Research on Language and Social Interaction, 27(3), 201-221
Copyright © 1994, Lawrence Erlbaum Associates, Inc.

Gestures as Part of Speech: Methodological Implications

Janet Beavin Bavelas
Department of Psychology
University of Victoria

It seems self-evident that, if we want to obtain or evaluate evidence about the communicative value of conversational gestures, we will first have to clarify what we mean by *gesture* and by *communication* (e.g., Feyereisen & de Lannoy, 1991). Neither is simply an observable event in the world: Even the initial decision to call a hand movement a gesture, rather than a random movement or an adaptor (e.g., wiping one's brow; cf. Ekman & Friesen, 1969) is not based on the physical action but on an inference about its meaning and function. Similarly, we cannot see communication; we make an inference, based on a variety of criteria, about whether it has occurred. Given the wide variety of ways in which these inferences can be made, we cannot test directly whether gestures communicate; we can only test particular models of how gestures communicate.

I am indebted to the Social Sciences and Humanities Research Council of Canada for programmatic support that permits us to study gestures and other phenomena of face-to-face dialogue. This article grew out of conversations with many colleagues, including especially my collaborators in the Victoria Group, all cited herein. Robin Routledge suggested to me that gestures may exist only in short-term memory. Discussions with Herb Clark, Adam Kendon, Robert Krauss, and David McNeill have led to further explication of my theoretical assumptions.

Correspondence concerning this article should be sent to Janet Beavin Bavelas, Department of Psychology, University of Victoria, Box 3050, Victoria, BC V8W 3P5, Canada.

Two corollaries follow immediately: (a) Unless the model is clearly articulated, any test has little meaning; and (b) a test that goes closer to the heart of a particular model is more informative than one that stays on a consensual surface. In this article, I present one such model, with implications for how it can (and cannot) be tested empirically. My approach owes a great deal both to the work of the other contributors in this issue (Kendon, Streeck, McNeill, et al.) and to skeptics about a communicative view (e.g., Krauss, Morrel-Samuels, & Colasante, 1991; Rimé & Schiaratura, 1991). Wherever the reader stands on the communicative value of gestures, I hope that this article clarifies the number of important choices that must be articulated for an empirical test of any theory of gestures to be truly informative.

I propose that one of the main functions of a speaker's gesture is linguistic, that is, to help convey meaning to the addressee in an immediate conversational context.[1] Thus, I stress *functions* rather than classification, *meaning* rather than physical movement, and the absolute importance of both the *addressee* and the moment-by-moment *context* in which the gesture occurs. This view of gestures is a microanalytic one, in which each gesture is typically made, decoded, acted upon, and forgotten in a matter of seconds, usually without ever entering long-term memory or metalinguistic awareness. In all of these respects, gestures are very much like words or phrases in spontaneous conversation. Indeed, my main criterion for any tests of the communicative value of gestures is that spoken words could pass such tests as well.

A FUNCTIONAL VIEW OF GESTURES

We need to move away from taxonomic approaches, with their apparent goal of classifying gestures and focus instead on function: *What does this gesture do and how*? Although it is often convenient to talk of different kinds of gestures, these labels should always be recognized as shorthand terms for more precise descriptions that include a verb. Taxonomic description can be retrogressive in the study of language, including gestures. To begin with, it suggests that we look for intrinsic properties of a gesture rather than for what that gesture is doing in its particular moment in the conversation.

The following excerpt contains two examples: Students in our courses are given volunteer forms, which they fill out if they agree to be called for experiments. These two had each agreed to come to be videotaped while getting acquainted with a stranger. They began on the natural topic of how they came to be here:

A: I'm taking a couple of, ah, Psych courses, "Drugs and Behaviour," so that's how they got my name.
[*"writes" in air*]

B: That's how they got yours, eh? (pause) I'm in second 'n third year Psych so, yeah, they got my name—
ACTUALLY, this is the first one, I've actually been in,
[*swings his arm up towards A, as if throwing something to him*]

A: Uh huh.

The function of Speaker A's gesture is obvious; his stylized writing clarifies *that's* by adding a gestural clause (*because I filled out the volunteer form in class*). Note that there was no verbal reference to the form, yet when (a few minutes later) Speaker A referred to the information requested on the pink forms, Speaker B responded without asking for clarification. The function of Speaker B's gesture is not so obvious without seeing it; it was timed with the interruption of his sentence to mark the new information he was interjecting, namely, that in spite of several psychology courses, he had *actually* not been in an experiment before. This gesture conveys the equivalent of a rhetorical *Y'know what?*

As the second example illustrates, gestures are not limited to global depiction of concrete information; they can serve many different and subtle functions in conversation. Gestures that serve to convey information about the topic may have other functions as well. McNeill (1992, pp. 170–171) described *superimposed beats,* in which a gesture depicting an action or event is immediately repeated (i.e., appears to be made twice in quick succession); McNeill proposed that the function of the repeated gesture is to underline the point being made. Kendon (personal communication, December 8, 1988) suggested that, when a gesture is held longer than would be needed simply to convey information, it becomes a *kinesically held question,* that is, a request for a response from the

addressee. We have also seen these extended gestures and are currently examining their relationship to listener responses (back-channels).

Thus, not only can gestures serve many different functions, but a gesture can have more than one function at once (i.e., a repeated or extended gesture not only depicts information but also conveys emphasis or seeks a response). The same is true for words; for example, when *but* functions as a conjunction, it also conveys a meaning different from other possible conjunctions. (In poetry, words serve aesthetic functions as well, including rhythm and alliteration.) Taxonomic categories imply mutually exclusive classifications, whereas functions need not be exclusive. For example, instead of classifying a gesture as a conduit metaphor (McNeill & Levy, 1982) or an interactive gesture aimed at the addressee (discussed later in this article), we can note that the gesture was encoded as a conduit metaphor (e.g., cupping the hand for a word) and also served the interactive function of involving the addressee.

One major implication of accepting multiple functions is that evidence for one function is not automatically evidence against another. In particular, lexical access and retrieval theories (e.g., Morrel-Samuels & Krauss, 1992) are not inconsistent with communicative theories of gestures. If at a point of verbal disfluency, the speaker makes a gesture instead of a word, the gesture may help him or her find the word. But it may also convey the speaker's meaning to the addressee in the absence of the desired word, or it may prompt the addressee to guess at the missing word (as do proxy verbal phrases, such as *What's his name, the guy we talked to yesterday . . . ?*). Evidence for each of these functions could be sought in the behaviors of the speaker (e.g., gaze direction, attempts at speech, and success in word-finding), as well as in the response of the addressee (e.g., supplying or attempting to supply the word, or continuing in a manner that indicates that the gesture was understood without words). In short, the goal of analysis should not be to decide in which category we should put a gesture (or all gestures) but rather to discover at least some of the things a gesture is doing at its particular moment in the conversation.

A GESTURE CONVEYS MEANING IN CONTEXT

Our model treats gestures at the level of meaning rather than of physical movement. Gestures are "manual symbols" (McNeill, 1985, p. 351), just as words are graphic or auditory symbols. I could convey the size of a fish I caught a long time ago by: (a) saying the word *tiny,* (b) writing the words *very small,* or (c) gesturing a size between my hands. In none of these cases is the fish actually there; its size is being abstracted and conveyed symbolically. Rather than dividing these into two verbal acts and one nonverbal act, we could as easily characterize them as one audible and two visible linguistic acts.

At the level of meaning, we do not describe the physical characteristics of the airwave that produces the spoken word (or count the number of vertical strokes or round characters in a written English word); we are relatively indifferent to the mechanics of how the meaning is achieved. Yet so-called nonverbal acts such as gestures have remained chained to their physical source and properties (e.g., size or movement) and often are insufficiently distinguished from nonsymbolic physical movements. In my view, the notion of nonverbal communication as a separate "channel" from verbal communication is based on a deep confusion of physical source with linguistic function. It makes no more sense to suggest that the linguistic function of a gesture is determined by its physical manifestation than to suggest that the function of a word is determined by the letter it begins with or the phonemes it contains. To speak of the overall amount of hand movement is like speaking of the overall amount of noise made by the speaker, which would include coughs, breathing, chewing sounds, and other adaptors as well as words. That is, any counting and aggregation must first be based on identification of the meaning of individual gestures. Nonsymbolic hand movements are outside the domain of the theory proposed here.

In brief, I propose that conversational gestures are truly part of speech in two senses: They contribute to meaning just as words and phrases do; and, like their lexical counterparts, their meaning depends upon the whole of which they are a part. An analogy is the word *ground,* which has 51 different meanings, literal and metaphorical, in the *Random House Unabridged Dictionary* (1993) yet it is unambiguous in phrases such as *to fall to the ground, on shaky ground, an electrical ground, to break ground, gain (or lose) ground, common ground,* and *ground ball.* In each case, the meaning of *ground* depends on its immediate context, but the word also makes a unique contribution to the phrase in which it occurs. Similarly, gestures both shape and are shaped by their immediate linguistic context (e.g., words, intonation, facial displays, etc.). They are, in Slama-Cazacu's (1976) term, part of

a *mixed syntax* and may occur before a word, with no words, with a general or specific word, with proxy words, or even with the wrong words. The meaning will usually be unambiguous to the addressee in the context of what else is being said. In Speaker A's writing gesture, as shown previously, the gesture acted as a clause that could have been put into words but instead appeared in exact parallel to the words it clarified.

Treating gestures as part of speech has unexpected methodological implications. For example, Krauss et al. (1991) tested whether gestures communicate in the sense of being independent of speech. The viewers who saw videotapes without sound were less accurate about what was being said than those who had both sound and picture. It should be apparent that the theory proposed here, although different in every respect, must also predict low levels of accuracy for the no-sound condition. It is not logically possible for a theory that stresses the complete integration of verbal and nonverbal elements to predict that conversational gestures can convey accurate information on their own; this would be like saying that it is possible to understand a written text by seeing just the adverbs.

SPEAKERS GESTURE FOR ADDRESSEES

Perhaps the most neglected aspect of any theory of gestural communication is serious consideration of the recipient of this communication. A conversational gesture may not be a "to whom it may concern" message (as is a hitchhiking gesture). Rather, like words, it may be tailored for a particular addressee in a particular conversation.

Here I draw on the important work of Clark and his colleagues (especially, Clark & Wilkes-Gibbs, 1986; Schober & Clark, 1989) on collaboration in (verbal) definite references, which I believe can be extrapolated to gestures as well. These authors pointed out that a traditional or *autonomous* view of speakers gives them complete control and responsibility for their references, with their addressees in the role of anonymous, passive recipients, if indeed they are mentioned at all. In this view, references would be relatively static and generic. Clark and Wilkes-Gibbs (1986) found evidence for an alternative, *collaborative* model by showing how speakers and addressees (who could not see each

other) worked together to create mutually effective verbal references for odd geometric figures. Evidence of this process of collaboration was that, as the dyad continued to talk about the same figures, they took significantly fewer turns and significantly fewer words to identify a figure for each other. That is, they progressed from elaborate discussions and descriptions to simple, unique terms they had worked out together (e.g., *the angel, the ice skater*).

Next, Schober and Clark (1989) demonstrated that the products of these collaborations are not generic. They showed experimentally that the original addressees were significantly superior to third parties (whom they called *overhearers*) in their understanding of the speaker's references. This was true even when the overhearers had been present for the entire conversation (i.e., they heard exactly what the addressee heard). The references between speaker and addressee were tailored by and for themselves and not for others. In the transitory, ad hoc, and creative way in which conversations proceed, the speaker's words had to work for the addressee, but only for the addressee.

In these studies, gestures were not studied because the participants spoke through a partition and the overhearers either listened through another partition or to an audiotape. Still, the findings could hardly fail to be relevant to gestures in face-to-face dialogue. If speakers tailor their words with and for addressees, they must similarly tailor their conversational gestures. For example, in the same getting-acquainted conversation described before, the two students soon discovered a common interest in law enforcement. Speaker A knew quite a bit about the Royal Canadian Mounted Police (RCMP), and Speaker B had some questions about whether they had an administrative office or only a police detachment in the Greater Victoria area:

B: Where is the RCMP? Is-the only RCMP in the area
is-what? Colwood or?
[*draws jagged circles, as if on a map*]

A: Well that's, that's a, a regular detachment in Colwood,
[*makes tighter circle*]

Speaker B states his question gesturally as well as verbally, by exploring with his circling gesture a hypothetical area that may or may not contain the office in which he is interested. Speaker A picks up and modifies the

same gesture into a firm statement about the nature of the Colwood office. He could have easily made a different gesture (e.g., pointing his thumb over his shoulder at the true geographic location of the Colwood detachment), but he built on Speaker B's gesture instead.

Anyone who analyzes videotape can see evidence that the participants were understanding and responding to each other's gestures *in real time,* whereas outside observers seldom do. Instead, we usually require repeated viewings, both of the preceding conversation and especially the sequence including the gesture, to comprehend and appreciate its precise meaning. (The video viewer also has the substantial disadvantage of a two-dimensional, smaller-than-life view, even on the best equipment.) It follows that there will be a loss of accuracy for anyone other than the participants; decoding by third parties will inevitably underestimate the communicative effectiveness of gestures. When testing whether gestures communicate to receivers, the most appropriate receiver will always be the addressee. Overhearers, even with the possibility of repeated viewings, are second best.

The same principle also has implications for visual-availability designs, which manipulate whether the speaker can see the addressee. The interpretation of such studies must depend on whom the speaker was addressing, including whether the speaker knew that he or she was being videotaped. In the classic studies by Cohen and Harrison (1973) and Cohen (1977), speakers who were alone and did not know they were being videotaped made significantly fewer gestures than those with a receiver. However, in our experiments both with single subjects and with dyads who were visually separated by a partition (Bavelas, Chovil, Lawrie, & Wade, 1992; see following), the speakers knew they were being videotaped and made the same number (or even more) gestures depicting topical aspects of their stories than did dyads interacting face-to-face. We think that these gestures were both self-prompting *and* addressed to the cameras and the experimenter (who was watching in the adjacent room). To be informative, visual-availability designs will have to take into account all possible addressees and the degree to which the speaker is aware of them and therefore may be tailoring his or her gestures for them.

Ultimately, the best such designs will probably be micro-experiments, focused on individual gestures rather than counting the overall number. There are some precedents in the study of facial displays; for example, Kraut and Johnson's (1979) observational study

of whether bowlers smiled when facing the pins versus when facing their companions; Bavelas, Black, Lemery, and Mullett's (1986) manipulation of momentary eye contact; or Jones and Raag's (1989) demonstration that small children turn to face their mothers before smiling.

GESTURES ARE EPHEMERAL

The stylized hand signals called *emblems* (Ekman & Friesen, 1969) are often used in non-speaking settings (e.g., sawmills; Meissner & Philpott, 1975), where they have relatively invariant forms and meanings. In contrast, conversational gestures are spontaneous and transient; each is invented anew and may never be used again. The same is true for the verbal parts of spontaneous speech. Except when we talk for a living (as professors often do), we seldom use exactly the same words in the same way. Instead, we improvise from wherever we are in the conversation and in the sentence, taking into account the addressee's previous contributions and current understanding. As Schober and Clark (1989) pointed out for verbal references, this "is really an opportunistic process. It succeeds in part by exploiting adventitious commonalities between speakers and addresses" (p. 229). (Recall the arbitrary use of a circle to locate the RCMP office.)

In the case of gestures, hand position and arm position are additional constraining factors. If the arms are crossed, the speaker may gesture only with the fingers; if he or she has just finished an action that left one hand in the air, an immediately following gesture will take off from that base position. Thus, gestures, too, are opportunistic: The participants improvise from what they are given and where they are at that moment. And, far from repeating what worked once in a future conversation, they may never do it exactly that way again. Gestures are elegant ephemera, more minutely improvisational than jazz.

We have argued (Bavelas & Coates, 1992) that the rapid production of such precise and complex communicative acts (often in less than simple reaction time) presents problems for cognitive theories of language. There are also some methodological implications that go beyond the technical impossibility of analyzing it in real time.

First, it seems unlikely that speakers or addressees retain any

conscious long-term memory of actual communicative acts; when conversing spontaneously, there is no rewind button. For example, many of us have had the experience of talking with a lip-reading deaf person, who at some point asked us to repeat what he had said—meaning exactly what we just said; usually, we can only offer a close summary or paraphrase. It is likely that, in any spontaneous conversation, the exact words and gestures themselves were only in short-term memory (so that we can, for example, catch production errors if we do so fairly quickly). What is remembered after a few seconds is the gist of what was said (or intended!) or, in the case of the addressee, what was understood. Both are recodings. Thus, speakers are unlikely to be able to reproduce or even recognize their gestures any more than their precise words or phrases, and addresses are likely to remember the understood meaning rather than the means by which it came about. There is evidence that, at least for written or acoustical materials, readers or listeners retain meaning rather than explicit details of surface structure (e.g., Bransford & Franks, 1971; Sachs, 1974), although indirect tests that do not require conscious recognition or production may reveal some memory for these details (Masson, 1984).

Rod Coughlan conducted a pilot study at the University of Victoria to test whether gestural information is remembered as such. He selected videotaped excerpts from a stand-up comedy routine in which key parts of the story were conveyed only gesturally, with no verbal counterpart. (Note that the use of a professional comedy monologue made the viewers appropriate addressees rather than overhearers.) The vast majority got the joke. When he asked subjects to retell or recall what they had seen, they put it into words. They sometimes made the gesture as well; but they never made the gesture without the words. It seems that their memories did not reflect what they had seen but rather their recoding of it into verbal form. Thus, overt memory is an unlikely medium for the study of conversational gestures.

For similar reasons, metalinguistic awareness by either speaker or addressee seems equally unpromising. Again, a verbal example will illustrate. An uneducated native speaker can produce syntactically and semantically correct sentences without being able to articulate (metalinguistically) how he or she is doing so. We are all uneducated gesturally; only future research will give us metalinguistic awareness of what we do and how we do it so well.

ILLUSTRATIVE METHODS

In this section, I use our program of research on interactive gestures (Bavelas et al., 1992; Bavelas, Chovil, Coates, & Roe, in press) to exemplify experimental and non-experimental methods consistent with the theory outlined previously. This research began after we had completed two long-term projects on motor mimicry (Bavelas et al., 1986; Bavelas, Black, Chovil, Lemery, & Mullett, 1988) and equivocal communication (Bavelas, Black, Chovil, & Mullett, 1990), both of which had ended up contradicting widely held views of nonverbal communication as a separate, less well-controlled "channel" dedicated to emotional expression. We were therefore attracted by the more linguistic view proposed by Kendon (1980, 1987) and McNeill (1985) and decided to examine both facial displays (Chovil, 1989; 1991/92) and gestures (Lawrie, 1988) in closer detail. It was important to limit our data to dialogues between two participants (rather than interviewer and subject, or solitary individuals) and to very good video recordings. The latter included a pseudo-3D split-screen (i.e., front views of each participant plus a wide-angle side view of both) and industrial VCRs, with frame-by-frame slow motion or jog/shuttle. By luck, our first VCRs kept the sound on during slow motion so that, although distorted, it was recognizable and never separated from the accompanying gesture.

To begin with, we simply looked at every gesture made and tried to understand its meaning and function. Most were both obvious and elegant; seeing the precise timing and metaphors of hand gestures was even more captivating than reading about them. There was a smaller group of gestures, however, that were puzzling: At times, the gesture had nothing to do with the topic of conversation but was aimed directly at the addressee. At first, these looked a great deal like the simple emphasizing gestures variously described as batons (Efron, 1972; Ekman & Friesen, 1972), punctuating movements (Freedman, 1972), or beats (McNeill, 1985); see also Kendon (1987) and Rimé and Schiaratura (1991). However, our side view of the interlocutors revealed that, although the speaker was often stressing a point, he or she was doing so directly at the addressee (e.g., Speaker B's gesture with *ACTUALLY, this is the first one*). The general function of these gestures seemed to be to include the addressee in the dialogue; a list of specific functions we

have seen is given in Table 1. In addition, their physical form consisted of an iconic reference to the interlocutor, usually a rapid and simple one (e.g., a quick flick of a finger or a brief display of the palms). Because these gestures often occurred with no corresponding verbal reference to the addressee, we proposed that they are a highly efficient means for the speaker to include and involve the addressee without yielding the floor. In different terms, they are one of the means available to help the interlocutors coordinate their contributions to the conversation (Clark & Schaefer, 1989).

We wanted to test our distinction between topic gestures (e.g., Speaker A's writing and Speaker B's circling gestures) and interactive gestures (e.g., Speakers B's marking the delivery of new information for the interlocutor), as well as our view of how interactive gestures functioned in dialogue. To do this, we had to solve three methodological problems.

Analyzing Gestures for Meaning

Our interpretation of gestures as interactive or topical in function depended on our interpretation of the meaning of the gesture at the particular moment it occurred. That is, we were translating or explicating the gesture's meaning. This meant taking into account both its physical encoding and what it seemed to be conveying in conjunction with the accompanying words, intonation, and facial displays—all in the context of where the dyad was, conversationally, at that precise moment. By analogy, suppose someone used the word *ground* in the conversation; we would know the meaning of any particular use by what they were talking about, how it was said, and the words and other acts with which it occurred.

Our commitment to the meaning of the gestural act would seem to remove us from "objective" description into the domain of "subjective" interpretation, whereas our equally strong commitment to reliable measurement would seem to condemn us to the study of physical movements. Yet objectivity is merely intersubjective agreement; there are no inherently objective facts, only those on which observers' perceptions (always subjective) are obviously likely to be aligned. The question becomes whether we could achieve high levels of independent agreement for our interpretations of gestures. If so, they would become

TABLE 1
Proposed Functions of Interactive Gestures

CONVERSATIONAL GESTURES accompany and illustrate talk and are improvised with and synchronized to words. They are usually made by the person speaking at the moment.

- TOPIC GESTURES depict some aspect of the topical content of the conversation; for example, the size of an object or (metaphorically) of a problem. The vast majority of conversational gestures are topic gestures.
- INTERACTIVE GESTURES are a much smaller group that refer to the addressee and provide no information about the topic-at-hand. They serve several functions necessary for dialogue:
  - *Delivery* gestures, as a group, refer to the delivery of information by speaker to addressee:
    - *General* delivery gestures mark the standard relation of speaker to addressee; the speaker "hands over" to the addressee new information relevant to his or her main point. A verbal paraphrase is often "Here's my point."
    - *Shared* information gestures mark material that the addressee probably already knows—information that is part of their common ground (Clark & Brennan, 1991). They mean, essentially, "As you know."
    - *Digression* gestures mark information that should be treated by the addressee as an aside from the main point. Analogues to "By the way" or "Back to the main point."
    - *Elliptical* gestures mark information that the addressee should elaborate for himself or herself; the speaker will not provide further details. Analogous to "or whatever."
  - *Citing* gestures refer to a previous contribution by the addressee:
    - *General citing* indicates "as you said earlier," that is, that the point the speaker is now making had been contributed by the addressee.
    - *Acknowledgment* of the addressee's response indicates that the speaker saw or heard that the addressee understood what had been said. Paraphrased, "I see that you understood me."
  - *Seeking* gestures aim to elicit a specific response from the addressee:
    - *Seeking help* requests a word or phrase that the speaker cannot find at the moment. A verbal paraphrase would be "Can you give me the word for . . .?"
    - *Seeking agreement* asks whether the addressee agrees or disagrees with the point being made. Analogous to "Do you agree?"
    - *Seeking following* asks whether the addressee understands what is being said. Verbal equivalents include "you know?" or "eh?" at the end of a phrase.
  - *Turn* gestures refer to issues around the speaking turn:
    - *Giving turn* hands it over to the other person. As if to say, "Your turn."
    - *Taking turn* accepts the turn from the other person. Paraphrased as "OK, I'll take over."
    - *Turn open* indicates that it is anyone's turn, as if to say "Who's going to talk next?"

as objective as reading numbers off a dial, identifying facial muscles, or counting the number of hand movements. Objectivity is not a given; it is an accomplishment.

To do this, we (Bavelas, Hagen, Lane, & Lawrie, 1989) formalized our interpretive processes into a *decision tree,* in which each step of the analyst's thinking is directed to considering specified criteria in a specified order. No intuitive leaps are permitted; the analyst is as accountable for his or her decision processes as for the final decision. We provide training tapes with many examples, and within a few weeks, newcomers with no background (or knowledge of hypothesis) can achieve independent agreement of around 90% on whether a gesture is interactive or topical in function. All of our subsequent articles report high levels of independent agreement, regardless of the analysts' experience and knowledge of our theory or particular hypotheses.

The 10% loss of reliability may well be the price of being overhearers, albeit considerably more advantaged ones than in Schober and Clark's (1989) study. Still, our agreement is often higher than many purely physical descriptions of nonverbal acts. We attribute this to the fact that, as everyday interlocutors, we have a great deal more experience interpreting such acts than describing them physically, so that our training aims to create metalinguistic awareness of what we already know. In sum, when interpretation is required, the interpretive process itself can be formalized, an approach we have found valuable as we continue to be interested in meaningful acts in context rather than in mere physical movements.

Experimental Manipulation

With the identification problem solved, we still had to demonstrate that interactive gestures were different enough from other conversational gestures to justify a new distinction. We happen to like experimental methods, but it was not at all obvious how to apply them appropriately to interactive gestures. The functions served by individual interactive gestures at their moments in conversation are quite heterogeneous; all they have in common is a reference to the addressee rather than to the topic of conversation. The key was in our theory that they are all caused by the collaborative requirements of face-to-face dialogue. We decided to hold the topic of conversation constant, manipulate the relationship between the interlocutors, and predict that interactive and topic gestures would be affected differently.

In a series of three experiments, we moved from obvious to more subtle manipulations of the speaker–addressee relationship. At first, we (Bavelas et al., 1992) simply removed the addressee; we compared individuals speaking alone with dyads talking about the same topics. The individuals made interactive gestures at a significantly lower rate than the dyads; topic gestures were not affected. (See Table 2 for the results of this and the next experiment.) The absence of an interlocutor decreased the speakers' rate of interactive gestures, in spite of the same topic of conversation. In the next experiment (Bavelas et al., 1992), there was always an interlocutor, but (in one condition) they were separated by a partition. These speakers, too, made interactive gestures at a significantly lower rate than those who could see—and be seen by—their addressees. We had predicted that the rate would decrease in the experimental condition because the addressee would not see the gesture, so it could serve no interactive function. Finally, we (Bavelas et al., in press) went to the heart of a conversation, which is whether the participants were collaborating. In both conditions, the dyads were sitting vis-a-vis, but in the experimental condition, they were instructed to talk in sequential monologues and not to help each other. In the control condition, as usual, they were permitted to have a normal dialogue. The difference in rate of interactive gestures was again significant; so the presence of a visible, collaborating interlocutor seems to be the key.

In none of these experiments did the rate of topic gestures also decrease in the experimental condition. These were either unaffected or

TABLE 2
Effects of Experimental Conditions on Mean Gesture Rates (Per Minute)

	Interactive Gestures		*Topic Gestures*	
Study	*M*	*p*	*M*	*p*
Bavelas et al. (1992)				
Experiment 1:				
Alone	1.85		20.13	
Dyad	3.28	.03	12.70	.06
Bavelas et al. (1992)				
Experiment 2:				
Partition	2.09		18.43	
Face-to-face	4.38	.03	20.75	ns

(in one case) their rate actually increased. The nature and function of gestures changed with the conditions of the dialogue. Thus, it was possible to demonstrate the existence and probable function of certain gestures within the framework of the theory outlined in this article using classical experimental methods.

Microanalysis

Some aspects of our theory of how interactive gestures function in dialogue could not be tested by experimental methods. Whatever gestures do, each does it at a particular moment in conversation. To demonstrate how interactive (or any other) gestures actually work in conversation, it is necessary to go down into where they happen, which requires microanalysis of conversational moments.

Our first microanalysis (Bavelas et al., 1992) tested an observation from the original pilot study that, whereas topic gestures usually overlapped the meaning of the words spoken with them, interactive gestures often had no simultaneous verbal counterpart. That is, interactive gestures seemed to accompany topical words. We speculated that this enables them to serve their function of touching bases with the addressee quickly and without interrupting the verbal flow, whereas topic gestures work closely with the verbal narrative. To test this hypothesis, we rated a large set of interactive and topic gestures for their redundancy with (vs. independence of) the accompanying phonemic clause. This required identifying the phonemic clause, agreeing on a general meaning of the gesture, and then (reliably) rating the degree of overlap between gesture and words. The 4-point scale ranged from *complete redundancy* (0) to *completely separate* (3) meanings of gesture and words. The results are shown in Figure 1. On average, topic gestures were fairly redundant with the words, but the meaning of 80% of the interactive gestures did not appear at all in the words. An interactive gesture was a highly specialized act inserted (as we had predicted) into the dialogue in a manner complementing rather than overlapping the words. The verbal paraphrases given in Table 1 seldom appeared in real speech.

However, there are several reasons we should not dismiss topic gestures as merely redundant with words. First, our system favored words over gestures by measuring the gesture against the words rather

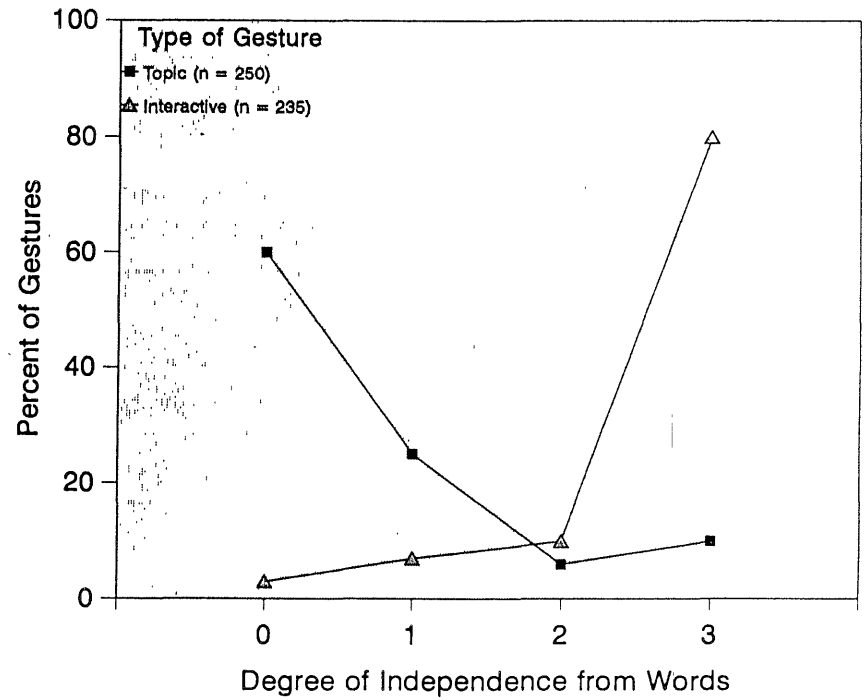

FIGURE 1 Redundancy between gestures and accompanying words for interactive and topic gestures. (Based on data reported in Bavelas, Chovil, Lawrie, & Wade, 1992.)

than the reverse. Second, as Figure 1 shows, about 16% of topic gestures in these data conveyed a meaning completely or substantially independent of the words. Third, redundancy serves an essential function in all human communication (written and spoken); without it, nuances are lost and missed meanings are more likely. Fourth, topically redundant gestures may also be serving other functions, as suggested at the outset. All of these possibilities need further investigation.

Our second predictive microanalysis (Bavelas et al., in press) sought to establish a crucial aspect of our theory of interactive gestures: their impact on the addressee. Our experiments had shown that speakers made interactive gestures at a higher rate in face-to-face dialogue than in other conditions. It seemed unlikely to us that they would make such precise gestures regardless of effect, but it would be interesting to attempt to demonstrate empirically the effect on addressees.

To do so, we took advantage of the finite number of functions these gestures seemed to serve in dialogue (given in Table 1). Although each gesture is unique, it was possible to make predictions for gestures that (we hypothesized) served a similar function. Gestures seeking help with a word should elicit help rather than back-channels, and gestures marking the delivery of new information should often elicit a back-channel acknowledgment (e.g., Speaker A's *Uh huh* after Speaker B's interactive gesture in the first example). So, after generating a set of predictions for each function, we had separate sets of analysts identify: (a) the functions of a random sample of interactive gestures and (b) the response of the addressee to the gesture. In both cases, we constructed new decision trees to guide the analysis and to ensure high reliability. Naturally, the people who identified the predicted response were completely blind to our hypotheses and unaware of different functions of gestures. When we put it all together, the results strongly confirmed that interactive gestures elicit predictable responses from addressees within 2 to 3 sec.

CONCLUSIONS

Our research has, in a sense, focused on the periphery, on the approximately 10% to 20% of conversational gestures that serve primarily interactive functions. These studies were not tests of the various functions of topic gestures or of all of the aspects of the theory outlined here. Still, they have some broader implications for gesture research: (a) it is vital to distinguish among different gestural functions rather than pooling them together, (b) interactive gestures definitely have communicative functions, (c) a linguistic or meaning-based model of gestures such as the present one cannot be tested by research generated from other models, and (d) such a model can be tested by a variety of appropriate research methods.

My overall theme is the intimate relationship between theory and method. The view of gestures outlined here necessitates certain methods and eschews others (which might be suitable to other theories). Because of these incompatibilities, it is unlikely that there can be crucial experiments of whether gestures communicate. There can, however, be tests crucial to a particular theory if (and to the extent that) the theory is fully explicated and appropriate methods are devised. It has been our experience that this task is not only possible but highly interesting and informative.

NOTE

1 The theory proposed here is part of an "integrated message model" of language in face-to-face dialogue being developed with Nicole Chovil (Bavelas & Chovil, 1993). The substantial changes from my earlier proposals about nonverbal communication (Watzlawick, Beavin Bavelas, & Jackson, 1967) have been wrought by the data (cf. Bavelas, 1990; 1991).

REFERENCES

Bavelas, J. B. (1990). Behaving and communicating: A reply to Motley. *Western Journal of Speech Communication, 54,* 593–602.

Bavelas, J. B. (1991). Recherche sur les axiomes de "Une logique de la communication" [Research into the "Pragmatics of human communication"]. *Cahiers Critiques de Therapie Familiale et de Pratiques de Reseaux, 13,* 117–132.

Bavelas, J. B., Black, A., Chovil, N., Lemery, C. R., & Mullett, J. (1988). Form and function in motor mimicry. Topographic evidence that the primary function is communicative. *Human Communication Research, 14,* 275–299.

Bavelas, J. B., Black, A., Chovil, N., & Mullett, J. (1990). *Equivocal communication.* Newbury Park, CA: Sage.

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I *show* how you feel": Motor mimicry as a communicative act. *Journal of Personality and Social Psychology, 50,* 322–329.

Bavelas, J. B., & Chovil, N. (1993). *Redefining language. An integrated message model of language in face-to-face dialogue.* (Manuscript submitted for publication).

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (in press). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin.*

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes, 15,* 469–489.

Bavelas, J. B., & Coates, L. (1992). How do we account for the mindfulness of face-to-face dialogue? *Communication Monographs, 59,* 301–305.

Bavelas, J. B., Hagen, D., Lane, L., & Lawrie, D. A. (1989, May). *Interactive gestures and a systems theory of conversation.* Paper presented at the meeting of the International Communication Association, San Francisco.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331-350.

Chovil, N. (1989). *Communicative functions of facial displays in conversation.* Unpublished doctoral dissertation, Department of Psychology, University of Victoria.

Chovil, N. (1991/92). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction, 25,* 163-194.

Clark, H. H., & Brennan, S. A. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259-294.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22,* 1-39.

Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communications, 27,* 54-63.

Cohen, A. A., & Harrison, R. P. (1973). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology, 28,* 176-279.

Efron, D. (1972). *Gesture, race, and culture.* The Hague: Mouton. (Original work published 1941, as *Gesture and environment*)

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1,* 49-98.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication, 22,* 353-374.

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

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. In A. R. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 153-175). Elmsford, NY: Pergamon.

Jones, S. S., & Raag, T. (1989). Smile production in older infants: The importance of a social recipient for the facial signal. *Child Development, 60,* 811-818.

Kendon, A. (1980). Gesture and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *Nonverbal communication and language* (pp. 207-227). The Hague: Mouton.

Kendon, A. (1987). On gesture: Its complementary relationship with speech. In A. Seigman & S. Feldstein (Eds.), *Nonverbal communication* (pp. 65-97). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61,* 743-754.

Kraut, R. E., & Johnson, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology, 9,* 1539-1553.

Lawrie, D. A. (1988). *Discourse-oriented hand gestures with a proposal for a new category: The semantic emphasizer.* Unpublished honours thesis, Department of Psychology, University of Victoria.

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

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

McNeill, D., & Levy, E. (1982). Conceptual representations in language activity and gesture. In R. Jarvella & W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics* (pp. 271-295). Chichester, England: Wiley.

Masson, M. E. J. (1984). Memory for the surface structure of sentences: Remembering with and without awareness. *Journal of Verbal Learning and Verbal Behavior, 23,* 579-592.

Meissner, M., & Philpott, S. B. (1975). The sign language of sawmill workers in British Columbia. *Sign Language Studies, 9,* 291-308.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 615-622.

Random House unabridged dictionary (11th ed.). (1993). s.v. "ground."

Rimé, B., & Schiaratura, L. (1991). Gesture and speech. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 239-281). Cambridge, England: Cambridge University Press.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory and Cognition, 2,* 95-100.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21,* 211-232.

Slama-Cazacu, T. (1976). Nonverbal components in message sequence: "Mixed syntax." In W. C. McCormack & S. A. Wurm (Eds.), *Language and man: Anthropological issues* (pp. 217-222). The Hague: Mouton.

Watzlawick, P., Beavin Bavelas, J., & Jackson, D. D. (1967). *Pragmatics of human communication.* New York: Norton.