ILLUMINATING MENTAL REPRESENTATIONS THROUGH SPEECH AND GESTURE

Martha W. Alibali,1 Miriam Bassok,2 Karen Olseth Solomon,3 Sharon E. Syc,4 and Susan Goldin-Meadow5

1Carnegie Mellon University, 2University of Washington, 3Northwestern University, 4The Erikson Institute, and 5University of Chicago

Abstract—Can the gestures people produce when describing algebra word problems offer insight into their mental representations of the problems? Twenty adults were asked to describe six word problems about constant change, and then to talk aloud as they solved the problems. Two problems depicted continuous change, two depicted discrete change, and two depicted change that could be construed as either continuous or discrete. Participants’ verbal and gestured descriptions of the problems often incorporated information about manner of change. However, the information conveyed in gesture was not always the same as the information conveyed in speech. Participants’ problem representations, as expressed in speech and gesture, were systematically related to their problem solutions. When gesture reinforced the representation expressed in the spoken description, participants were very likely to solve the problem using a strategy compatible with that representation—much more likely than when gesture did not reinforce the spoken description. The results indicate that gesture and speech together provide a better index of mental representation than speech alone.

A growing body of research shows that spontaneous hand gestures can be used to assess problem-solving strategies, among both children (Church & Goldin-Meadow, 1986; Perry, Church, & Goldin-Meadow, 1988) and adults (Garber, 1997; Perry & Elder, 1996; Schwartz & Black, 1996). Indeed, gestures sometimes reveal problem-solving strategies that are not expressed at all in speech (Garber, Alibali, & Goldin-Meadow, 1998; Goldin-Meadow, Alibali, & Church, 1993). In this article, we examine whether spontaneous gestures provide information not about how people solve problems, but about how they represent problems. Specifically, we report the results of a study exploring the gestures people produce when describing written problems that they have not yet solved, and examine the information that such gestures reveal about people’s mental representations of those problems. By “mental representation,” we mean an internal copy or model of a problem.

Mental representations have been postulated to underlie a wide variety of behaviors related to problem solving, including strategies, errors, latencies, and patterns of transfer (e.g., Bassok, Wu, & Olseth, 1995; Kotovsky, Hayes, & Simon, 1985; Lovett & Schunn, in press). Because mental representations are invoked to explain such behaviors, it is important to be able to assess representations independently of the behaviors to be explained. In this study, we assessed mental representations of problems by asking adults to describe the problems to a naive listener before solving the problems. We then analyzed both the speech and the gestures that the adults produced in their descriptions of the problems. Our central hypothesis was that speech and gesture, taken together, provide a more complete picture of mental representations than does speech alone.

Why might gestures be helpful in assessing people’s mental representations of problems? Many researchers draw inferences about representations of the basis of respondents’ speech, either verbal descriptions of problems (e.g., Brown, Kane, & Echols, 1986) or verbal protocols (e.g., Simon & Hayes, 1976). However, respondents’ speech is likely to be heavily influenced by the words used in the problems themselves. Furthermore, verbal reports may be incomplete (Ericson & Simon, 1993), systematically omitting information that is difficult to verbalize (e.g., information represented in visual images). For these reasons, speech may be a less-than-optimal tool for assessing how people represent problems.

Spontaneous gestures are not subject to these same limitations. First, unlike speech, gestures cannot be mindlessly modeled on the problem text, because text is typically presented without accompanying gestures. Second, gesture is particularly well suited to convey visual and perceptual information. Therefore, gesture may provide a window onto knowledge that is not readily expressed in speech. For example, it may be difficult to describe an irregular shape in speech, but easy to depict the shape in gesture.

It seems likely that visual or perceptual knowledge may be incorporated into people’s mental representations of problems. Indeed, it is thought that problem solvers construct mental models of texts or problem situations that incorporate information, including perceptual information, drawn from semantic knowledge (Hayes & Simon, 1974; Johnson-Laird, 1983; Kintsch & Greeno, 1985; van Dijk & Kintsch, 1983). We hypothesize that such mental models might naturally lead to the production of spontaneous gestures, which iconically represent perceptual properties of the models. Indeed, Schwartz and Black (1996) have argued that spontaneous hand gestures are “physically instantiated mental models” (p. 464).

The central goal of this study was to determine whether people’s spontaneous gestures can be used as a tool for illuminating their mental representations of problems. To address this issue, we chose to study problems that lend themselves to the construction of distinct mental models—algebra word problems about continuous and discrete constant change (adapted from Bassok & Olseth, 1995). Discrete-change problems focus on change over a series of steps, such as change in the number of books on each shelf of a six-shelf bookcase. In contrast, continuous-change problems focus on change over a single, nonpartitioned event, such as change in the amount of air pressed per minute into a hot air balloon over a 30-min period (see the appendix). We (Bassok & Olseth, 1995) have suggested that people spontaneously incorporate information about the manner of change into their representations of such problems. However, in this study, we did not...
directly assess the representations constructed by the problem solvers. Instead, we inferred this information from a different group of participants who performed a graph-choice task (selecting a discrete step curve or a smooth continuous curve to represent each problem).

In the present study, we asked 20 adults first to describe, and then to solve, six constant-change problems. We hypothesized that participants’ spoken and gestured descriptions of the problems would convey information regarding the manner of change (continuous or discrete) that they mentally represented for the problems. We tested this hypothesis by comparing the relationship of the spoken and gestured descriptions of the problems to the solutions the participants subsequently provided.

**METHOD**

**Participants and Procedure**

Twenty undergraduate students (8 male, 12 female) were paid to participate. Each solved a set of six structurally analogous word problems that involved constant change (adapted from Bassok & Olseth, 1995; see the appendix). The set included two problems about entities that changed continuously, two about entities that changed discretely, and two about entities that could be construed as changing either discretely or continuously (mixed problems). Problems were presented one at a time on separate sheets of paper in one of two fixed orders. Participants in the continuous-first group (n = 10) solved the continuous-change problems first, followed by the discrete-change and then the mixed problems. Participants in the discrete-first group (n = 10) solved the discrete-change problems first, followed by the continuous-change and then the mixed problems.

The experimental session involved the experimenter, the participant, and a confederate who pretended to be another participant. The participant was told that he or she would be asked to complete six problems while the other “participant” listened, and then roles would be reversed. The participant was asked to read the first problem silently and return it to the experimenter. The participant was then asked to describe the gist of the problem to the other “participant.” We asked participants to describe the problems to a confederate rather than to the experimenter in order to ensure that the participants assumed no prior knowledge of the problems on the part of the listener. Following the problem description, the participant was asked to talk aloud while solving the problem. This procedure was repeated for each of the six problems. Participants were not expected to complete all of the calculations necessary to solve each problem, but simply to give an account of how they would go about arriving at a solution. The session was videotaped.

**Coding Verbal and Gestural Descriptions of the Problems**

**Coding speech**

Participants’ verbal descriptions of the problems were transcribed and divided into clauses, using breath pauses and syntactic criteria (see Ericsson & Simon, 1993). Each clause was then coded as providing cues to a continuous representation, a discrete representation, both representations, or neither representation. Cues that indicated a representation of change as a single event were coded as continuous, whereas cues that indicated a representation of change as a set of individual events were coded as discrete (see Bassok & Olseth, 1995). Specifically, the following verbal cues were taken as evidence of a continuous representation: (a) mention of the values in the problem using ratelike units (e.g., “It started at 10 L per minute”), (b) reference to the entire period of time or space involved in the problem (e.g., “over the 12-min period”), and (c) explicit reference to a rate (e.g., “The number of books went up at a constant rate”). The following verbal cues were taken as evidence of a discrete representation: (a) mention of the values in the problem using amountlike units (e.g., “At first there were 10 L”), (b) reference to the individual units of time or space involved in the problem (e.g., “in each of the 12 min”), and (c) explicit reference to “a constant” (e.g., “The number of books increased by a constant”). Note that many of these verbal cues can also be found in the problem texts (cf. the appendix).

For each verbal description of a problem, we tabulated the number of clauses reflecting each type of representation, and then scored whether the overall description reflected a continuous representation (i.e., more continuous than discrete clauses), a discrete representation (i.e., more discrete than continuous clauses), or both representations equally (i.e., equal numbers reflecting the two models).

**Coding gesture**

The stream of manual movement was divided into individual gestures using previously established criteria (Church & Goldin-Meadow, 1986). The shape, motion, placement, and orientation of the hands in each gesture were transcribed without access to the audio portion of the tape (i.e., with the sound turned off). Each gesture was then classified as conveying a continuous representation, a discrete representation, both representations, or neither representation (for further details, see Alibali, Bassok, Olseth, Syc, & Goldin-Meadow, 1995). Gestures that incorporated smooth, continuous motions (e.g., sweeping, arcing, dragging) were coded as reflecting a continuous representation. Gestures that incorporated a set of discrete movements (e.g., a sequence of three or more taps, points, beats) were coded as reflecting a discrete representation. Gestures that incorporated zigzagging, circling, or spiraling motions were also coded as discrete. In rare cases, individual gestures combined features of both continuous and discrete representations. For example, gestures that incorporated a series of short sweeping motions were coded as reflecting both representations. All other gestures were coded as reflecting neither representation. Such gestures included simple beat (rhythmic) gestures, points, flicks, and gestures that represented aspects of the problem other than manner of change.

For each gestured description of a problem, we tabulated the number of gestures reflecting each representation, and then scored whether the overall description reflected a continuous representation (more continuous than discrete gestures), a discrete representation (more discrete than continuous gestures), both representations equally (equal numbers of discrete and continuous gestures), or neither representation (no gestures that conveyed either model).

**Coding the Strategies Used to Solve the Problems**

As in previous work (Bassok & Olseth, 1995), participants were found to use two types of correct strategies to solve the problems: (a) the average strategy, which involves finding an average rate of change and multiplying it by the number of units (e.g., for the potato problem: “I would find the average number of potatoes a day, by adding the two...
rates and dividing by 2, and then multiply by the number of days”), and (b) the sum strategy, which involves calculating the change per unit (i.e., per increment) and summing these amounts (e.g., for the bookshelf problem: “It increases five times, so we take 45 and subtract 15...the total increase is 30, and we divide that by 5...so it increases by 6 per shelf...we just take the first shelf as 15, and 21, and then 27, and [so on]...we just add them all together to get the total number of books”). All six problems could be solved using either of these two strategies.

Note that the average strategy is compatible with a representation of the change as a single event (i.e., a continuous representation), whereas the sum strategy is compatible with a representation of the change as a set of discrete events (i.e., a discrete representation). Participants occasionally offered both the sum and the average strategies for a single problem. Participants also frequently used incorrect solution strategies, some of which incorporated pieces of both the sum and the average strategies.

Reliability

To assess reliability, a second observer recoded 20% of the data. Agreement between coders was 96% (N = 24) for assessing the strategies used to solve the problems, 88% (N = 24) for coding verbal descriptions of the problems (as discrete, continuous, or equal), and 88% (N = 24) for coding gestured descriptions of the problems (as discrete, continuous, equal, or neither).

RESULTS

Problem Representations

Speech alone

We first examined participants’ verbal descriptions of the problems. As predicted, participants tended to articulate discrete representations on discrete-change problems, and continuous representations on continuous-change problems (Fig. 1a). Mixed problems were described primarily with continuous representations. This pattern is not surprising given that the written texts for the discrete-change problems contained cues reflecting a discrete representation, and the texts for the continuous-change and mixed problems contained cues reflecting a continuous representation.

To evaluate this pattern statistically, we assigned participants 1 point for each problem described with a discrete representation in speech. Thus, participants received a score from 0 to 2 for each problem type. We then used repeated measures analysis of variance (ANOVA) with problem type (continuous, discrete, or mixed) as a within-subjects factor and problem-order group (continuous-first, discrete-first) as a between-subjects factor. As expected, there was a significant main effect for problem type, F(2, 36) = 111.92, p < .001. In addition, with results collapsed across problem types, participants in the discrete-first group expressed discrete representations more often than participants in the continuous-first group (M = 3.00, SD = 1.05 vs. M = 2.20, SD = 0.63), F(1, 18) = 4.24, p < .06.

Gesture alone

Another source of evidence about problem representations—one not open to influence from the problem texts—is participants’ spontaneous gestures. Indeed, in describing the problems, all participants produced gestures, and these gestures often revealed information about their mental models of the problems. On average, participants produced gestures that conveyed manner of change on 3.7 of the 6 problems (SD = 2.0, range: 0–6), and produced 1.5 such gestures per problem (SD = 1.3, range: 0–9).

As for the verbal descriptions, we assigned participants 1 point for each problem described with a discrete representation in gesture. As shown in Figure 1b, participants expressed discrete representations in gesture most often on discrete-change problems, and somewhat less often on continuous-change and mixed problems, yielding a main effect for problem type, F(2, 36) = 6.87, p < .005. However, the effect was less dramatic in gesture than in speech (see Fig. 1).

Relations Between Representations and Solutions

We next examined whether the mental models participants expressed in their descriptions of the problems were systematically
related to the strategies that they subsequently used to solve the problems. We hypothesized that if participants represented the change as discrete, they would use the sum strategy. If they represented the change as continuous, they would use the average strategy.

**Speech alone**

We first examined whether solution strategies could be predicted on the basis of the representations participants expressed in their verbal descriptions of the problems. As expected, the distribution of solution strategies differed as a function of the representation conveyed in speech (Fig. 2a). Participants used the sum strategy most often on problems that they described in speech as discrete. Participants decreased their use of the sum strategy and increased their use of the average strategy on problems that they described in speech as continuous. These findings are consistent with previous work showing that the sum strategy is the most common strategy used by adults on both discrete- and continuous-change problems. The average strategy is used primarily on problems that involve continuous change, and only part of the time (Bassok & Holyoak, 1989; Bassok & Olseth, 1995).

**Gesture alone**

We next examined whether solution strategies could be predicted on the basis of the representations participants expressed in their gestured descriptions of the problems (Fig. 2b). Once again, participants used the sum strategy most often on problems that they described in gesture as discrete. Participants decreased their use of the sum strategy and increased their use of the average strategy on problems that they described in gesture as continuous.

**Using Gesture and Speech Together to Assess Representation and Predict Solutions**

The analyses thus far demonstrate that, independently, speech or gesture can be used to infer problem representations, and that the representations expressed in each modality are systematically related to problem solutions. However, gestures are almost always produced in the context of speech. In principle, gesture can reflect a representation that either matches or mismatches the representation reflected in the speech it accompanies.

The participants conveyed the same predominant representation in both speech and gesture (both continuous or both discrete) on 28% of the problems. In such cases, gesture reinforced speech. On 45% of the problems, the participants conveyed no preference in gesture (typically one continuous gesture and one discrete gesture), or no relevant gestures at all. In such cases, gesture was neutral with respect to speech. On 28% of the problems, the participants conveyed different representations in speech and gesture (i.e., a continuous representation in speech with a discrete representation in gesture, or vice versa). In such cases, gesture conflicted with speech.

We examined whether variations in the relationship between speech and gesture in the problem descriptions were associated with different patterns of problem solutions. To address this issue, we assessed the distribution of solution strategies as a function of the relationship between gesture and speech. Figure 3 (hatched bars) displays the proportion of problems solved using a strategy compatible with the verbal representation (i.e., sum strategy for problems with a discrete verbal representation and average strategy for problems with a continuous verbal representation), as a function of the relationship between speech and gesture. When gesture reinforced speech (i.e., when gesture and speech converged on the same representation), participants were very likely to use the solution strategy compatible with that representation (sum for discrete, average for continuous). Participants used the strategy compatible with the verbal representation less often when gesture was neutral with respect to speech, and even less often when gesture conflicted with speech. Indeed, when gesture conflicted with speech, participants were quite likely to use the strategy compatible with the gestured representation (Fig. 3, dark bar).

To evaluate this pattern statistically, we used logistic regression to estimate the odds ratios for using the strategy compatible with the representation expressed in speech (i.e., the odds of using that strategy to not using that strategy) as a function of the gesture-speech relationship (reinforcing, neutral, or conflicting). We included a parameter for each participant in the model, so that the overall odds ratios for the gesture-speech relationship would estimate the common odds ratios across
participants. The analysis revealed a significant overall effect of the gesture-speech relationship, Wald $\chi^2 (2, N = 110) = 9.34, p < .01$. The odds ratios was higher when gesture reinforced speech than when gesture was neutral with respect to speech, Wald $\chi^2 (1, N = 82) = 3.45, p = .06$, and higher when gesture reinforced speech than when gesture conflicted with speech, Wald $\chi^2 (1, N = 61) = 9.21, p < .005$. Thus, variations in the relationship between speech and gesture were associated with variations in strategy use.

To provide a qualitative sense of the data, Figure 4 displays the proportion of problems solved with each solution strategy as a function of the gesture-speech relationship, for problems described with discrete representations in speech (Fig. 4a) and with continuous representations in speech (Fig. 4b). For problems described with discrete verbal representations, when gesture conveyed a discrete representation and therefore reinforced speech, participants always used the sum strategy, and never used the average strategy. When gesture was neutral or conflicted with speech, participants decreased their use of the sum strategy and slightly increased their use of the average strategy. A complementary pattern is seen for problems described with continuous verbal representations. When gesture conveyed a continuous representation and therefore reinforced speech, participants used the average strategy more often than the sum strategy to solve the problems. When gesture was neutral or conflicted with speech, participants increased their use of the sum strategy and decreased their use of the average strategy. Thus, the strategies used to solve the problems varied systematically as a function of how those problems were represented in both speech and gesture.

**DISCUSSION**

Our findings demonstrate that spontaneous gestures reveal important information about people’s mental representations of problems. In this study, speech and gesture together provided a more complete preview of solution strategies than speech alone. These findings add to the growing body of literature showing that gestures can provide information about problem-solving processes that is not revealed through speech (Alibali, 1999; Crowder & Newman, 1993). In particular, gestures provide information not only about problem-solving strategies, as previous work has shown, but also about problem representations.

Our findings bear on the question of whether gesture and speech derive from two independent, competing representations of a problem or from two facets of a single representation. When gesture and speech
converge on the same representation, they provide a very powerful predictor of problem solution. In such cases, it seems parsimonious to argue that gesture and speech derive from a single representation. Indeed, previous research supports the view that gesture and speech form an integrated system, rather than two independent systems (Alibali & Goldin-Meadow, 1999a; Goldin-Meadow et al., 1993; McNeill, 1992). However, on some problems, the participants in this study expressed different representations in speech and gesture (i.e., discrete in speech and continuous in gesture, or vice versa). When gesture and speech conflicted, participants’ problem solutions were as likely to correspond with the representation they expressed in gesture as with the representation they expressed in speech (cf. Fig. 3). Thus, in cases in which gesture and speech conflict, it seems possible that gesture and speech derive from independent representations.

We have shown here that gesture and speech do not always convey the same information. Such “mismatches” have also been documented in children and adults solving other types of problems, including Piagetian conservation tasks (Alibali, Evans, & McNeill, 1999; Church & Goldin-Meadow, 1986), mathematical equations (Alibali & Goldin-Meadow, 1993b; Perry et al., 1988), Tower of Hanoi problems (Garber, 1997), and problems about bears (Perry & Elder, 1996). The frequent production of mismatches has been associated with transitional knowledge in children acquiring conservation (Church & Goldin-Meadow, 1986) and symbolic equivalence (Alibali & Goldin-Meadow, 1993a; Perry et al., 1988). It remains an open question whether gesture-speech mismatches in constant-change problems also index a transitional knowledge state (i.e., a period of imminent change in the modal problem-solving strategy).

Our findings indicate that problem solvers sometimes represent aspects of the problems that they do not use when solving those problems. For example, participants sometimes represented the fact that change was continuous when describing a problem, but then solved that problem using a strategy that did not incorporate this information (the sum strategy). These findings are consistent with previous reports that children sometimes represent features of problems—either in speech (Siegel, 1984, 1985) or in gesture (Garber et al., 1998)—that they do not exploit when they solve those problems. In future studies, we plan to examine whether problem solvers rely on these encoded-but-not-used features when called upon to generate new strategies. One might predict, for example, that participants who encode and represent the fact that change is continuous in a given problem might be especially likely to generate the average strategy as opposed to the sum strategy in subsequent problems.

In this study, problem content (i.e., whether change was portrayed as a series of steps or a single event) was found to influence people’s problem representations. We acknowledge that content is only one of many potential factors that might influence how a problem is represented. In turn, there are undoubtedly many factors other than representation that can influence how a problem is solved. Although we have uncovered systematic relationships between representations and solution strategies, we were not able to perfectly predict participants’ solution strategies. Indeed, in some cases, participants seemed not to have a strategy in their repertoire that “fit” their problem representation. For example, participants sometimes represented the fact that problem entities changed continuously, but were unable to incorporate this information in a problem-solving strategy. In such cases, participants sometimes gave up, sometimes constructed new (often incorrect) strategies on the spot, and sometimes appeared to rerepresent the problem as discrete before solving it.

We believe that understanding how people represent problems will help to explain how they construct and choose among problem-solving strategies. Our data suggest that not all of what people know about problems is revealed in their speech. In at least some cases, some facets of their mental representations are expressed in gestures, and these gestures sometimes conflict with speech. Our findings suggest that mental representations often include visual or perceptual information, which may at times be more readily expressed in gesture than in speech. In these cases, spontaneous gestures can be a valuable tool for illuminating mental representations.

REFERENCES


Acknowledgments—This research was supported by a P.E.O. Scholar Award to Martha Alibali, by a University of Chicago Social Sciences Divisional Research Grant to Miriam Bassok, and by a grant from the Spencer Foundation to Susan Goldin-Meadow. We thank Brian Junker for statistical advice and Gavin Huntley-Fenner, Ken Kotovsky, Robert Krauss, Eliza Beth Littleton, William Meadow, Bethany Rittle-Johnson, Christian Schunn, and Dan Schwartz for helpful discussions and constructive feedback. We also thank Michael Kang, Cynthia Lubin, Leslie Petasis, Audrey Russo, and Eliza Beth Littleton for assistance in transcribing and coding the data.


(Received 5/22/98; Accepted 9/10/98)

**APPENDIX: PROBLEMS USED IN THE STUDY**

The following problems were presented to participants in this study.

**Discrete Problems**

For a lecture, 10 rows of chairs have been arranged in a lecture hall. The chairs have been set up such that the number of chairs in each row increases by a constant from the number of chairs in the previous row. If there are 25 chairs in the first row and 115 chairs in the 10th row, how many chairs total are there in the lecture hall?

A bookcase has 6 shelves. The number of books on each successive shelf from top to bottom increases by a constant from the number of books on the shelf above it. If there are 15 books on the top shelf and 45 books on the bottom shelf, how many books total are in the bookcase?

**Continuous Problems**

It takes 35 minutes to inflate a hot air balloon. The rate at which the hot air is pressed into the balloon increases steadily from 10 liters/minute at the beginning of the first minute to 80 liters/minute at the end of the 35th minute. How many liters of hot air are pressed into the balloon over the 35 minute period?

The speed of an airplane increases at a constant rate during a period of 12 minutes from 10 miles/minute to 34 miles/minute. What distance, in miles, will the plane travel during the 12 minute period?

**Mixed Problems**

After a seven day harvest, a potato farmer notices that his rate of gathering potatoes increased steadily from 35 bushels/day to 77 bushels/day. How many bushels of potatoes total did the farmer collect during the seven day harvest?

During the last 6 years, the rate of population growth in the town of Mudville increased steadily from 300 people/year to 1500 people/year. How many people total were added to the population during the 6 year period?