ELSEVIER

Cognitive Science 26 (2002) 817–831

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Gesture offers insight into problem-solving in adults and children

Philip Garber*, Susan Goldin-Meadow

Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637, USA

Received 19 September 2001; received in revised form 3 April 2002; accepted 20 June 2002

Abstract

When asked to explain their solutions to a problem, both adults and children gesture as they talk. These gestures at times convey information that is not conveyed in speech and thus reveal thoughts that are distinct from those revealed in speech. In this study, we use the classic Tower of Hanoi puzzle to validate the claim that gesture and speech taken together can reflect the activation of two cognitive strategies within a single response. The Tower of Hanoi is a well-studied puzzle, known to be most efficiently solved by activating subroutines at theoretically defined choice points. When asked to explain how they solved the Tower of Hanoi puzzle, both adults and children produced significantly more gesture–speech mismatches—explanations in which speech conveyed one path and gesture another—at these theoretically defined choice points than they produced at non-choice points. Even when the participants did not solve the problem efficiently, gesture could be used to indicate where the participants were deciding between alternative paths. Gesture can, thus, serve as a useful adjunct to speech when attempting to discover cognitive processes in problem-solving.

Keywords: Psychology; Cognitive development; Communication; Problem-solving; Developmental experimentation

1. Introduction

When people talk, they gesture. Gesture is found in all cultures that have been observed, and it occurs across a wide range of tasks and ages (Feyereisen & de Lannoy, 1991; McNeill,

* Corresponding author. Present address: College of Education, National-Louis University, 2840 North Sheridan Road, Evanston, IL 60201, USA. Tel.: +1-847-475-1100; fax: +1-847-256-1057.

E-mail addresses: pgarber@nl.edu (P. Garber), sgm@uchicago.edu (S. Goldin-Meadow).

PII: S0364-0213(02)00087-3

1992). Gesture might be nothing more than handwaving, reflecting an outpouring of excess energy or a bid for the listener's attention. However, recent research has shown that the gestures speakers spontaneously produce when they talk can reflect substantive ideas relevant to the task at hand (Goldin-Meadow, in press; Kendon, 1980; McNeill, 1992). For example, consider a child shown two tall thin glasses containing the same amount of water. The water in one glass is poured into a short wide dish and the child is asked whether the dish has the same amount of water as the full glass. The child says "no" and justifies his incorrect belief by saying "it's different because this one's tall and that one's short" while at the same time gesturing TALL (a flat palm held at the water level of the full glass) and then SHORT (a flat palm held at the water level of the dish). This child is focusing on the height of the water and has indicated his focus in gesture as well as speech.

There are times, however, when speakers use their hands to convey information that is *not* found in their speech. Consider another child who gives the same response to the water conservation question in speech—"it's different because this one's tall and this one's short"—but in gesture produces SKINNY (two flat palms held at the sides of the glass) and then WIDE (two flat palms held at the sides of the dish). This child highlights height in her speech but width in her gesture—she has produced a gesture–speech mismatch (Church & Goldin-Meadow, 1986).

Do mismatching gestures of this sort have cognitive significance? Two types of evidence suggest that they do, one from learning paradigms, the other from problem-solving paradigms. (1) Gesture–speech mismatch can index cognitive stability and thus predict learning outcomes. Children who produce many mismatches when they explain their answers to a set of conservation or math problems are in an unstable cognitive state with respect to those problems—if provided with appropriate instruction, they will improve on the task (Church & Goldin-Meadow, 1986; Perry, Church, & Goldin-Meadow, 1988); if not, they will fall back to a less advanced, but more stable state (Alibali & Goldin-Meadow, 1993). These children are ready to learn, more ready than children who produce few mismatches on the same task. (2) Gesture–speech mismatch can predict future problem-solving strategies. When asked to restate a continuous change problem, adults will often convey information in their gestures that is not found in their speech. When later asked how they would solve this problem, these adults not only call upon strategies compatible with their spoken description of the problem, but also strategies compatible with their gestured description—and they call upon gestured strategies equally as often as spoken strategies (Alibali, Bassok, Olseth, Syc, & Goldin-Meadow, 1999). Gesture, thus, provides insight into a speaker's thoughts. But what is gesture telling us about those thoughts?

We have speculated that gesture in a mismatch reflects a second problem-solving strategy, one that the speaker has not integrated with the problem-solving strategy expressed in speech (Goldin-Meadow, Alibali, & Church, 1993a). Gesture–speech mismatch, thus, reflects the activation of two strategies on a single problem. We have gathered indirect evidence for this hypothesis using a cognitive load paradigm (Goldin-Meadow, Nusbaum, Garber, & Church, 1993b). We observed 9–10-year-old children, all of whom solved a mathematical equivalence problem incorrectly. We divided the children into those who produced many mismatches and those who produced few when explaining their solutions to the math task. We then asked all of the children to solve the same types of math problems a second time, but in this condition they did not have to explain their solutions—instead, they had to remember a list of unrelated words while solving each problem. The number of words the children recalled served as a gauge of how much cognitive effort they expended while solving the math problems. We found that children who had produced many mismatches on the earlier explanation task remembered significantly fewer words compared to children who had produced few mismatches. These observations suggest that the mismatchers were activating two strategies when solving the math problems, and thus, had less cognitive effort left over to remember words than the matchers who were activating only one strategy. In other words, children who produced two strategies (one in speech, one in gesture) on a single problem when *explaining* the task were just the ones who activated more than one strategy when *solving* the task. Thus, the study not only suggests that mismatchers activate more than one strategy when either solving or explaining a task, but it also validates the use of gesture in explanations as an index of problem-solving strategies.

Our goal in the present work is to extend this phenomenon to a task where previous theory provides clear-cut predictions about when the problem-solver ought to be activating two strategies. If gesture can be validated as a technique for identifying when problem-solvers are entertaining two strategies, it can supplement verbal protocols as a tool for tapping cognitive processes in problem-solving.

Gesture is pervasive and comes for "free" whenever adults and children are asked to give explanations. Gestural data are consequently there for the taking. Moreover, since much of what problem-solvers think about when they solve problems may not be expressed verbally, verbalizations do not always accurately reflect the cognitive processes at work during problem-solving (Ericsson & Simon, 1980, 1993). Gesture, it turns out, has access to speakers' implicit thoughts—thoughts that speakers cannot articulate and may not even know they have (Garber, Alibali, & Goldin-Meadow, 1998). Thus, gesture could serve as a particularly useful technique to assess problem-solving as it unfolds, one that offers a unique perspective on the process.

The Tower of Hanoi is a puzzle in which a graduated tower of disks, with the largest on the bottom and the smallest at the top, must be moved from a source peg to a goal peg according to two rules (Egan & Greeno, 1973). The first rule stipulates that only one disk can be moved at a time; the second stipulates that a larger disk cannot be placed on top of a smaller one. The problem is most efficiently solved by making repeated comparisons of the disks' current state to the final goal state and to several desired intermediate goal states (Newell & Simon, 1972). Many studies have demonstrated that successful problem-solvers do break the Tower of Hanoi puzzle into theoretically defined subroutines. Evidence for the existence of subroutines comes from a variety of sources: *errors* made by adults (Egan & Greeno, 1973) and children (Bidell & Fischer, 1995; Byrnes & Spitz, 1979); *temporal patterning of moves* in adults (Karat, 1982; Kotovsky, Hayes, & Simon, 1985) and children (Bidell & Fischer, 1995); *verbal protocols* provided by adults (Anzai & Simon, 1979; Hayes & Simon, 1977) and children (Klahr & Robinson, 1981); and *computer simulations* of human problem-solving (Ernst & Newell, 1969; Newell & Simon, 1972).

We suggest that *gesture* might also provide evidence that problem-solvers use subroutines when solving the Tower of Hanoi puzzle. At moments when subroutines are activated, the problem-solver must choose between at least two possible paths—one which allows the puzzle to be solved optimally (in the least number of moves), and the other which does not. Once the subroutine is activated, the next several steps on the path are clear, and there are no choices to make. We have claimed that gesture–speech mismatches have the potential to instantiate two strategies in a single response (one expressed in speech, and one in gesture). We, therefore, speculated that problem-solvers, both adults and children, might produce mismatches at just those branch points in the Tower of Hanoi puzzle when theoretically called-for subroutines should be chosen—but not produce mismatches in the several steps after the subroutine has been activated.

2. Method

2.1. Participants

2.1.1. Adults

Twenty-four college students ($M = 20$ years; range = 18–24 years), 12 males and 12 females, were chosen randomly from a larger database of volunteers. None reported being familiar with the Tower of Hanoi puzzle or any of its isomorphs (Hayes & Simon, 1977). All were paid for their participation.

2.1.2. Children

Thirty children were screened for participation in the study; 5 who were not successful during the practice trial, and 1 who was familiar with a computerized version of the puzzle, were eliminated from the study. The final sample consisted of 11 boys and 13 girls, ages 8;5 (years;months) to 10;1 ($M = 9$;8), from three parochial schools in Chicago reflecting a variety of socio-economically and ethnically diverse backgrounds.[1] Children were given a pencil in exchange for participating in the study.

Adults were asked to solve and explain a 4-disk version of the Tower of Hanoi, and children a 3-disk version. The 4-disk version can be solved by adults but presents a challenge (Ewert & Lambert, 1932; Gagne & Smith, 1962), as does the 3-disk version for 9-year-old children (Byrnes & Spitz, 1979).

2.2. Materials and task

2.2.1. Apparatus

The Tower of Hanoi apparatus consisted of a flat wooden base, 18 in. × 6 in., and three 6 in. vertical wooden pegs (referred to as pegs 1, 2, and 3), drilled into the base at 3 in. intervals. Disks were made of wood covered with Formica coatings of bright, easily identifiable colors (blue, red, yellow, green). The smallest disk measured 2 in. in diameter, with each disk increasing in diameter by 1/2 in.

2.2.2. Rules

The top of Fig. 1 displays the start configuration of the 4-disk task. The goal was to stack all four disks, in the same configuration, on peg 3 following two rules. The first stipulated that only one disk could be moved at a time. The second stipulated that a larger disk could not be

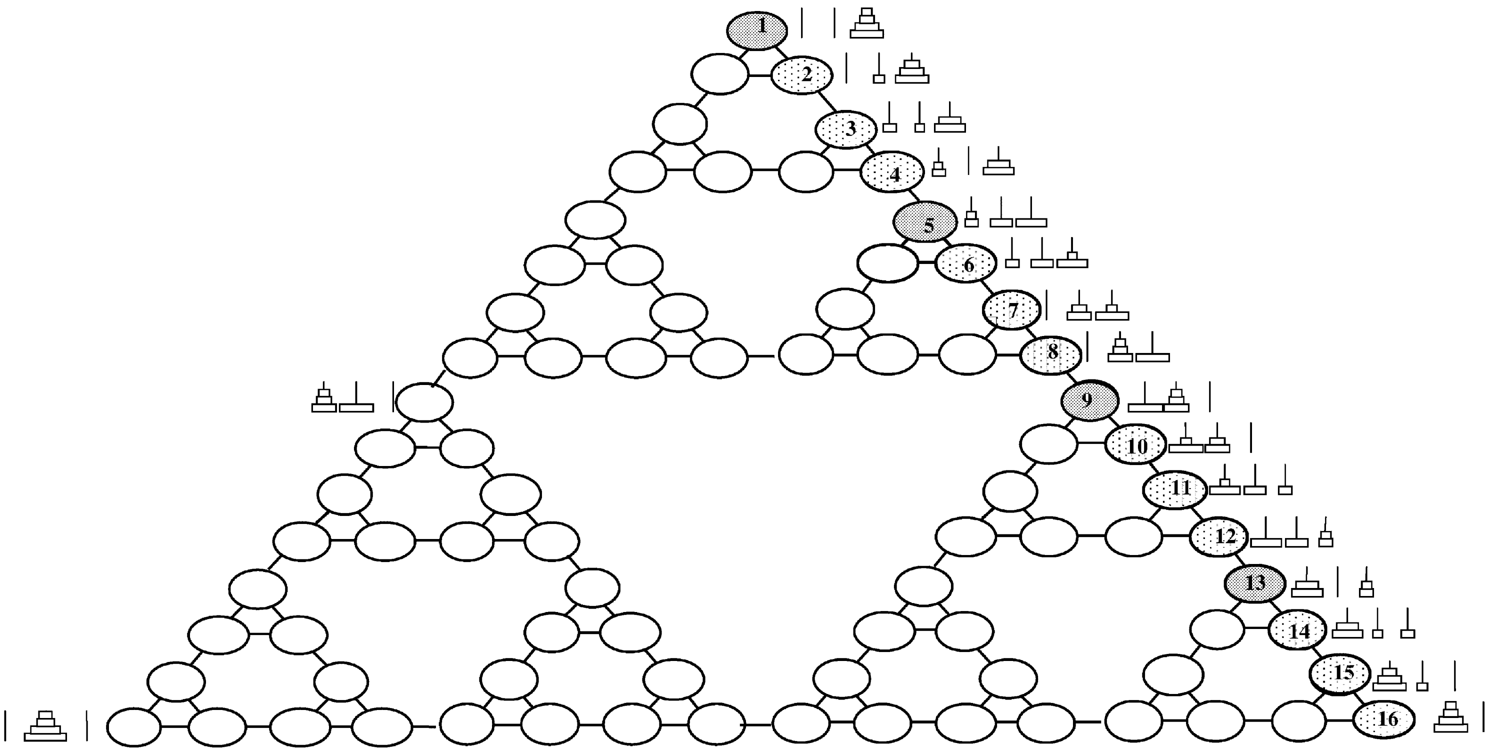

Fig. 1. A problem-space diagram of the 4-disk tower of Hanoi task. The circles indicate all possible configurations (configuration 1 is the starting point; configuration 16 is the goal). The numbered circles indicate the steps in the optimal solution to the task; the configuration of disks at each step is shown beside the circles. Gray circles indicate choice points in the optimal path; dotted circles indicate non-choice points. Open circles indicate moves that are not optimal.

placed on top of a smaller disk. Both rules were reiterated several times throughout the study and whenever a rule infraction was made during solving.

The minimum number of moves required to solve any version of the Tower of Hanoi task optimally is $2^n - 1$, where n = number of disks—7 moves for the 3-disk task, and 15 for the 4-disk task. Consider the sequence of moves necessary to solve the 3-disk puzzle optimally. At the outset, a child solving this task must figure out how to get the largest disk (at the bottom of the initial configuration) onto the goal peg. However, the disk cannot be moved without first moving the two disks that sit on top of it. The smallest disk must be placed on the goal peg, the middle disk then placed on the middle peg, and the smallest disk placed back on the middle disk on the middle peg—all before the largest disk can be moved to the goal peg. To reach the final goal state, again the smallest disk must first be moved elsewhere before the middle disk can be placed on the largest disk, now on the goal peg. Thus, at each juncture, there are intermediate subroutines that have to be activated before each of the three disks can be placed in its rightful position on the goal peg.

Precisely the same process must be followed to solve the 4-disk problem optimally, but there are many more subroutines that must be activated. Fig. 1 displays a diagram of all possible moves on a 4-disk problem. The series of moves down the right leg of the largest triangle represents the optimal solution to this problem. The gray circles along this leg represent the points where the participant can choose to activate a subroutine that will lead to the optimal solution. To solve the 4-disk problem optimally, the subroutine is activated on steps 1, 5, 9, and 13 in the 4-disk task (cf. Anzai & Simon, 1979). The structure of the 3-disk problem can be seen in the top three small triangles in Fig. 1; to solve the 3-disk problem optimally, the subroutine is activated on steps 1 and 5 (Klahr, 1978).

2.3. Procedure

Each participant was given a simple version of the puzzle to practice on—a 3-disk task for adults, a 2-disk task for children. Once participants had succeeded on the practice trial (after two or three attempts in most cases), they were given the same Tower of Hanoi problem to solve three times. A problem-solving trial was considered complete when the participant either reached the final goal (i.e., all disks in the proper order at peg 3), or reached an impasse and stopped solving.

At the end of each trial, the experimenter returned the stack of disks to peg 1 and asked the participant to explain his or her solution steps in detail without moving the disks.[2] The cognitive load study described earlier (Goldin-Meadow et al., 1993b) suggests that the explanations problem-solvers produce when asked to describe how they solved a problem are a good index of how they actually go about solving problems of that type. We, therefore, assumed that the explanations produced in this study reflected the steps that the participants took when solving Tower of Hanoi problems—either the steps they had taken when actually solving the problem (including their planning steps) or, more likely, the steps they took when re-solving the problem on-line during the explanation.

Experimental sessions took place in quiet rooms at the participants' university or elementary school and lasted approximately 25 min. All testing rooms contained a table, where the puzzle was placed, and two chairs for the participant and the experimenter. A videocamera, used to record data, was positioned so that the participant faced the camera head-on and the experimenter was off to the side, but still captured on videotape.

2.4. Coding

All videotaped data were transcribed for speech and gesture, and coded according to a system described in Garber (1997). Moves described in speech were transcribed first. Hand gestures were then indicated on the transcripts. A gesture was defined as any movement of one or both hands directed toward the apparatus that indicated a disk or its path as it traveled from one peg to another. Other manual behaviors, such as functional actions (touching or picking up a disk) or self-adapters (scratching the head, touching the face or body, cf. Ekman & Freisen, 1969) were not considered gestures.

2.4.1. Types of path descriptions

We divided path descriptions into optimal (those in which participants described moving the tower of disks from the starting peg to the goal peg in the minimum number of legal moves) and non-optimal (those in which participants described taking more than the minimum number of legal moves or failing to reach the goal entirely).[3] Two adults produced optimal path descriptions on all 3 problems, 15 produced non-optimal path descriptions on all 3 problems, and 7 produced some of both; as a result, 9 adults contributed to the analyses of optimal paths, and 22 contributed to the analyses of non-optimal paths. Seven children produced only optimal path descriptions, 9 produced only non-optimal path descriptions, and 8 produced both; as a result, 15 children contributed to the analyses of optimal paths and 17 contributed to the analyses of non-optimal paths.[4]

2.4.2. Identifying moves in path descriptions

Our next step was to code each move in the participant's described path. We first identified the path described in speech and then, on a separate pass through the data, identified the path described in gesture. We next compared the speech and gesture codes for each move, classifying a description as a *match* if gesture and speech described the same path, and as a *mismatch* if gesture and speech described different paths.[5] Fig. 2 presents examples of a gesture–speech match (top) and mismatch (bottom) produced by an adult describing an optimal solution to the 4-disk task.

Three types of mismatches were found in the data: (1) Gesture conveyed a different path from the one identified in speech. For example, one participant described moving the blue disk to peg 3 by saying, "I moved the blue disk to the last peg [peg 3]," while pointing to the blue disk and then to peg 2, the middle peg. (2) Gesture conveyed the path identified in speech but also an additional path. For example, a participant described moving the blue disk from peg 1 to peg 3 by saying, "I moved the blue disk to the last peg [peg 3]," while pointing to peg 1 and then to peg 3 and, at the same time, pointing with the other hand to peg 1 and then peg 2 (an additional move that was never verbalized; see also Fig. 2B). (3) Gesture described a specific path whereas speech provided a general description of a path that was difficult to identify. For example, a participant said, "I would do this back and forth while trying to keep the little ones together," while pointing at the smallest disk on peg 1 and then pointing at peg 2.

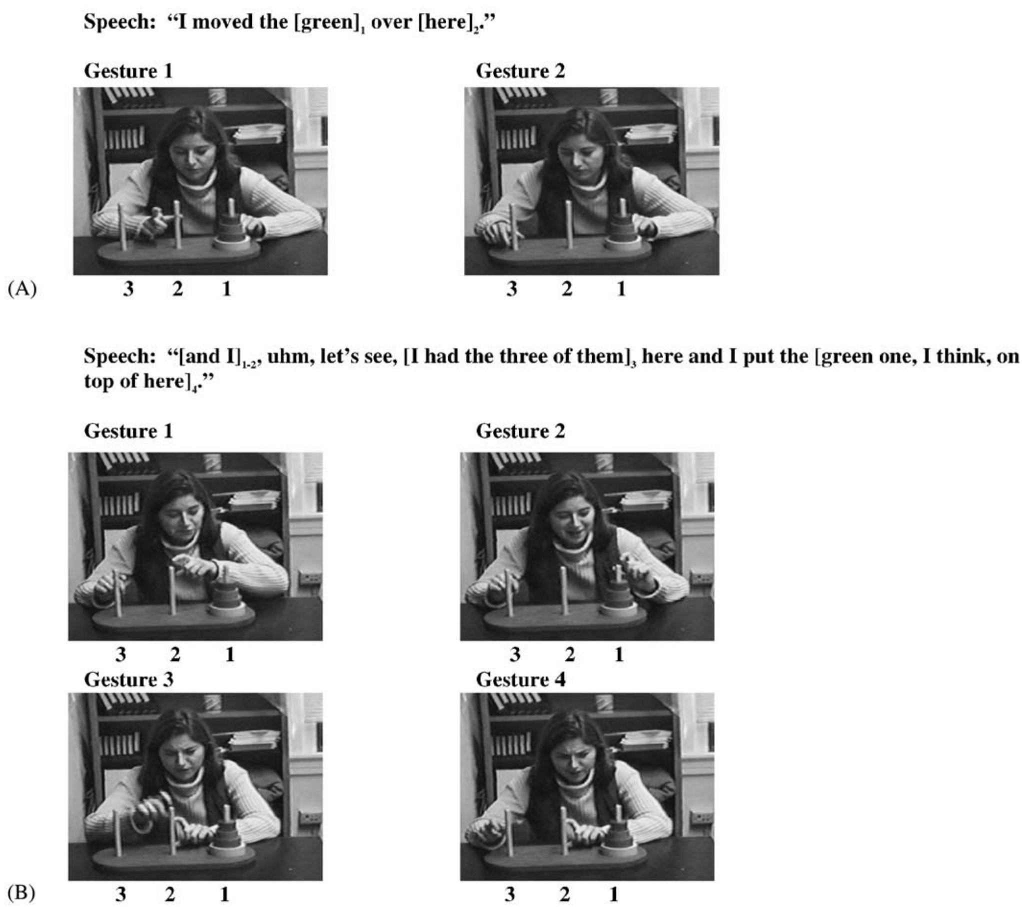

Fig. 2. Examples of a gesture–speech match and mismatch. The adult was describing an optimal solution to the 4-disk task. Numbered brackets indicate where in the speech stream each gesture was produced. (A) Gesture–speech match. The speech and gestures an adult produced when describing configuration 3 in Fig. 1 (a non-choice point). The adult moved her right hand shaped in a C in a path from peg 2 to peg 3 while describing precisely the same path in speech. Speech: “[and I]$_{1-2}$, uhm, let’s see, [I had the three of them]$_{3}$ here and I put the [green one, I think, on top of here]$_{4}$.” (B) Gesture–speech mismatch. The speech and gestures an adult produced when describing configuration 9 in Fig. 1 (a choice point). The adult first moved her left hand shaped in a C in a path from peg 2 to 1, a path which she did *not* mention in speech. She then held her left hand shaped in a vertical C at peg 2 and moved her right hand shaped in a horizontal C in a path from peg 2 to 3, a path which she did describe in speech.

Note that what is crucial in creating a mismatch is that the path identified in one modality be different from the path identified in the other. The two modalities may overlap in some of the information they convey—for example, pointing at the blue disk and saying “the blue disk” in example (1) above.[6] However, the pointing gesture in this example was part of a ‘blue disk to peg 2’ path whereas the speech was part of a ‘blue disk to peg 3’ path—the description was thus classified as a mismatch. More generally, we use speech to identify the unit of analysis that is appropriate to the task (the path in the Tower of Hanoi task; the conservation rationale in our conservation tasks, Church & Goldin-Meadow, 1986; the problem-solving strategy in our math tasks, Perry et al., 1988), and then code gesture in relation to that unit. We consider a response to be a mismatch only if the extra information conveyed in gesture identifies a unit (path, rationale, or problem-solving strategy) that is different from the unit identified in speech.

2.5. Reliability

Adults described a total of 1,416 moves; children described 688. On average, adults produced gestures in 94% (*SD* = 6%) of their described moves, and children produced gestures in 96% (*SD* = 6%) of their described moves. Reliability was established by having a second experimenter transcribe and code the spoken and gestured moves of a subset of adult and child participants (172 moves produced by 3 adults over 9 paths; 76 moves produced by 3 children over 9 paths). Inter-rater agreement was determined by calculating the proportion of agreements between coders and Cohen's kappa coefficients (Cohen, 1960). For adults, agreement between coders was 1.00 for describing moves in speech (kappa = 1.00); .97 for describing moves in gesture (.94), and .90 for describing gesture–speech matches and mismatches (.88). For children, comparable numbers were: 1.00 (1.00); 97 (.94); 90 (.88).

3. Results

3.1. Optimal path descriptions

Previous theoretical analyses of the Tower of Hanoi problem predict that participants who describe an optimal solution should activate subroutines at steps 1 and 5 in the 3-disk task (Klahr, 1978) and steps 1, 5, 9, and 13 in the 4-disk task (Anzai & Simon, 1979; see Fig. 1). We hypothesized that participants would be particularly likely to produce gesture–speech mismatches at these points precisely because they had to plan several moves ahead at the choice points and thus might be considering alternative paths simultaneously. These alternative paths could be expressed in a gesture–speech mismatch, one path in speech and another in gesture.

To test this hypothesis, we divided each participant's described moves into theoretically predetermined "choice point" moves and "non-choice point" moves. We then calculated the proportion of gesture–speech mismatches produced on each type of move. Fig. 3 presents the proportion of mismatches that adults (left bars) and children (right bars) produced when describing choice point moves (black bars) and non-choice point moves (white bars). We conducted an ANOVA with age (adult, child) as a between-subjects factor and type of move (choice point, non-choice point) as a within-subjects factor. Arcsine transformations were performed before statistical analysis. As predicted, there was a significant effect of type of move [$F(1, 22) = 10.52$, $p = .004$]—both adults and children produced more mismatches when describing moves at choice points than at non-choice points. There was no effect of age [$F(1, 22) = 2.04$, ns] and no interaction between factors [$F(1, 22) = .78$, ns].

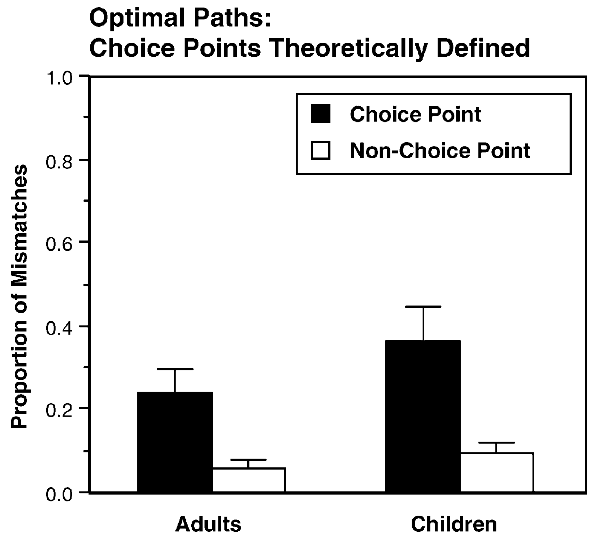

Fig. 3. Mismatches in optimal paths. The proportion of gesture–speech mismatches that adults and children produced when describing choice point moves and non-choice point moves in optimal paths. Choice points were defined theoretically as the point at which a subroutine ought to be activated to solve the problem in the least number of possible moves. Error bars indicate standard errors.

3.2. Non-optimal path descriptions

Non-optimal paths were those in which the goal was not reached in the minimal number of moves, or was not reached at all. We suspected that when participants described a non-optimal path, their path decisions may not have been made at theoretically defined choice points since they were not, in fact, following the most efficient path. Nevertheless, we divided their described moves into theoretically defined choice points and non-choice points, and calculated the proportion of mismatches at each type of move. The left graph of Fig. 4 displays the data for the non-optimal paths. As expected, neither adults nor children produced more mismatches when describing theoretically defined choice points than when describing non-choice points [$F(1, 37) = .94$, ns]. There was no effect of age [$F(1, 37) = 1.32$, ns] and no interaction between factors [$F(1, 37) = .01$, ns].

By definition, non-optimal paths were not the most efficient way of solving the Tower of Hanoi puzzle. It also turned out that when participants recounted non-optimal paths, their descriptions were often inaccurate accounts of their own behavior. In other words, the path described in speech deviated from the path the participant had actually followed in solving the puzzle. We suspected that in many (if not all) of their explanations, participants were resolving the problem on-line rather than remembering their original solution. We further speculated that the points where the participants' descriptions diverged from their original solutions were likely to be, for them, trouble spots in the problem; that is, points where they were having difficulty deciding on a next move. Thus, we hypothesized that the moments when a participant's description of a path deviated from the path that participant actually took were those moments when the participant was considering alternative moves, a personal choice point. To test this

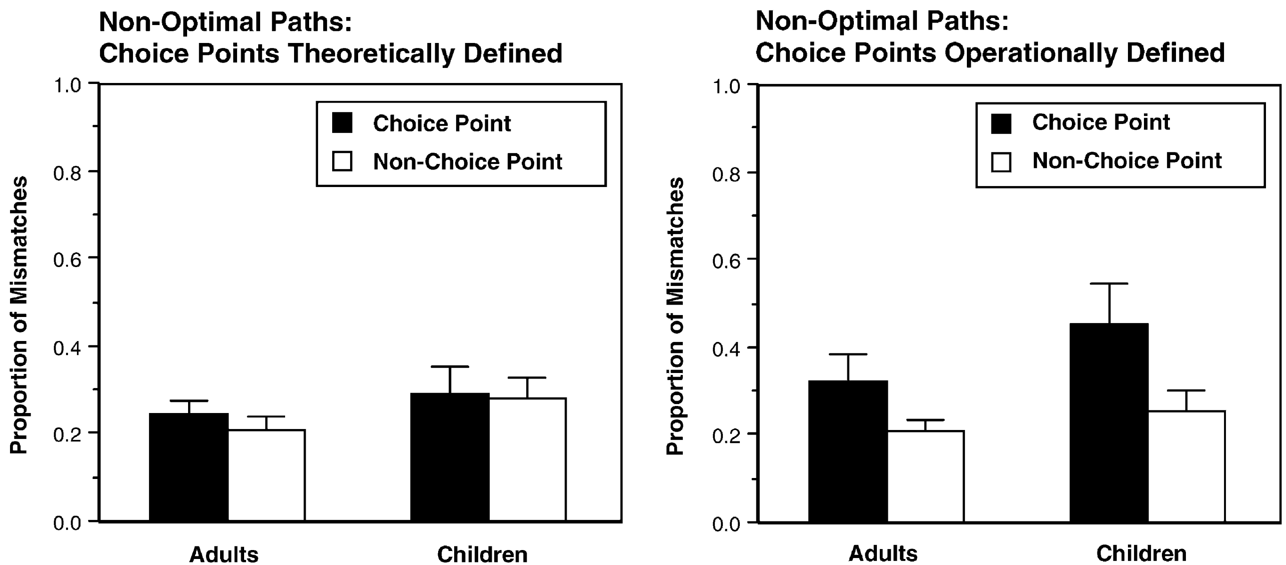

Fig. 4. Mismatches in non-optimal paths. The proportion of gesture–speech mismatches that adults and children produced when describing choice point moves and non-choice point moves in non-optimal paths. In the graph on the left, choice points were defined theoretically (the point at which a subroutine ought to be activated to solve the problem in the least number of possible moves). In the graph on the right, choice points were defined operationally (the point at which the participant's verbally described path diverged from the path the participant actually followed when solving the problem). Error bars indicate standard errors.

hypothesis, we reexamined the non-optimal paths, this time dividing the moves into *operationally defined* choice points (those points where the described path first deviated from the path actually taken when the problem was solved) and non-choice points. We again calculated the proportion of mismatches at each type of move. The right graph of Fig. 4 presents the data.[7] The participants produced more mismatches when describing the operationally defined choice points than when describing the non-choice points [$F(1, 36) = 11.46$, $p = .002$]. There was no effect of age [$F(1, 36) = 2.16$, ns] and no interaction [$F(1, 36) = 1.34$, ns]. Thus, we have evidence from the participants' gesture–speech mismatches that they were entertaining two paths at their own operationally defined choice points.

4. Discussion

Much of the previous work on gesture in problem-solving tasks has focused on tasks in which the problem-solver must generate an end state for the problem (e.g., Piagetian conservation or mathematics problems). In contrast, the Tower of Hanoi is a task whose beginning and end points are known to the problem-solver, and whose challenge lies in figuring out how to go from the known initial state to the known end state. In our current study, we used this classic problem to make three points about gesture–speech mismatch as an index of underlying problem-solving processes.

First, it is widely believed that there are particular moments in the Tower of Hanoi puzzle when subroutines need to be planned (Anzai & Simon, 1979; Klahr, 1978). We found here that these are precisely the moments when problem-solvers produce the greatest number of gesture–speech mismatches. Mismatches occur when there are strategies to be decided between (i.e., on choice point moves), not when strategies have already been chosen and are being executed. Note that planning seems to be an important factor in eliciting mismatches on the Tower of Hanoi puzzle. On every move in the puzzle, problem-solvers have several paths open to them. Mismatches arise not because several moves are possible, but because the problem-solver is entertaining alternative strategies and formulating a plan of action. Once the plan (or subroutine) has been activated, the problem-solver has, in effect, no choices for the next several moves, and thus produces few gesture–speech mismatches—until the next choice point is encountered.

Second, we also used mismatch to explore when participants who did not solve the problem efficiently entertained alternative paths. In these cases, it was not the theoretically defined choice points that elicited mismatches, but rather the moments when a problem-solver's description first deviated from the path he or she actually took when solving the puzzle. We hypothesized that these points of divergence might reflect trouble spots for individual problem-solvers—moments when the solver was uncertain about choices among two or more alternatives. The fact that both the adults and children tended to produce many gesture–speech mismatches at precisely these moments supports this hypothesis. By extension, to the extent that mismatch is a valid index of cognitive uncertainty, mismatches can themselves be used to identify choice points during the process of solving a variety of problems in future studies.

Third, we have extended descriptions of gesture–speech mismatch in adults. Adults have previously been shown to produce gestures spontaneously in narrations (McNeill, 1992), in descriptions of how gears move (Perry & Elder, 1997; Schwartz & Black, 1996), and in teachers' descriptions of mathematics lessons (Goldin-Meadow & Singer, 2002; Goldin-Meadow, Kim, & Singer, 1999). The findings presented here go beyond identifying yet another context in which adults gesture. We demonstrate here that, under similar problem-solving situations, adults and children produce gesture–speech mismatches in comparable cognitive contexts—that is, at choice points in the Tower of Hanoi problem. Previous attempts to compare problem-solving in children and adults have often failed to establish comparable conditions across the two groups (but see Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). Thus, an important contribution of this study lies in comparing adults' and children's problem-solving on the same task, and demonstrating that gesture can offer insight into problem-solving in adult and child alike.

Notes

1. Nine-year-old children were chosen for this study because 9 is the age when children can, on their own, solve the Tower of Hanoi puzzle at least some of the time (Byrnes & Spitz, 1979). Although 4- and 5-year-old children are able to solve the problem, they do so only when exceptionally intelligent (Kanevsky, 1989), when given experimenter support (Kanevsky, 1989; Klahr & Robinson, 1981), or when given an abbreviated task (Borys, Spitz, & Dorans, 1982; Klahr & Robinson, 1981; Spitz, Webster, & Borys, 1982).
2. After the participants gave their step-by-step descriptions, they were asked to describe any overall strategies they had for solving the problem. These metacognitive commentaries (e.g., "I basically grouped the disks so that I could free up the last peg") are less likely than verbatim problem-solving to accurately access cognitive processing (Ericsson & Simon, 1980, 1993) and thus were not analyzed here.
3. There were 54 non-optimal paths: 33 followed a circuitous route to the goal; 11 followed a circuitous route but did not reach the goal; 10 followed a straight route down one of the legs of the triangle in Fig. 1 but did not reach the goal.
4. Of the 7 adults and 8 children who produced descriptions of both optimal and non-optimal paths, 6 adults and 6 children began by producing non-optimal path descriptions and then produced optimal path descriptions; 1 adult and 2 children produced non-optimal path descriptions on trials 1 and 3, and an optimal path description on trial 2.
5. The few ($<$6%) moves in which participants produced speech with no gesture and gesture with no speech were also classified as matches simply because these responses did not convey two distinct paths.
6. Note that gesture and speech never convey exactly the same information. For example, pointing at a particular disk is not the same as saying "that disk"—the point provides information about the object's location whereas the speech provides information about the category to which the object belongs, yet both modalities identify the disk. Thus, the extent to which gesture and speech diverge in the information they convey is always one of degree (for discussion, see Goldin-Meadow, in press). As a result, the cognitive significance of gesture–speech mismatch on a particular task must be evaluated in relation to that task. For the Tower of Hanoi, gesture must identify a different path from the path identified in speech in order for the response to qualify as a gesture–speech mismatch.

7. Only one child's non-optimal path descriptions did not deviate at any point from his solutions. This child was consequently eliminated from the analysis of operationally defined choice points. This accounts for the N of 36 in the right graph of Fig. 4 and the N of 37 in the left graph.

References

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

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

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review*, *86*(3), 124–140.

Bidell, T. R., & Fischer, K. W. (1995). Developmental transitions in children's early on-line planning. In M. M. Haith, J. B. Benson, R. J. Roberts Jr., & B. F. Pennington (Eds.), *The development of future-oriented processes*. Chicago: University of Chicago Press.

Borys, S. V., Spitz, H. H., & Dorans, B. A. (1982). Tower of Hanoi performance of retarded young adults and nonretarded children as a function of solution length and goal state. *Journal of Experimental Child Psychology*, *33*, 87–110.

Byrnes, M. M., & Spitz, H. H. (1979). Developmental progression of performance on the Tower of Hanoi problem. *American Journal of Mental Deficiency*, *81*, 561–569.

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

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Egan, D. E., & Greeno, J. G. (1973). Acquiring cognitive structure by discovery and rule learning. *Journal of Educational Psychology*, *64*, 85–97.

Ekman, P., & Freisen, W. V. (1969). The repertoire of nonverbal behavioral categories: Origins, usage, and coding. *Semiotica*, *1*, 49–98.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Ernst, G. W., & Newell, A. (1969). *GPS: A case study in generality and problem solving*. New York: Academic Press.

Ewert, P. H., & Lambert, J. F. (1932). The effect of verbal instructions upon the formation of a concept. *Journal of General Psychology*, *6*, 400–413.

Feyereisen, P., & de Lannoy, J.-D. (1991). *Gestures and speech: Psychological investigations*. Cambridge: Cambridge University Press.

Gagne, R. M., & Smith, E. C. (1962). A study of the effects of verbalization on problem solving. *Journal of Experimental Psychology*, *63*, 12–18.

Garber, P. (1997). Using gesture and speech to index understanding of a problem solving task: A comparative study of adults and children explaining the Tower of Hanoi puzzle. Unpublished Ph.D. dissertation, University of Chicago, Chicago.

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

Goldin-Meadow, S. (in press). *Hearing gestures: How our hands help us think*. Cambridge: Harvard University Press, in press.

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

Goldin-Meadow, S., Nusbaum, H., Garber, P., & Church, R. B. (1993b). Transitions in learning: Evidence for simultaneously activated strategies. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(1), 92–107.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology*, *91*, 720–730.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Sciences*, *12*, 516–522.

Goldin-Meadow, S., & Singer, M. (2002). From children's hands to adults' ears: Gesture's role in teaching and learning. Submitted for publication.

Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2, pp. 21–41). Hillsdale, NJ: Lawrence Erlbaum.

Kanevsky, L. (1989). Pursuing qualitative differences in the flexible use of problem-solving strategy by young children. *Journal for the Education of the Gifted*, *13*(2), 115–140.

Karat, J. (1982). A model of problem-solving with incomplete constraint knowledge. *Cognitive Psychology*, *14*, 538–559.

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

Klahr, D. (1978). Goal formation, planning, and learning by preschool problem solvers, or: 'My socks are in the dryer'. In R. S. Siegler (Ed.), *Children's thinking: What develops*? (pp. 181–212). Hillsdale, NJ: Lawrence Erlbaum.

Klahr, D., & Robinson, M. (1981). Formal assessment of problem solving and planning processes in preschool children. *Cognitive Psychology*, *13*, 113–148.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from the Tower of Hanoi. *Cognitive Psychology*, *17*, 248–294.

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

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

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

Perry, M., & Elder, A. D. (1997). Knowledge in transition: adults' developing understanding of a principle of physical causality. *Cognitive Development*, *12*(1), 131–157.

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: induction and fallback. *Cognitive Science*, *20*, 457–497.

Spitz, H. H., Webster, N. A., & Borys, S. V. (1982). Further studies of the Tower of Hanoi problem-solving performance of retarded young adults and nonretarded children. *Developmental Psychology*, *18*(6), 922–930.