Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Memory and Language 50 (2004) 395–407

Journal of Memory and Language

www.elsevier.com/locate/jml

Probing the mental representation of gesture: Is handwaving spatial?

Susan M. Wagner,* Howard Nusbaum, and Susan Goldin-Meadow

Department of Psychology, University of Chicago, 5848 S. University Ave., Chicago, IL 60637, USA

Received 16 May 2003; revision received 9 January 2004

Abstract

What type of mental representation underlies the gestures that accompany speech? We used a dual-task paradigm to compare the demands gesturing makes on visuospatial and verbal working memories. Participants in one group remembered a string of letters (verbal working memory group) and those in a second group remembered a visual grid pattern (visuospatial working memory group) while explaining math problems. If gesture production is mediated by visuospatial representation, gesturing should interfere more with performance on the concurrent visuospatial task than the concurrent verbal task. We found, however, that participants *in both groups* remembered significantly more items when they gestured than when they did not gesture. Moreover, the number of items remembered depended on the meaning conveyed by gesture. When gesture conveyed the same propositional information as speech, participants remembered more items than when it conveyed different information. Thus, in contrast to simple handwaving, the demands that gesture makes on working memory appear to be propositional rather than visuospatial.

Keywords: Gesture; Working memory; Spatial representation

When people talk, they move their hands. But these movements, typically called gestures, are not just handwaving—they are closely related to and may even be beneficial for cognitive processing (Goldin-Meadow, 2003; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). Given the relevance of speech-accompanying gestures to thinking, it becomes important to explore the nature of the mental representations that underlie these gestures. Information in speech is typically described in terms of propositional structures (e.g., Anderson, Budiu, & Reder, 2001; Carpenter & Just, 1975; Clark & Chase, 1972; Gernsbacher, 1990), but the same types of structures need not underlie gesture. In fact, gesture seems like a particularly good vehicle for capturing visual and spatial information—its visuospatial form makes it perfect for capturing visuospatial content.

For example, consider a speaker explaining a spatial layout of a town or neighborhood. Speakers tend to produce gestures in such contexts (Emmorey, Tversky, & Taylor, 2000). A speaker might say, "The hospital is on one side and on the other is the school and in the middle is the parking lot," while pointing first to one location in space, then to another location, and finally to a location in between the first two points. In this example, speech and gesture are both characterizing the spatial relationships among entities.

However, people produce gestures along with their speech even when the content of their talk is not spatial. The hand movements and shapes of gesture can use space and visual form to depict information that itself is inherently *nonspatial*. Thus, the information conveyed by gesture can be in a visuospatial form even when the speaker's message is not visuospatial. Consider a speaker explaining the mathematical equation

---

* Corresponding author. Fax: 1-773-702-0886.
E-mail address: swagner@uchicago.edu (S.M. Wagner).

0749-596X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2004.01.002

$x^2 + 5x + 6 = (x + 2)(x + 3)$. The speaker might say, “two plus three is five” while at the same time pointing at the 2, the 3, making a “join together” gesture by closing the hand, and then pointing at the 5. In this example, the speech does not describe a spatial relationship. Rather, speech characterizes a mathematical relationship between the numbers in the equation (it pinpoints the elements and describes the operation performed on those elements). Gesture could be conveying precisely this same mathematical relationship, albeit in a visuospatial format. That is, pointing at the 2, the 3, and the 5 could be a visuospatial representation of the mathematical notion that two plus three is five. In this case, the content of gesture is mathematical (with a propositional underpinning) even though its form is visuospatial. Alternatively, gesture might merely be locating the items referred to in space and time, much like the gestures in the previous example that located the various buildings in space. In this case, the content of gesture, as well as the form of gesture, is visuospatial. The goal of the present study is to determine whether the representations underlying gesture are better characterized as propositional or visuospatial.

Previous observations of the tasks on which gesture is produced are consistent with the view that visuospatial representation underlies gesture. For example, speech with spatial content is particularly likely to be accompanied by gesture (Lavergne & Kimura, 1987), as are words that can be characterized as imagistic (Hadar & Krauss, 1999). Indeed, speech becomes less imagistic when speakers are restricted from gesturing (Rimé, Schiaratura, Hupet, & Ghysselinckx, 1984). In addition, speakers are more likely to gesture when they describe a cartoon they have seen (a visuospatial event) than when they retell a narrative of the same cartoon that they have heard (Hostetter & Hopkins, 2002). Even more convincing, patients with visuospatial deficits produce fewer gestures on a picture description task than do aphasic individuals or individuals without brain injury (Hadar, Burstein, Krauss, & Soroker, 1998). These studies suggest that gesture is associated with visuospatial information.

However, in each of these cases, the form and content of the information in gesture are confounded in that both are visuospatial. And, as the above math problem illustrates, we cannot infer that the representations underlying gesture are visuospatial simply because its form is visuospatial. Consider, for example, sign languages in deaf communities (e.g., American Sign Language) which are visuospatial in form but, like speech, are typically described in terms of propositional structures. To determine whether the representations underlying gesture are visuospatial or propositional, we need an independent measure of the mental representations associated with gesture.

Using a dual-task paradigm to tap underlying mental representations

Dual-task paradigms have frequently been used to investigate cognitive processes. In a dual-task paradigm, two tasks are performed simultaneously. If the processes used on these two tasks interfere with one another, those processes are assumed to be similar. Indeed, in the classic study by Brooks (1968), recall of verbal information was worse when it co-occurred with verbal processing than when it co-occurred with visuospatial processing presumably because both tasks relied on the same type of processing. Similarly, Wickens (1980) found more interference between two visuospatial processes or between two language processes than between a visuospatial process and a language process. Here again, processes that interfere with each other are assumed to make use of the same kind of cognitive resource (e.g., Wickens, 1980; although see Navon, 1984; for debate over the definition of and even the existence of cognitive resources).

In many dual-task paradigms, one of the two tasks is a working memory task (Baddeley, 1986). For example, Logan (1979) asked participants to hold a list of items in working memory while performing a classification task involving either controlled processing or automatic processing. He found that varying the number of items held in working memory had an effect on the controlled processing task but (as predicted) not on the automatic processing task. Logan thus used the interaction between two simultaneously occurring tasks (the working memory task and the classification task) to explore the processes involved in one of the two tasks (automatic vs. controlled processing in the classification task).[1]

In the present study, we also use two tasks (a working memory task and an explanation task) to explore the processes involved in one of the tasks (gesturing vs. not gesturing in the explanation task). We compared performance on two different types of working memory tasks (a verbal memory task and a visual memory task) to diagnose the representations that underlie gesture production. We have previously found that gesturing during an explanation task enhances recall on a concurrent verbal memory task—if asked to explain their solutions to a math task while at the same time remembering a list of verbal items (letters), speakers remember more of those items when gesturing than when not gesturing (Goldin-Meadow et al., 2001). But what would we predict if the concurrent task were visuospatial? If the mental

---

[1] It is important to note that this kind of interaction does not depend on specific assumptions of Baddeley’s working memory model, some of which have been called into question recently (see Nairne, 2002). The interaction is used to diagnose processes and representations at a much more abstract level than is posited in any particular model of a memory system (see Navon, 1984).

representations underlying gesture production are visuospatial, gesturing ought to *interfere* with performance on a concurrent visuospatial memory task simply because the two visuospatial representations make use of the same (visuospatial) working memory system. That is, the visuospatial representations underlying gesture ought to interfere with the visuospatial representation required for the memory task. In this event, gesturing will not improve the speaker's working memory, and speakers ought to remember *fewer* unrelated visuospatial items (dots on a grid) when gesturing than when not gesturing.

If, however, despite gesture's visuospatial form, the mental representations underlying gesture production are propositional, then switching from a verbal memory task to a visuospatial memory task should have no effect—the propositional representations underlying gesture ought not interfere with a concurrent visuospatial memory task any more than they interfere with a concurrent verbal memory task. Indeed, gesturing could improve the speaker's visuospatial working memory, just as it improves verbal working memory (Goldin-Meadow et al., 2001). Thus, speakers ought to remember *more* unrelated visuospatial items when they gesture during their explanations than when they do not gesture—just as they remember more unrelated verbal items when they gesture than when they do not gesture (Goldin-Meadow et al., 2001).[2]

We asked speakers to perform two tasks simultaneously: (1) They explained how they solved a math problem, and (2) they remembered a set of unrelated items. Each speaker was permitted to move freely when explaining some of the problems (i.e., gesture was permitted), but was prevented from moving his or her hands on other problems (i.e., gesture was prevented). We also manipulated the simultaneously performed working memory task: One group of participants was asked to remember and then recall a list of letters (the verbal working memory task); another group was asked to remember and then recall an array of dots on a grid (the visuospatial working memory task). If the mental representations underlying gesture are propositional, gesture production ought to facilitate recall on the visuospatial task just as it facilitates recall on the verbal task—participants should remember more dots and more letters when gesturing than when not gesturing. If, however, the mental representations underlying gesture are visuospatial, gesture production ought to interfere with performance on the visuospatial task (but not on the verbal task)—participants should remember fewer dots (but more letters) when gesturing than when not gesturing.

Methods

Participants

Seventy-two college-aged adults (43 females, 29 males) participated in the study.[3] Participants were recruited using an email list of individuals who were affiliated with the University of Chicago and who had expressed a desire to participate in psychology experiments. The adults were paid for their participation. Sixty-six of the participants had completed or were concurrently enrolled in a college-level calculus course, and 12 of these participants had completed math classes more advanced than calculus.

The task

Participants were asked to solve math problems of the form $x^2 \pm bx \pm c = ()()$ and then to explain their solutions. This task was chosen because it is not an inherently spatial task. In order to solve the problem, a participant must understand the mathematical relationships among the entities (e.g., the set of factors that equal c when multiplied and b when added). Note,

---

[2] A fundamental assumption underlying our approach is that working memory plays a role in communication. From the theoretical perspective of modularity (Fodor, 1983), language competence depends on highly evolved expert modules that are independent of other more general cognitive systems, like working memory. Under this view, there should be no interaction between language use and working memory load. However, there is now a substantial body of research demonstrating an important role for working memory in language processing. For example, individual differences in working memory capacity predict language comprehension performance (Just & Carpenter, 1992). But the importance of working memory in language processing is not always easy to see. Nusbaum and Schwab (1986) argued that it often looks like working memory is not involved in language processing simply because language processing does not typically challenge working memory capacity—the system must be taxed in order for working memory's role in language processing to be visible. As a result, dual-task paradigms, which tax the system, are often essential in demonstrating the interaction between language use and working memory load. And when dual-task paradigms are used, working memory turns out to interact with language processing at a variety of levels—early speech perception (Nusbaum & Morin, 1992), syntactic processing (Gordon, Hendrick, & Levine, 2002), lexical comprehension (see Baddeley, 1986). Furthermore, our previous work suggests that the gestures produced along with spoken language interact with working memory (Goldin-Meadow et al., 2001). We therefore felt comfortable using the dual-task paradigm to explore the nature of the representations that underlie gesture.

[3] Ten additional adults were tested but eliminated from the study: 7 because of problems with videotaping; 1 because she was unable to solve the problems correctly and changed her method of solution repeatedly during the experimental trials; 1 because she wrote during the explanations; 1 because she gestured during the no-movement condition.

however, that the gestures the participant produces to explain the task are spatial (as are all gestures). Thus, the format of the gestures produced on this task is visuospatial, but their content could be either propositional or visuospatial. The task thus allows us to ask whether the representations underlying gesture in such a context are visuospatial (as one might infer from its form) or propositional.

Procedure

On each trial, participants first solved a math factoring problem. They were then given a set of items and were asked to hold these items in memory while explaining how they solved the factoring problem. After completing their explanation, they were asked to recall the items from memory. Thus, in this study, the primary task was the explanation task, performed simultaneously with the memory task. In some dual-task procedures, the working memory task is assumed to take priority. In these studies, performance on the primary task is expected to vary with increased load on the memory task, and is therefore taken as the dependent measure. However, the assumption in these studies (as in all dual-task studies) is that the primary and memory tasks both make use of some common processing system (see Navon, 1984). It should therefore be possible to use performance on *either* the primary task *or* the memory task as an index of whether the processes involved in the two tasks interact. In our procedure, there was an immediate priority given to the explanation task—the study was advertised as being about math and the explanations were clearly more compelling than the word lists or dot arrays. As a result, performance on the primary task did not vary across the conditions of interest (i.e., gesturing vs. not gesturing during explanation; see *Performance on the Explanation Task*). Performance on the memory task did vary, however, and we take the number of items recalled as the dependent measure in all of our analyses.[4]

Participants were given 30 factoring problems to solve and explain. The problems were administered via presentation on a computer monitor that was located to the right of the participant, who was standing in front of a wall-mounted white board. The participant wrote the problem on the white board and solved it. All of the problems could be solved using whole number factors. After solving a problem, participants pressed the space bar (verbal group) or clicked the mouse (visuospatial group) to view the memory stimulus that was then presented on the computer screen. Participants saw one of two types of memory stimuli. (1) *Verbal Memory Stimulus*. Thirty-six participants saw a string of 6 consonants, grouped in pairs. Letters were presented as lowercase letters in Courier font size 60. The consonants were chosen so that they contained no obvious acronyms and were visible for 3 s. An example of a consonant group used in the experiment is `bl sj zk`. (2) *Visuospatial Memory Stimulus*. Thirty-six participants saw a configuration of four dots contained within a 5 × 5 grid. The grid pattern was visible for 2.5 s.

After the memory stimulus was displayed, the computer prompted the participants to explain their answer to the factoring problem they had just solved. The experimenter was located to the left of the participant and served as the interlocutor. After completing their explanations, the participants pushed the space bar (verbal group) or clicked the mouse (visuospatial group) on the computer and were prompted to recall the stimulus they had seen. The verbal group typed the letters on the keyboard; the visuospatial group used the mouse to locate the four dots within the grid on the screen. Note that, in both groups, participants were producing their explanations while holding the stimulus items in memory. Presentation of all stimuli was controlled using PsyScope software on a Macintosh computer (Cohen, MacWhinney, Flatt, & Provost, 1993).

Each participant explained 20 of the problems with no restrictions (the movement-permitted condition) and 10 of the problems with the restriction that they place their hands on the white board tray and remain as still as possible while explaining their answer (the no-movement condition). The order of these conditions was counterbalanced across participants. Gesture was not mentioned at all during the instructions. During the movement-permitted condition, participants were allowed to move their hands but were *not* told to gesture. During the no-movement condition, participants were instructed to remain still to increase their focus on the explanation but were not explicitly told not to gesture. The entire session was videotaped. All participants gave informed consent.

Data analysis

All trials were coded from the videotapes for the presence or absence of gesture. Gestures were defined as

---

[4] On one hand, we can view the explanation task and the memory task as both making demands on a specific working memory store. To the extent that explanation with gesture makes particular demands on a visuospatial store, memory recall from that store should be adversely affected and performance should decrease. Alternatively, one could view the explanation task as filling the retention interval for the memory task. To the extent that explanation with gesture selectively interferes with the representation of items in a visuospatial store, memory recall from that store should be adversely affected and performance should decrease. Note that each of these descriptions is compatible with our goal—to diagnose how the presence of gesture affects memory for visuospatial items.

visible hand motions without an obvious alternative purpose (e.g., fixing the hair or scratching the nose). Accordingly, any movement visible on the video that could not be readily interpreted as either a self-adapter or an instrumental act was considered a gesture. Speech and gesture from all participants were transcribed and then coded for propositional content. Reliability was established by having another person code a subset of the data. Agreement between coders was 86% ($N = 465$) for coding speech and 84% ($N = 400$) for coding gesture. Table 1 displays examples of the information commonly conveyed in speech and in gesture in the factoring explanations.

Non-standard trials were discarded from the analyses: trials when participants gestured in the no-movement condition, trials when participants changed their answer during their explanation, trials when participants dropped the pen and picked it up during the explanation, and trials when participants interacted with the experimenter during their explanation. The average number of included trials for each participant who gestured on some trials was 27.5 out of 30 (17.3 gesture trials, 10.2 no-gesture trials). Note that because some participants did not gesture on all of the movement-permitted trials, the average number of no-gesture trials was greater than the number of trials in the no-movement condition. Nine participants did not gesture on any of the movement-permitted trails (5 in the verbal group, 4 in the visuospatial group). Their data were analyzed separately.

Proportion correct was subjected to an arcsine transformation before statistical analysis; .985 was substituted for scores of 1 before transformation, as recommended by Bartlett (1947).

Results

Performance on the explanation task

As expected given their math backgrounds, the participants were successful at solving nearly all of the problems correctly. Only 2% of the problems were solved incorrectly. Thus, participants clearly understood the mathematical principles involved in the task.

Participants' explanations of the problem were very likely to include gesture. Sixty-three participants gestured on some experimental trials, 31 who were given the verbal stimulus to recall (the verbal group), and 32 who were given the visuospatial stimulus to recall (the visuospatial group). Participants produced both pointing and iconic gestures on their explanations of the math-factoring problem. Almost all (86%, 938/1078) of the explanations with gesture contained pointing gestures, grounded in the problem displayed on the board. Many (17%, 188/1078) of the explanations with gesture contained pointing gestures directed in space and not grounded in the problem displayed on the board. A smaller percentage (9%, 96/1078) of the explanations with gesture also contained iconic gestures. Fig. 1 displays one of the participants explaining his solution to the factoring problem. He put '$x - 5$' in the first set of parentheses and '$x + 2$' in the second set of parentheses to solve the problem $x^2 - 3x - 10 = ()()$. He then said, "negative 5 and positive 2 make negative 3 added together, multiplied together they make negative 10." At the same time, he produced the following string of gestures: he pointed at the 5 and then at the 2 with his palm facing the board; he pointed at the 3 with his palm facing away from the board; he produced an

Table 1
Types of information conveyed in speech and in gesture in explanations of the math factoring problems Example problem : $x^2 + 5x + 6 = (x + 3)(x + 2)$

Type of information	Speech	Gesture
Adding	"3 plus 2 is 5"	Point to 3, then 2, then 5, drop hand
Multiplying	"3 times 2 is 6"	Point to 3, then 2, then 6, drop hand
Combining terms	"3x plus 2x is 5x"	Point to 3, then to the right x, pause, point to 2, then to the left x, pause, point to 5x
Factoring	"I wanted factors of 6 like 1 and 6 or 2 and 3"	Point to 6, point to 6 again
Constrained factoring	"I wanted factors of 6 that added up to 5"	Point to 6, point to 5
x^2	"The x squared needs to be broken into x and x"	Point to the x squared, point to the left x, point to the right x
Signs	"If the 2 signs over here are both plusses, then they both have to be plusses in the parentheses"	Point to the first plus, point to the second plus, point to the first plus in parentheses, point to the second plus in parentheses
Multiplying signs	"The 6 has a plus in front of it so these signs are plus and plus which multiply out to another plus"	Point to the second plus, point to the first plus in parentheses, point to the second plus in parentheses
Adding signs	"The 5 has a plus in front of it so these signs are plus and plus which add to another plus"	Point to the first plus, point to the first plus in parentheses, point to the second plus in parentheses

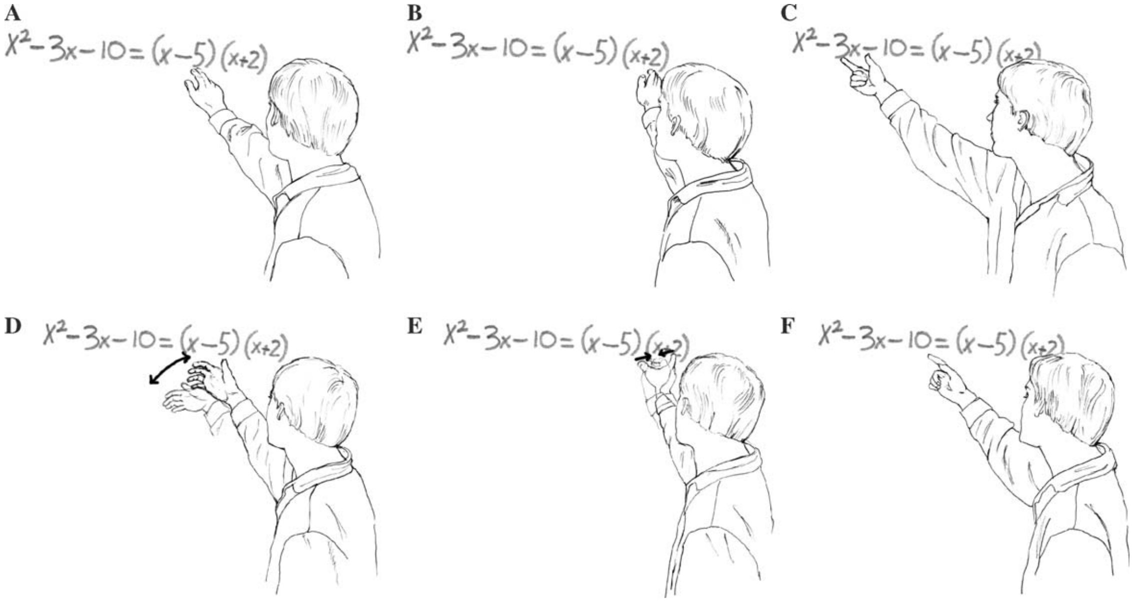

Fig. 1. An example of the gestures produced in explanations of the factoring problem. The participant indicates the two numbers $(-5, 2)$ that need to be added to give -3 and multiplied to give -10. Note that, in addition to pointing out each of the numbers, he produces an "add" gesture (4th gesture) and a "multiply" gesture (5th gesture). Reprinted from Goldin-Meadow (2003) with permission from Harvard University Press and the artist, Linda Huff.

"add-together" iconic gesture (he wiggled his palm back and forth); he produced a "multiply-together" iconic gesture (he held a "C-shaped" hand near the 5 and 2 and brought the ends of the "C" together); he pointed at the 10. The participant used pointing gestures to indicate the factors in the problem and complemented those points with iconic gestures to indicate that the two numbers are added together to give one number and multiplied together to give another. Note that the iconic gestures provide information that is not available in the supporting visual context.

Importantly, the participants' spoken explanations that were accompanied by gesture did not differ from those not accompanied by gesture. The participants in both the verbal and visuospatial groups produced the same number of different types of information in their speech when they gestured vs. when did not gesture (3.3 vs. 3.5 verbal, $t(29) = 1.76$, $p = .08$; 3.3 vs. 3.4 visuospatial $t(31) = 0.94$, n.s.). Thus, any differences observed on the memory recall task cannot be attributed to differences in the content of the spoken message on the explanation task.

Performance on the memory recall task

For the 63 participants who gestured on some trials, recall data were entered into an analysis of variance (ANOVA) with one between-subjects factor (type of memory load, verbal vs. visuospatial) and one within-subjects factor (gesture vs. no-gesture). Fig. 2 presents the data. Participants remembered significantly more verbal items than visuospatial items (70% vs. 53% correct, $F(1, 61) = 16.25$, $p < .001$). More interestingly, participants remembered significantly more items when gesturing during their explanations than when not gesturing (64% vs. 58%, $F(1, 61) = 13.07$, $p < .001$). This pattern was found in both memory groups—participants in the verbal (72% vs. 67%, $t(61) = 2.06$, $p < .05$) and the visuospatial (56% vs. 49%, $t(61) = 3.06$, $p < .01$) groups remembered significantly more items when gesturing than when not gesturing. In other words, there was no interaction between gesture and type of memory load ($F(1, 61) = .46$, n.s.).

But is gesturing truly responsible for the higher memory scores? Perhaps participants remembered less when not gesturing simply because not-gesturing is itself a load on working memory—at the least, participants had to remember not to gesture. Twenty participants (6 in the verbal group, 14 in the visuospatial group) did not gesture on at least one trial when movement was permitted, and thus provided us with data that can address this concern.[5] For these participants, we can separate the

---

[5] One participant who met this criterion was eliminated from the analysis as an outlier; his memory performance when not gesturing naturally was more than two standard deviations greater than the mean for the rest of the participants.

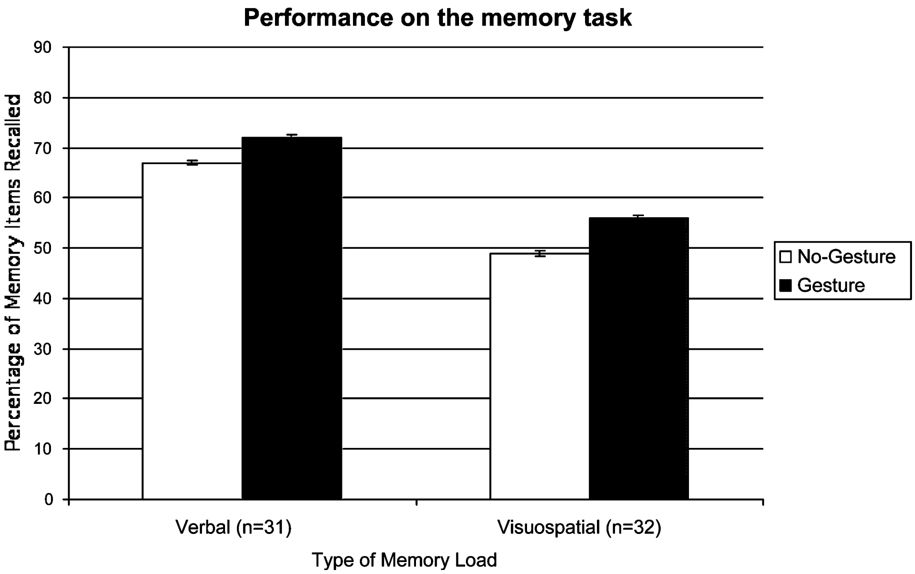

Fig. 2. Percentage of items recalled by participants who gestured during some explanations of the factoring problem. Bars represent standard errors.

effect of *naturally* not gesturing from the effect of *being instructed* not to gesture. Data were entered into an ANOVA with one between-subjects factor (type of memory load, verbal vs. visuospatial) and one within-subjects factor (natural gesture, natural no-gesture, instructed no-gesture).[6] Fig. 3 presents the data. The participants in this analysis remembered marginally more verbal items than visuospatial items (67% vs. 52% correct, $F(1,36) = 4.04$, $p = .06$). More importantly, the participants remembered significantly more items when gesturing during their explanations than when not gesturing, either naturally refraining from gesture (66% vs. 57%, $F(1,36) = 5.17$, $p = .029$) or after being instructed not to gesture (66% vs. 57%, $F(1,36) = 5.67$, $p = .023$). There was no difference in recall between natural no-gesture and instructed no-gesture (57% vs. 57%, $F(1,36) = .01$, n.s.), indicating that the instructions to remain still did not impose a systematic additional memory load on the participants. In addition, there was no interaction between gesture and type of memory load ($F(2,36) = .27$, n.s.)—participants in both the verbal and the visuospatial groups recalled more items when gesturing during their explanations than when not gesturing, either naturally or by instruction.

Recall that nine participants did not ever gesture, even when movement was permitted. These participants provide additional support for the claim that the instructions to remain still did not add to the working memory load. For these participants, we can compare performance on trials that contained instructions not to move with performance on trials that did not contain these instructions. If the instructions truly did not impose a memory load, these participants ought to have recalled the same percentage of items when they remained still on their own and when they were instructed to remain still—and, indeed, they did (56% vs. 58%, paired $t(8) = 0.55$, n.s.).

We need to eliminate one additional factor in order to be certain that the participants' improved recall performance was associated with gesture. Short-term memory is better when the retention interval is short (e.g., Shiffrin & Cook, 1978). If gesturing reduces the length of an explanation, recall might be better with gesture than without it simply because the retention interval is shorter with gesture than without it. We found, however, no evidence to support this hypothesis. Explanations that naturally included gesture were *not* significantly shorter than explanations that naturally did not include gesture (15.5 vs. 16.0 s, paired test,

---

[6] There were significantly more participants who did not gesture on all of the trials in the movement-permitted condition in the visuospatial group than in the verbal group (43% vs. 19%, $t(1,61) = 2.12$, $p = .038$). This difference appears to be due to the fact that some participants in the visuospatial group kept their hands on the computer mouse during their explanations; in the verbal group, participants responded via the keyboard and no participants kept their hands on the keyboard during their explanations.

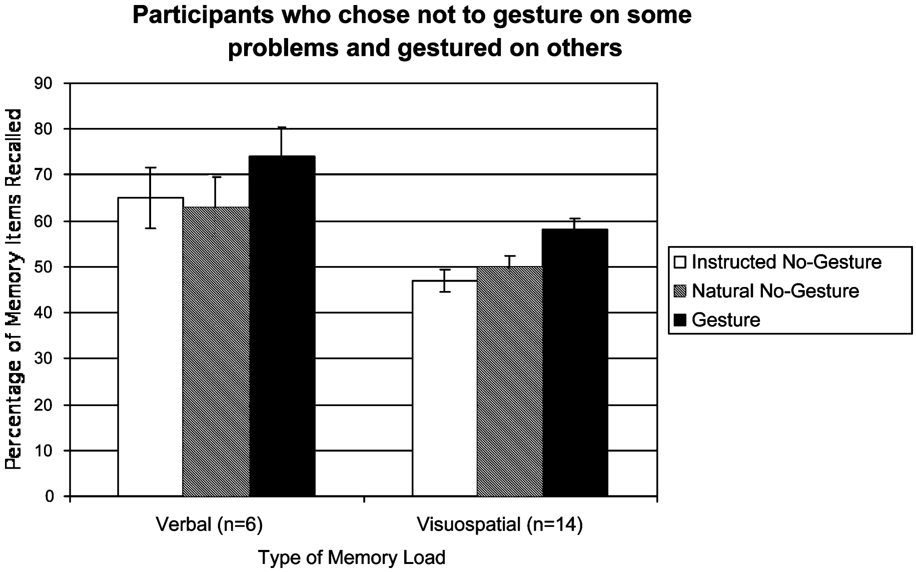

Fig. 3. Percentage of items recalled by participants who, when movement was permitted, chose not to gesture on some explanations of the factoring problems and gestured on others. Bars represent standard errors.

$t(19) = 1.09$, n.s.).[7] In addition, if we look only at those participants whose explanations were longer when they gestured than when they did not gesture, we find that these participants also remembered more when gesturing than when not gesturing (66% vs. 60%, paired test $t(16) = 2.55$, two-tailed $p = .021$). Thus, even when gesturing seems to take up time, it is still strongly associated with memory benefits.

These results confirm our previous finding that gesturing confers a significant benefit on verbal memory (Goldin-Meadow et al., 2001). Moreover, the results extend this finding by demonstrating that gesturing also confers a significant benefit on *visuospatial* memory. Since gesturing clearly did not interfere with recall on the visuospatial memory task, the mental representations underlying gesture production were not likely to be visuospatial. Are these representations better described as propositional? Our next analysis was designed to address this question.

Propositional analysis

Sometimes when the participants in our study gestured, they conveyed the same information in their gestures as they conveyed in the accompanying speech—they produced gesture–speech matches. However, other times, those very same speakers conveyed information in gesture that was different from the information conveyed in the accompanying speech (see Table 2)—they produced gesture–speech mismatches (Church & Goldin-Meadow, 1986; Goldin-Meadow, 2003). If the mere act of moving one's hand is reducing the load on working memory, it should not matter what the gestures mean (i.e., the propositional content of the gesture should be irrelevant)—and it therefore should not matter whether the speaker's gestures are matches or mismatches. If, however, it is gesture's ability to convey meaning that allows it to have an impact on working memory, matches and mismatches ought to have different effects on load.

Consider the following argument. In a gesture–speech match, gesture and speech convey the same information and therefore can reinforce one another. Indeed, gesture and speech might even share some of the representational load. However, in a gesture–speech mismatch, gesture conveys different information from speech, which means that the two cannot share the representational load. In fact, the load itself is heavier in a mismatch than in a match simply because there are two messages being conveyed—one in speech and another in gesture (cf. Nusbaum & Schwab, 1986). If the *meaning* of a gesture plays a role in its ability to reduce

---

[7] We found that we could not examine this question when participants were asked not to move. The computerized administration of the study introduced an additional step into the retention interval in the no-movement condition—placing the hand on the whiteboard tray; this step appeared to artificially increase the length of time required for no-gesture explanations. Indeed, explanations that naturally did not include gesture were significantly shorter than explanations that did not include gesture because of the experimental manipulation (15.5 vs. 18.4 s, paired test $t(19) = 2.45$, two tailed $p = .024$).

Table 2
Examples of gesture–speech mismatches in explanations of the math factoring problems Example problem : $x^2 + 5x + 6 = (x + 3)(x + 2)$

Type of information	Speech	Gesture
Speech: Adding Gesture: Combining terms	"3 plus 2 is 5"	Point to 3, then to the right x, pause, point to 2, then to the left x, pause, point to $5x$
Speech: Factoring Gesture: Constrained factoring	"I wanted factors of 6 like 1 and 6 or 2 and 3"	Point to 6, point to 5
Speech: Multiplying signs Gesture: Signs	"The 6 has a plus in front of it so these signs are plus and plus which multiply out to another plus"	Point to the first plus, point to the second plus, point to the first plus in parentheses, point to the second plus in parentheses

the load on working memory, gesture ought to ease a speaker's cognitive burden more when it matches the speech it accompanies than when it does not—that is, speakers ought to remember more letters and more dots on problems where their gestures match speech than on problems where their gestures do not match speech.

To test this hypothesis, we compared the information conveyed by gesture and speech in each explanation. We classified an explanation as a mismatch when the information conveyed by gesture (see Table 1) was *not* conveyed in the accompanying speech. Reliability was established by having a second coder categorize each explanation in a subset of the data as either a match or a mismatch. Agreement between coders was 87% ($N = 173$). Table 2 presents examples of common gesture–speech mismatches.

In general, speech and gesture tended to convey the same information—gesture–speech matches accounted for 95% of the explanations produced by the participants. However, 22 participants (12 in the verbal group, 10 in the visuospatial group) produced some explanations containing information in gesture that was not found in speech, that is, gesture–speech mismatches. For these participants, 13.6% of all explanations that included gesture were mismatches. We can therefore examine memory for both verbal and visuospatial items when these participants produced matches, compared to when they produced mismatches. Data were entered into an ANOVA with one between-subjects factor (type of memory load, verbal vs. visuospatial) and one within-subjects factor (type of gesture, match vs. mismatch). Fig. 4 presents the data. The participants in this analysis remembered significantly more verbal items than visuospatial items (67% vs. 47%, $F(1,20) = 5.18$, $p = .034$). More importantly, the participants remembered significantly more items when their gestures matched their speech than when their gestures did not match their speech (62% vs. 52%, $F(1,20) = 5.84$, $p = .025$).

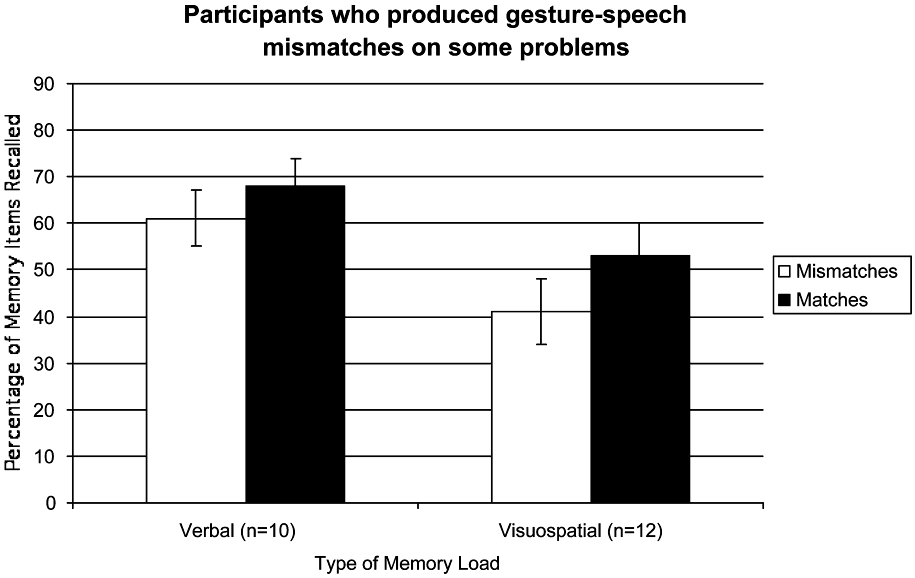

Fig. 4. Percentage of items recalled by participants who produced gesture–speech mismatches on some explanations of the factoring problem when movement was permitted. Bars represent standard errors.

Moreover, participants in both the verbal and visuospatial groups showed this effect—that is, there was no interaction ($F(1,20) = 0.58$, n.s.). As in the earlier analysis, these results cannot be explained simply by retention interval differences: Explanations that contained mismatches were no longer than those that contained matches (25.7 s vs. 28.7 s, paired test $t(21) = 1.61$, n.s.).

There is, however, another possible confound: Mismatches may have been produced on particularly difficult problems; if so, performance could have been worse on the recall items associated with these problems, not because of the multiple propositions conveyed in gesture and speech, but because of the cognitive uncertainty involved in solving these difficult problems. There was little evidence for this hypothesis. The math problems that elicited mismatches were no harder than the problems that elicited matches: 35% (19/55) of the mismatching explanations were produced on the more difficult problems containing a negative sign before the '*c*' in the equation, as were 35% (121/340) of the matching explanations. Thus, the lowered recall on items associated with mismatching explanations cannot be attributed to problem difficulty, and is likely to reflect the fact that gesture in a mismatch adds propositional information to the overall message.[8]

Discussion

The representations that underlie gesture production

Our findings demonstrate that gesturing has an impact on working memory. Moreover, it has the same impact on visuospatial working memory as it has on verbal working memory. The findings therefore do not lend support to theories in which visuospatial representation is considered the key to gesture (Alibali, Kita, & Young, 2000; De Ruiter, 2000; Hadar et al., 1998; Hadar & Butterworth, 1997; Kita, 2000; Rauscher, Krauss, & Chen, 1996). If the representations underlying gesture were visuospatial, gesturing ought to make more demands on visuospatial working memory than on verbal working memory, and ought to interfere with performance on a concurrent visuospatial task more than on a concurrent verbal task. We found, however, that gesturing facilitates recall of visuospatial items as well as verbal items. Furthermore, our results demonstrate that the propositional content of gesture matters. Participants remembered more items—both visuospatial and verbal—when gesture conveyed the same propositional information as speech (gesture–speech matches) than when gesture conveyed different information (gesture–speech mismatches).

Thus, despite the fact that gesture's format is visuospatial, the representations underlying gesture are propositional—at least when gestures convey content that is *not* inherently spatial. However, our study leaves open the possibility that when gesture conveys content that is spatial, the representations underlying those gestures will also be spatial. To explore this possibility, we can vary the primary task in our dual-task paradigm and ask participants to explain a problem that involves space (e.g., we could ask them to describe a habitual route from memory, Iverson, 1999; Iverson & Goldin-Meadow, 1997). Our study also leaves open the possibility that the visual support available in the communicative context influences the nature of the representations underlying gesture. The factoring problem in this study was visible throughout the explanation and participants could—and did—refer to it with their gestures. Indeed, nearly all of the explanations that the participants produced contained some pointing gestures referring directly to the numbers on the board and 57% (613/1078) contained only this type of grounded gesture. It is clear from our findings that visuospatial representations do not underlie gestures that are grounded in this way. However, a visuospatial representation might underlie gestures that are less grounded in the available visual context simply because the speaker may need to construct a visual mental image in order to produce such an ungrounded gesture. To explore this possibility, we can vary the props,

---

[8] It is important to note that the mismatches produced by the adults on this task are different from the mismatches that children produce when they are in the process of learning a task. When children who are in transition produce a gesture–speech mismatch, the information conveyed in gesture is typically not conveyed in speech *anywhere* in that child's explanations of the task (see, for example, children learning about mathematical equivalence in Goldin-Meadow & Singer, 2003). In contrast, the information conveyed in gesture in the mismatches produced by the participants in our study could always be found in their speech on another explanation (as was the case for the mismatches produced by the adult teachers in Goldin-Meadow & Singer, 2003). For children, mismatches are an index of cognitive instability. In contrast, for the adults in this study (and for the adult teachers in the Goldin-Meadow and Singer study), mismatches are an index of discourse instability—a moment when speech and gesture are not completely aligned, reflecting the dynamic tension of the speaking process (McNeill, 1992; see also Goldin-Meadow, 2003). Note that this difference is not a developmental difference but rather reflects the state of the speaker's knowledge—adults will produce mismatches comparable to child mismatches when they are in the process of *learning* a task (e.g., Perry & Elder, 1997). One final point deserves mention—gesture confers cognitive benefits on working memory whether a speaker is proficient in a task (as in the present study) or has not yet mastered the task (e.g., the children in Goldin-Meadow et al., 2001).

displays, and materials that are physically present during the explanations in the primary task (e.g., we could take away the display of the math problem and then ask participants to explain their solution). Whatever the outcomes of future studies that vary the parameters of the primary task, the results of the present study make it clear that visuospatial representations do not underlie all gestures (and, of course, they leave open the possibility that visuospatial representations do not underlie *any* gestures).

This finding is particularly striking because limb movements that are not communicative seem to involve processing that depends on visuospatial working memory. For example, producing meaningless, target-directed arm movements interferes with visuospatial working memory (remembering the location of circles in a 4×4 grid) more than verbal working memory (remembering consonants) (Lawrence, Myerson, Oonk, & Abrams, 2001). Moreover, body movement has detrimental xeffects on spatial memory tasks (Quinn & Ralston, 1986; Smyth & Pendleton, 1989, 1990). Meaningless arm movements appear to be visuospatially encoded, but at least some gestures are not—memory for visual items located on a grid is disrupted by the production of non-meaningful arm movements but enhanced by the production of gestures. Gestures are thus cognitively distinct from non-symbolic, non-communicative limb movements.

However, gesture is a motor act and therefore must involve motor representations. Do these motor representations also have an effect on working memory? To find out, we can vary the memory recall task in our dual-task paradigm, this time asking participants to remember different handshape configurations that must be produced during recall after the explanation has been completed. If the motor representations that underlie gesture have an impact on working memory, participants should remember fewer items on this task when gesturing than when not gesturing.

Note, however, that the effect gesture had on working memory in our study cannot have been a pure motor phenomenon—it must stem instead from the coordination of motor activity and higher order conceptual processes. If the motor aspects of gesture were solely responsible for the cognitive benefits associated with gesture production, mismatching gestures should have been as effective in promoting recall as matching gestures—after all, mismatching gestures are motor behaviors that are physically comparable to matching gestures. Rather, it appears to be the propositional properties of gesture that allow it to influence working memory, thus lending support to models of gesture production that consider gesturing to be a conceptual process (e.g., Alibali & Goldin-Meadow, 1993; Goldin-Meadow, 2003; Goldin-Meadow, Alibali, & Church, 1993; McNeill, 1992).

The cognitive benefits of gesture

Krauss and his colleagues (Krauss, Dushay, Chen, & Rauscher, 1995; Rauscher et al., 1996; see also Butterworth & Hadar, 1989; Hadar et al., 1998) have suggested that gesturing facilitates the retrieval of lexical items from memory and, in this sense, confers a cognitive benefit on the speaker. We have, indeed, found that gesturing is good for speakers. However, there are aspects of our results that are not consistent with the lexical retrieval hypothesis. In particular, the lack of condition-order (gesture condition first vs. no-gesture condition first) effects in our data is problematic for the hypothesis. Lexical retrieval demands should diminish over the course of the study since the problem types remain the same throughout the study and the participants use and then reuse a relatively small set of lexical items in their explanations. If gesture's sole function is to improve lexical access, its beneficial effects should be evident at the beginning of the study and should diminish over time. Participants who gestured during the second phase of the experiment should therefore not show the beneficial effects of gesture since they will have already repeatedly accessed the lexical items used in their explanations. This prediction was not confirmed. In both the verbal and the visuospatial groups, participants who completed the movement-permitted trials in the second half of the experiment showed the same pattern of results as everyone else—their recall was better when they gestured than when they did not gesture (verbal group: 69% vs. 63%, paired test $t(13) = 2.16$, two-tailed $p = .05$; visuospatial group: 60% vs. 44%, paired test $t(11) = 4.01$, two-tailed $p < .01$).[9]

How then can we account for the beneficial effects that gesturing appears to have on working memory? Our results suggest that gesture works along with speech to affect working memory. Yet, unlike speech, gesture does not form a linguistic system (McNeill, 1992). Perhaps gesture has its impact on working memory precisely because it is not language-like in form.

Gesture is qualitatively different from language, even from sign language which is produced in the same manual modality. Despite the fact that sign languages share both spatial and motor properties with gesture, sign languages structure information differently from gesture, perhaps because these spatial and motor

---

[9] It is important to note that the lexical retrieval models hypothesized in the literature pertain to iconic gestures, and only 9% of the explanations with gesture in our study contained iconic (as opposed to pointing) gestures. In order to convincingly argue against this model, our study should be done again, this time using an explanation task that elicits primarily iconic gestures (e.g., Piagetian conservation tasks, Church & Goldin-Meadow, 1986, or Kohlberg's moral dilemma tasks, Church, Schonert-Reichl, Goodman, Kelly, & Ayman-Nolley, 1995).

properties are not essential to the production and comprehension of sign. For example, Hickok, Bellugi, and Klima (1996) report a dissociation between deficits in the ability to produce and comprehend sign language (associated with left hemisphere lesions) and deficits in general spatial abilities (associated with right hemisphere lesions). The ability to produce and comprehend sign language has also been dissociated from measures of general motor control (Corina et al., 1992; Hickok et al., 1996). Although sign language uses spatial properties to represent information, this information is compositionally dense and, as such, is distinct from information represented in gesture, which is compositionally less dense. As McNeill (1992) has argued, one crucial way in which gesture can complement language is through its global and holistic properties. Language, be it signed or spoken, conveys information (even visual and spatial information) in a segmented fashion, whereas gesture conveys information holistically. As a result, the representations that underlie gesture are likely to be more global than those underlying either sign or speech.

We suggest that the global representations underlying gesture can have an impact on the representations that underlie speech. If so, the mental representations underlying thinking-for-speaking will be different when speakers gesture than when they do not gesture. Because of its global properties, gesture can provide an overarching framework that serves to organize the propositions in speech, in effect chunking mental representations to reduce the load on working memory. Gesture may bring a different kind of mental coherence to the representation of an intended message, one that increases the efficiency of representation compared to speech without gesture. Our findings on gesture–speech mismatch are consistent with this view. Gesture–speech mismatches were less effective in facilitating working memory than gesture–speech mismatches. In a mismatch, gesture conveys a message that is not easily integrated with the message in speech. In this case, it is difficult for gesture to provide an overarching framework for the speech it accompanies. In contrast, in a match, gesture conveys the same types of information as speech (albeit from a different perspective). It is therefore able to provide a framework that complements and organizes speech and thus lightens the burden on working memory.

Our hypothesis is reminiscent of Kita's (2000) claim that gesture helps speakers organize spatial and motor information into packages appropriate for speaking. However, the hypothesis is distinct in that, under our view, the cognitive benefits of gesture production stem from its propositional organization rather than its visuospatial form. And, as we have shown here, gesture is indeed associated with cognitive benefits even when the information it represents is not visuospatial.

In sum, although gesture uses visuospatial properties to represent information, these visuospatial properties do not appear central to the mental representation that underlies gesture. In contrast to simple handwaving, the demands that gesture makes on working memory appear to be propositional rather than visuospatial.

Acknowledgments

This paper benefited from the input of a variety of people. The authors thank the reviewers for their remarkably insightful comments on an earlier version. These comments greatly improved the final version. We also thank Zachary Johnson for invaluable help with data collection, transcription and coding. We thank the many members of the Goldin-Meadow lab for comments on earlier versions as well as helpful discussions related to the ideas in this paper.

References

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

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

Anderson, J., Budiu, R., & Reder, L. (2001). A theory of sentence memory as part of a general theory of memory. *Journal of Memory and Language, 45*, 337–367.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Bartlett, M. (1947). The use of transformations. *Biometrics, 3*, 29–52.

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology, 22*(5), 349–368.

Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review, 96*, 168–174.

Carpenter, P., & Just, M. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review, 82*, 45–73.

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

Church, R. B., Schonert-Reichl, K., Goodman, N., Kelly, S. D., & Ayman-Nolley, S. (1995). The role of gesture and speech communication as reflections of cognitive understanding. *Journal of Contemporary Legal Issues, 6*, 123–154.

Clark, H., & Chase, W. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472–517.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257–271.

Corina, D., Poizner, H., Bellugi, U., Feinberg, T., Dowd, D., & O'Grady-Batch, L. (1992). Dissociation between linguistic and nonlinguistic gestural systems: A case for compositionality. *Brain and Language, 43*, 414–447.

De Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284–311). Cambridge: Cambridge University Press.

Emmorey, K., Tversky, B., & Taylor, H. A. (2000). Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation, 2*, 157–180.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

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

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

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

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory–load interference in syntactic processing. *Psychological Science, 13*, 425–430.

Hadar, U., Burstein, A., Krauss, R., & Soroker, N. (1998). Ideational gestures and speech in brain-damaged subjects. *Language and Cognitive Processes, 13*, 59–76.

Hadar, U., & Butterworth, B. (1997). Iconic gestures, imagery, and word retrieval in speech. *Semiotica, 115*, 147–172.

Hadar, U., & Krauss, R. (1999). Iconic gestures: The grammatical categories of lexical affiliates. *Journal of Neurolinguistics, 12*, 1–12.

Hickok, G., Bellugi, U., & Klima, E. (1996). The neurobiology of sign language and its implications for the neural basis of language. *Nature, 381*, 699–702.

Hostetter, A., & Hopkins, W. (2002). The effect of thought structure on the production of lexical movements. *Brain and Language, 82*, 22–29.

Iverson, J. M. (1999). How to get to the cafeteria: Gesture and speech in blind and sighted children's spatial descriptions. *Developmental Psychology, 35*, 1132–1142.

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

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.

Krauss, R., Dushay, R., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology, 31*, 533–552.

Lavergne, J., & Kimura, D. (1987). Hand movement asymmetry during speech: No effect of speaking topic. *Neuropsychologia, 25*, 689–693.

Lawrence, B., Myerson, J., Oonk, H., & Abrams, R. (2001). The effects of eye and limb movements on working memory. *Memory, 9*, 433–444.

Logan, G. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 189–207.

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

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53–81.

Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review, 91*, 216–234.

Nusbaum, H. C., & Morin, T. M. (1992). Paying attention to differences among talkers. In Y. Tohkura, Y. Sagisaka, & E. Vatikiotis-Bateson (Eds.), *Speech perception, production, and linguistic structure* (pp. 113–134). Tokyo: Ohmasha Publishing.

Nusbaum, H. C., & Schwab, E. C. (1986). The role of attention and active processing in speech perception. In E. C. Schwab & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Volume 1, Speech perception*. New York: Academic Press.

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

Quinn, J., & Ralston, G. (1986). Movement and attention in visual working memory. *The Quarterly Journal of Experimental Psychology A, 38*, 689–703.

Rauscher, F., Krauss, R., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*, 226–231.

Rimé, B., Schiaratura, L., Hupet, M., & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion, 8*, 311–325.

Shiffrin, R., & Cook, J. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior, 17*, 189–218.

Smyth, M., & Pendleton, L. (1989). Working memory for movements. *Quarterly Journal of Experimental Psychology A, 41*, 235–250.

Smyth, M., & Pendleton, L. (1990). Space and movement in working memory. *Quarterly Journal of Experimental Psychology A, 42*, 291–304.

Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale: LEA.