual modality cannot, under other circumstances, assume a language-like 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, McNeill & Goldin-Meadow, 1993), these gestures take on aspects of language-like structure and thus no longer resemble the spontaneous gestures that accompany speech (see Goldin-Meadow, 1993, for further discussion).

Our data suggest that, for certain problems and at certain times in the learning process, gesture 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 herself may be incapable of verbally expressing her understanding of the problem at a moment when she *is* capable of expressing that understanding in gesture. Why might this be so? Hadamard (1945) has argued that mathematical thinking, particularly innovative mathematical thinking, is not conducted in words but rather in spatial images—images which 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). Gesture may indeed provide a more accessible code for a child's mathematical knowledge than does speech.[6] For example, in acquiring 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 explicit understanding of the individual parts and the relationships among them. On this view, the translation of an image into speech is not merely a matter of learning the appropriate verbal labels, but rather one 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. According to Karmiloff-Smith, at early points in the development of a concept, knowledge of the concept may be dis-

---

[6] Although we have found that correct explanations appear in gesture before speech in children acquiring mathematical equivalence, our results leave open the possibility that speech could anticipate gesture in other (perhaps less spatial) domains. For example, Goodman, Church, and Schonert (1991) have found that gesture and speech do not always match in children's and adults' responses to Kohlberg's moral reasoning tasks. 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.

played in a very limited range of tasks precisely because the code in which that knowledge is represented limits its accessibility (e.g., the developmental gap between succeeding in action and being capable of explaining that success, cf. Karmiloff-Smith & Inhelder, 1974/75). The process of development thus consists of redescribing knowledge into a more explicit and flexible code, one that extends its accessibility and allows knowledge of the concept to be displayed in a variety of tasks. We suggest that, when knowledge is represented uniquely in a gestural code, that knowledge is implicit and unconscious and therefore can be deployed in only a narrow range of tasks. Nevertheless, when a child does display knowledge of a concept in gesture, this display is a signal that the child has taken the first steps in acquiring the concept and has entered a transitional state. At that moment, the child has knowledge accessible to gesture that is not accessible to speech; thus, when called upon to explain the concept, the child will, at times, produce gesture–speech mismatches. It is at this point in development that the child is ready to begin the process of redescription, a process that will eventually result in the translation of knowledge encoded uniquely in gesture into speech.

We suggest that the mismatch between gesture and speech signals the beginning of the redescription process. We might also ask whether gesture-speech mismatch plays a role in propelling this process of redescription. It is possible that children who produce gesture–speech mismatches are aware of the fact that the information encoded in their gestures does not match the information encoded in their speech. This awareness might then act as an impetus for redescribing the information encoded in gesture into speech. However, the fact that children (and adults, for that matter) in general appear to be unaware that they are gesturing suggests that children are not likely to be conscious of their mismatches. Thus, awareness of a contradiction between the information conveyed in gesture and the information conveyed in speech is not likely to be the force which propels development forward. However, Karmiloff-Smith (1986) has suggested that there is a general push toward redescribing information so that it is accessible to more flexible and explicit codes. It is this push toward redescribing information into speech—the mode that is recognized by the community as a legitimate communicative channel—that may, in fact, provide the impetus for change in the acquisition of at least certain, verbalizable concepts.

The Function of Gesture in Transition

We have suggested that children are not likely to be aware of their gestures. As a result, any effect that gesturing might have on cognitive change is not likely to be mediated by conscious processes. However, it is possible that gesturing could have effects on cognitive change that do *not* require awareness on the part of the learner. If so, gesture might be playing an active role in bringing about cognitive change, and thus might be causing change as well as reflecting it.

The effect of gesture on the learner. McNeill (1992) has suggested that the act of gesturing can affect thought. According to McNeill, some dimensions of thought are presented in gesture and others in linguistic form. There is a synthesis at the moment of speaking when language and gesture are unconsciously combined into one unified presentation of meaning. This is an act of communication, but 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 gesturing might itself facilitate cognitive change (although it is obvious that gesturing cannot be requisite for cognitive change since some children who do not gesture at all still progress). In our study, we found suggestive evidence that children who did not gesture at all on our task performed less well on post-training assessments than children who did gesture on the task. These data are consistent with the hypothesis that gesturing facilitates the learning process (whether knowingly or unknowingly).

The effect of gesture on the learning environment. It is also possible that gesture 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 partners to adjust their 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 (that is, the child's zone of proximal development, Vygotsky, 1978). The results of this study suggest that trained experimenters are able to interpret the match or mismatch between gesture and speech in a child's explanations, and use that information to make inferences about the child's knowledge at varying points in the acquisition process. In addition, Goldin-Meadow, Wein and Chang (1992) have shown that adults who have *not* been trained to code gesture can detect and interpret the match or mismatch between gesture and speech in a child's explanations, and use that information in their assessments of the child's knowledge of the task (see also a study conducted by McNeill, Cassell, McCullough, & Tuite, reported in McNeill, 1992). Thus, children's production (or lack of production) of gesture-speech mismatches may provide information to those who interact with them, thereby providing children with a mechanism through which they can help shape their own learning environments.

In general, McNeill (1992) argues 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. McNeill further argues 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 was found to influence 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, 1993; Siegler & Jenkins, 1989). Thus, it is quite possible that adults do notice—and process—the gestures children produce, particularly at the time of transition.

Our data suggest that gesture provides a window that can be used—by both experimenter and naive observer—to assess whether a child is in a transitional state. However, it is worth noting the rather obvious point that a child's transitional status is detectable through the window provided by gesture only if the child gestures. Recall that there were 27 children in our study who gestured very little or not at all, and who therefore provided no opportunity to observe patterns of change in the relationship between gesture and speech. As long as the window offered by gesture is there and accessible to the observer, it can be used to assess whether a child is in a transitional state. The caveat, however, is that the window may not always be open.

In summary, we have shown that the relationship between gesture and speech is a rich source of information about the process of knowledge change. We have argued that the mismatch between gesture and speech is a more sensitive and more informative index of the transitional knowledge state than other indices which rely solely on speech. 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 the specific hypotheses which children entertain as they move into and out of the transitional knowledge state. Thus, the mismatch between gesture and speech allows us to probe the fleeting state that occurs between the endpoints of learning. As such, it is a valuable tool for exploring, not merely the presence of cognitive change, but also the process.

APPENDIX A

Possible Paths from the Pretest through the Addition Training Problems and the Number of Children Who Took Each Path (CI, Concordant Incorrect; D, Discordant; CC, Concordant Correct)

State on pretest	State on 1st six training problems	State on 2nd six training problems	Proportion of children taking path	Type of path taken
Children who received no instruction on the training problems (*N* = 20)				
CI (N = 10)	CI (*N* = 9)	CI (*N* = 7)	.35	Stay
		D (*N* = 2)	.10	Progress
		CC (*N* = 0)	.00	Progress/violate
	D (*N* = 0)	CI (*N* = 0)	.00	Regress
		D (*N* = 0)	.00	Progress
		CC (*N* = 0)	.00	Progress
	CC (*N* = 1)	CI (*N* = 0)	.00	Regress
		D (*N* = 0)	.00	Regress
		CC (*N* = 1)	.05	Progress/violate
D (*N* = 10)	CI (*N* = 2)	CI (*N* = 2)	.10	Regress
		D (*N* = 0)	.00	Regress
		CC (*N* = 0)	.00	Regress
	D (*N* = 8)	CI (*N* = 4)	.20	Regress
		D (*N* = 3)	.15	Stay
		CC (*N* = 1)	.05	Progress
	CC (*N* = 0)	CI (*N* = 0)	.00	Regress
		D (*N* = 0)	.00	Regress
		CC (*N* = 0)	.00	Progress
Children who received instruction on the training problems (*N* = 43)				
CI (*N* = 18)	CI (*N* = 2)	CI (*N* = 1)	.02	Stay
		D (*N* = 1)	.02	Progress
		CC (*N* = 0)	.00	Progress/violate
	D (*N* = 11)	CI (*N* = 0)	.00	Regress
		D (*N* = 8)	.19	Progress
		CC (*N* = 3)	.07	Progress
	CC (*N* = 5)	CI (*N* = 0)	.00	Regress
		D (*N* = 0)	.00	Regress
		CC (*N* = 5)	.12	Progress/violate

Appendix A—*Continued*

State on pretest	State on 1st six training problems	State on 2nd six training problems	Proportion of children taking path	Type of path taken
D (*N* = 25)	CI (*N* = 2)	CI (*N* = 1)	.02	Regress
		D (*N* = 0)	.00	Regress
		CC (*N* = 1)	.02	Regress
	D (*N* = 15)	CI (*N* = 1)	.02	Regress
		D (*N* = 7)	.16	Stay
		CC (*N* = 7)	.16	Progress
	CC (*N* = 8)	CI (*N* = 0)	.00	Regress
		D (*N* = 1)	.02	Regress
		CC (*N* = 7)	.16	Progress

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). *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.

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

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(2), 919–939.

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

Bresnahan, J. L., & Shapiro, M. M. (1966). A general equation and technique for the exact partitioning of chi-square contingency tables. *Psychological Bulletin*, 66(4), 252–262.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, **4**, 55–81.

Church, R. B. (1993). Using the relationship between gesture and speech to capture transitions in learning. Under review.

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

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

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477–531.

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

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 293–322). Hillsdale, NJ: Erlbaum.

Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology*, 4, 631–666.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Earlbaum.

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). New York: Cambridge University Press.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to reach the mind. *Psychological Review*, 100(2), 279–297.

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Non-effect of parental input on language development. *Science*, 221(4608), 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), no. 207.

Goldin-Meadow, S., & Mylander, C. (1990). Beyond the input given: The child's role in the acquisition of language. *Language*, 66(2), 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), 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) *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. (1993). Indexing transitional knowledge. *Developmental Psychology*, in press.

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

Huttenlocher, J. (1973). Language and thought. In G. A. Miller (Ed.), *Communication, language, and meaning: Psychological Perspectives* (pp. 172–184). New York, NY: 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., Ackerman-Vallado, E., Blanchet, A., Karmiloff-Smith, A., Kilcher-Hagedorn, H., Montagero, J., & Robert, M. (1976). Des structures cognitives aux procedures de decouverte. *Archives de Psychologie XLIV*, 171, 57–72.

Inhelder, B., & Sinclair, H. (1969). Learning cognitive structures. In P. Mussen, J. Langer, and 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. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, 23(2), 95–147.

Karmiloff-Smith, A., & Inhelder, B. (1974/75). "If you want to get ahead, get a theory." *Cognition*, 3(3), 195–212.

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.

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

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

Miller, S. A., Brownell, C. A., & Zukier, H. (1977). Cognitive certainty in children: Effects of concept, developmental level, and method of assessment. *Developmental Psychology*, 13(3), 236–245.

Murray, J. P. (1974). Social learning and cognitive development: Modelling effects on children's understanding of conservation. *British Journal of Psychology*, 65(1), 151–160.

Perry, M. (1985). *Children's procedures in interpreting symbolic equivalence*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Toronto, Canada.

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. (1964/1967). *Six psychological studies*. New York: Random House.

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

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). New York, NY: Academic Press.

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 (Whole No. 189).

Siegler, R. S. (1983). Five generalizations about cognitive development. *American Psychologist*, 38, 263–277.

Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, 46(6), 606–620.

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

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–294). Hillsdale, NJ: Erlbaum.

Singleton, J., Goldin-Meadow, S., & McNeill, D. (1993). The cataclysmic break between gesticulation and sign. To appear in K. Emmorey & J. Reilly (Eds.), *Sign, gesture, and space*. Hillsdale, NJ: Erlbaum.

Strauss, S. (1972). Inducing cognitive development and learning: A review of short-term training experiments. I. The organismic developmental approach. *Cognition*, 1(4), 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**(4), 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.

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

(Accepted November 16, 1992)