From:

Social Communication

Edited by
Klaus Fiedler

Published in 2007
by Psychology Press
270 Madison Avenue
New York, NY 10016
www.psypress.com

Published in Great Britain
by Psychology Press
27 Church Road
Hove, East Sussex BN3 2FA
www.psypress.com

Psychology Press is an imprint of the Taylor & Francis Group, an informa business

Typeset by RefineCatch Limited, Bungay, Suffolk, UK
Printed in the United States of America on acid-free paper
Cover design by Lisa Dynan

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress.

ISBN: 978-1-84169-428-3 (hbk)

Psychology Press
Taylor & Francis Group
NEW YORK AND HOVE

10

Conversational Hand Gestures and Facial Displays in Face-to-Face Dialogue

JANET BAVELAS and JENNIFER GERWING

Many scholars have proposed that the primary site of language use is face-to-face dialogue (e.g., Bavelas, 1990; Bavelas, Hutchinson, Kenwood, & Matheson, 1997; Clark, 1996, pp. 8–10; Fillmore, 1991; Goodwin, 1981; Levinson, 1983; Linell, 2005). Whether within the family, with friends, at work, or in brief conversations with strangers, the vast majority of everyday social exchanges, from mundane to important, take place in person. Moreover, face-to-face dialogue is a child's first language developmentally, and it is arguably humanity's first language phylogenetically. This format, therefore, is centrally important to understanding social communication. Face-to-face dialogue has a combination of affordances that make it different from other language formats, such as written text, formal lectures, phone conversations, etc. (Bavelas & Chovil, 2006; Bavelas, Coates, & Johnson, 2002). The present chapter will focus on combining two of these unique features: (1) Face-to-face dialogue includes an interlocutor who can respond reciprocally and in real-time; there is both a requirement to coordinate and the opportunity to collaborate. (2) In face-to-face dialogue, the participants have visible as well as audible resources for social communication, that is, not just words and prosody, but also hand gestures, facial displays, gaze direction and timing, body orientation, and objects in their shared environment. We propose that some non-verbal acts play an important role in the participants' collaboration, so that the study of these non-verbal acts in dialogue is a promising way to understand basic processes in social communication.

In 1985, Clark chided social psychology for its relative neglect of the social aspects of language use. Although the individual processes that preoccupy linguists and psycholinguists (e.g., language production and comprehension, syntax, and semantics) may not be of interest, it is clear that the social interactive aspects of communication are within the domain of social psychology. However, to illustrate the intrinsically social nature of communication (e.g., coordination and common

ground), Clark had to rely largely on qualitative research because there were few experimental examples at the time, virtually none of them by social psychologists. He and his colleagues have gone on to examine experimentally the social, collaborative aspects of language use, such as the dyadic process of grounding and other social influences on verbal reference (Clark, 1992, 1996; Schober & Brennan, 2003). Studies such as theirs have two implications for experimental social psychologists who are interested in communication: First, verbal communication in dialogue is joint action, something that two people do together. Moment by moment, dialogue is an inherently social process. Second, it is possible and desirable to advance our knowledge of these processes using experimental methods.

Our research group's experimental interests have focused on the related possibilities that certain *nonverbal* acts are also an important part of the social collaboration in face-to-face dialogue and that such acts are amenable to quantitative experimental study (e.g., Bavelas, Chovil, Coates, & Roe, 1995; Bavelas, Coates, & Johnson, 2000; Gerwing & Bavelas, 2004). The main theme of the present chapter is substantive, focusing on the role that certain specific non-verbal acts can play in conversational collaboration. The secondary theme is methodological; the studies reviewed also demonstrate that there are both exemplars and principles from which to develop classical experimental designs in an area often restricted to other methods.

In order to calibrate with the reader about the scale and kind of phenomena to be discussed, we will first specify our perspective on both social communication and the non-verbal acts we are referring to. Then we will review the growing literature on collaborative functions of conversational hand gestures and, finally, the much smaller but promising literature on conversational facial displays.

WHAT IS "SOCIAL"?

In focusing on face-to-face dialogue as a social process, it is first necessary to explicate some of the many possible senses of *social* and to indicate clearly those senses that we do not address here. Language is social in a societal sense. For example, there are clearly socio-cultural differences that frame both verbal and non-verbal acts (e.g., forms of polite address or display rules for facial expressions). This socio-cultural focus is much more global than ours. Another sense in which language is social is the undoubted influence of social motivations, cognitions, or traits, that is, mental processes or characteristics of the individual that may influence communication (e.g., a particular social schema, a trait of extraversion, or an intention to deceive). This focus on individuals and their intrapsychic processes is not ours either. Instead, we are drawing attention to a different level of social influence, which can be called the *micro-social* (Bavelas, in press): the observable, moment-by-moment details of communicative acts between the participants in a face-to-face dialogue, acts that typically occur in seconds or less.

MAKING DISTINCTIONS AMONG NON-VERBAL ACTS

In a classic article, Ekman and Friesen (1969) outlined a wide and diverse repertoire of non-verbal behaviors, emphasizing the considerable differences in the nature and use of behaviors that are often grouped together simply because they are "not verbal." Of Ekman and Friesen's five categories, we are limiting our focus to two: actions that serve either to *illustrate* or to *regulate* face-to-face dialogue, specifically conversational hand gestures and facial displays. We exclude the other three categories: emblems, which typically occur without speech; true (involuntary) affect displays; and adaptors (such as scratching or blinking) that clearly serve functions other than communication. Some of these (as well as possible newer categories such as self-prompting gestures) may also occur in conversation, but they are not directly tied to the social communicative aspects of the conversation.

One of the primary differences from other approaches is our focus on the relationship of the non-verbal acts to words. The vast majority of experimental studies treat verbal and non-verbal behaviors as separate topics; indeed, the main interest in non-verbal behaviors is often that they may reveal what words cannot or will not. We propose that there is a subset of non-verbal behaviors – many hand gestures and facial displays in certain settings – that have a close relationship to the words they accompany and are of interest because of their role in social communication at the micro-social level. Although it would be better to be able to provide a video example, two transcriptions may clarify for the reader the phenomena we will be discussing.[1] In both cases, the participants were previously unacquainted volunteers discussing assigned topics in the laboratory:

Example 1: Hand Gestures The speaker in this experiment (Bavelas, Gerwing, Prevost, & Sutton, 2006) was describing a drawing of an unusual 18th century dress[2] to an addressee who would need to be able to identify this particular dress later. A prominent feature of the dress was the unusually wide skirt, which jutted out horizontally to each side and then made a slightly curved corner before draping straight to the floor. In the following, the underlining indicates the timing of each of the gestures to the words.

Speaker:	"It $_1$starts HERE, to go out, and $_2$goes out like – $_3$REALLY BIG, OK?"
Addressee:	"$_1$It's got like, one of the POUFY ones that's got, $_2$like, the CORSET in the middle?"
Speaker:	"It's CRAZY poufy!"

In gesture 1, the speaker placed her hands on either side of her own waist. Her deixis (the demonstrative pronoun "here") drew the addressee's attention to the fact that the location of "here" would be indicated by her gesture rather than her words. She went on to describe the width of the dress by reaching out horizontally the full length of both arms (gesture 2). This gesture was not redundant with the words it accompanied because the words ("goes out") were ambiguous as to width, and the gesture depicted the width specifically. At the very end of her reach, she

shaped a rounded downwards curve with her hands (gesture 3). This gesture indicated both where the dress stopped "going out" (and therefore what "REALLY BIG" meant) and the shape of the contour of the skirt at the hipline. Again, these details did not appear in her words. However, the addressee obviously understood them because his gesture (4) reached out the same width on his own body, and the word "poufy" captured the rounded shape she had depicted. Then, with his reference to a "corset," he returned his gesture to his own waist (5), indicating that he had also understood where the dress started to go out.

Example 2: Facial Displays The speaker was relating a past close-call in which she nearly fell off a barn roof (Chovil, 1989; also described in Bavelas & Chovil, 1997). She was telling the story as if it were humorous, smiling in between various facial illustrations of her reactions and those of her father, who was trying to rescue her. At one point, the addressee interrupted to ask a question, and the speaker's facial displays (indicated by underlining below) changed immediately:

Addressee: "How high up were ya?"
Speaker: "¡UM!" [1 s pause] ¿I'm really not too sure. It's quite high up there."

With "UM!" (display 1), the speaker looked off to the side and changed to an intensely serious face, furrowing her eyebrows and pulling her mouth straight and smaller, as if she were thinking hard. After the pause, as she began to answer, she looked back toward the addressee and resumed smiling (display 2). Her first facial display, when combined with her exclamation, shift of gaze, and silence, had informed him what her momentary silence meant: that she had heard his question and still had the speaking turn but was having difficulty coming up with an exact answer. Had she continued smiling or simply gone blank when he asked, her failure to answer would have been inexplicable to him.

Four Criteria

There are four defining criteria that distinguish the subset we are focusing on from the vast and interesting domain of other non-verbal behaviors and even from other functions of hand or facial actions. These criteria set the parameters of our literature review and may also be useful for imagining future studies in the area.

Social Setting To investigate *conversational* hand gestures and facial displays, we must observe and record them where they occur, which is in spontaneous face-to-face dialogue. Obviously, the participants must be able to see each other, and they must be free to gesture and to make facial displays if they wish. It may be desirable to include conditions that, for example, restrict sociability, visibility, or ability to gesture, but these would be for purposes of comparison to the face-to-face condition. An interest in social processes would dictate that the unit of study be the dyad; both interlocutors must be participants who are communicating with each other spontaneously. Designs where one interactant is an experimenter, interviewer, or confederate are unlikely to produce all the details of natural social reciprocity typical of face-to-face dialogue and will almost certainly preclude collaborative behaviors. There is evidence to suggest that natural behaviors by an addressee are more subtle and reciprocal than mere back-channels (Bavelas et al., 2000) and that a confederate or experimenter trying to respond in a "neutral" or "standard" manner could have unintended effects (Beattie & Aboudan, 1994). Therefore, although the dyad will probably be doing a task assigned by the experimenter, their communication must be unscripted and generated in their interaction. That is, the reciprocity and mutual influence that ordinary conversation permits must be fully possible; only this way can dialogic processes emerge.

In contrast, other interests in gesture would not require dyadic interaction. Most of the experiments on hand gestures within both social psychology and psycholinguistics have been cognitive, examining the role of gesture in language production. The two main contemporary approaches differ on where gesture plays its role: *Lexical access* theories (e.g., Krauss, Chen, & Chawla, 1996; Rimé & Schiaratura, 1991) propose a peripheral, self-prompting role, namely, that in the latter stages of language production a gesture may assist the speaker in finding the right word. McNeill's *growth-point* theory (e.g., McNeill, 1992; McNeill & Duncan, 2000) proposes that gestures reflect the thinking behind language production itself, specifically, the global-synthetic images that combine with linguistic categories in the overt message. Although these two approaches differ regarding whether gestures merely assist lexical access or are evidence of more central cognitive processes, the primary interest is definitely on individual mental processes, not social interaction. The ideal setting would be an individual speaker alone or in monologue with a passive partner (e.g., the experimenter or interviewer), not in a reciprocal dialogue. Note that, because of the focus on language production, the gestures of the non-speaking addressee would not be of interest, although they can be important to social theories (e.g., Bangerter, 2004; Clark & Krych, 2004; Furuyama, 2000).

Similarly, the dominant interest in facial expression has been individual in focus. From Darwin (1872/1965) to Ekman and Friesen (1969) to Izard (1977) and Buck (1984), virtually all social psychological interest in the face has been tied to emotion. The preferred setting for the study of the face as emotional expression is the individual alone. For emotion theories, social interaction is a confound that is likely to mask the pure expression of emotion (e.g., by display rules), whereas social interaction is precisely our locus of interest.

Timing The second criterion is temporal. Hand gestures and facial displays in dialogue are very quick, often begun and completed in a second or less. If the individual is speaking at the moment, these actions are closely timed with his or her words. As shown in Example 1, they follow the verbal syntax, beginning and ending precisely with the phrase that they supplement or complement. Or, as shown in Example 2, they are timed to a particular interactive moment (e.g., the speaker's indication that she was thinking about what the addressee had asked). One characteristic of hand gestures and facial displays is that their movements can match the speed of words.

The achievement of precise timing with words is noteworthy for both gestures and facial displays, although for different reasons. The facial muscles are rapid enough to track the speed of speech; for example, eyebrows can flash to emphasize a single word. For gestures, the timing takes a different route because hands are relatively slower. A gesture often begins with a non-lexical preparatory phase in which the individual moves his or her hand(s) into position (e.g., raising the hand from the lap or moving it to a different place from the previous gesture); this is simply a physical movement that has no relation to the gesture's meaning. Once the hand is in place, which may take only a fraction of a second, then the real gesture begins (called the *stroke* phase), and the hand depicts the speaker's meaning (Kendon, 1980, p. 212; see also, McNeill, 1992, p. 25). Thus, the speaker has to begin and pace the preparatory phase so that the stroke will occur with the right word or phrase.

The growth-point theory of gesture would, like us, expect most gestures to be synchronous with speech (e.g., McNeill, 1992, pp. 25–29). In contrast, precise timing is important to lexical access theories in a different way, because a gesture can only assist with word-finding *before* the word is uttered. For theories focused on a self-prompting function of hand gestures, the gestures that are of interest are those whose stroke phase precedes a word, because when the word and stroke phase co-occur, lexical access cannot be an issue. (See also Nobe, 2000, for other timing considerations.)

Ekman (1997) distinguished between facial expressions of emotion and the communicative actions that he called *referential expressions* and *conversational signals*. One key difference he pointed out was their timing. Ekman proposed that, in contrast to communicative facial actions:

> While facial expressions of emotion often occur during conversation, their location in the speech flow is related *not to the structure of talk* but to the semantics, revealing an emotional reaction to what is being said or not said (p. 340, italics added)

Like gestures, the facial displays that play a role in social communication will be related to the structure of talk, that is, timed with individual words or syntax.

Contextual Meaning The third criterion concerns the relationship of the gesture or facial display to other features of the conversation: The meaning of the act must be directly related to the *immediate communicative context* (Bavelas & Chovil, 2006; Chovil, 1989; Gerwing, 2003; Gerwing & Bavelas, 2004). As illustrated in Examples 1 and 2 above, this context includes the concomitant words, prosody, and other communicative acts, as well as what is happening in the interaction between interlocutors at that moment (e.g., the facial displays in Example 2). Both the meaning and the form of the act will depend intimately on its particular context, just as the precise meaning of a word always depends on the immediate context in which it is used; neither stands alone, unrelated to what went before and what is happening at the same time. A corollary is that we are not separating these actions from the words, prosody, or each other. They are always treated as part of an *integrated message* (Bavelas & Chovil, 2000) or *composite signal* (Clark, 1996) rather than as a separate channel. Note that the non-verbal actions in a multimodal message are not necessarily redundant with the words; they may be complementary, providing unique but related information.

In contrast, although words are obviously important to lexical access theories, the focus is on a single word (i.e., on vocabulary and dictionary equivalents) and not its conversational context. Moreover, the gestures of interest must be redundant with the words that they prompt. Although growth-point theory does connect the gesture with the speaker's broader meaning at the moment (and would expect some non-redundant gestures, as we do), it does not incorporate the social interaction processes that are part of speech in dialogue.

It is quite possible and fairly common to study facial expressions of emotion without any relationship to conversational context, for example using still photos. Even when captured in dialogue, the focus of interest is on a set of universal, context-free emotional expressions rather than on the local, highly context-specific, and extremely varied information that a conversational facial display may momentarily convey. However, as with cognitive theories of hand gestures, there are lively divisions within the field of emotional expression. Russell and others (e.g., Russell & Fernandez-Dols, 1997) have challenged the universality or context-free nature of facial expressions of emotion. Researchers with an ethological perspective, such as Kraut and Johnston (1979) and Fridlund (1997) have emphasized the role of sociality, or audience effects, in facial displays of emotion. Although these approaches share some of our criteria, neither of them studies *conversational facial displays*, that is, facial actions that are closely timed to and integrated with words in dialogue.

Social Function Fourth, we are interested in these actions as part of moment-by-moment social communication at the micro-social level. Some hand gestures and facial displays play an essential role in this collaborative process. Working with words, prosody, and each other, these actions shape and are shaped by the ongoing dialogue; they convey information to the interlocutor; and they can help regulate the process of interacting in dialogue. The research reviewed below outlines the beginnings of our knowledge of these functions.

All cognitive theories implicitly limit their interest to the role of a gesture in the individual speaker's language production, not including any other effect on the addressee, nor any social interaction. In these theories, the role of gestures is, like words, unilateral; they do not shape and are not shaped by dyadic factors.

Studies of emotional expression (e.g., Ekman, 1993) understandably focus on the overt facial action primarily because of its relationship to the individual's underlying emotional state. That is, although one reason for studying emotional states is that these clearly have implications for social interaction, the primary interest is the inner life of the individual.

Table 10.1 summarizes how the four criteria introduced here can distinguish a social communicative approach to hand gestures and facial displays from other, more familiar approaches to the same physical modalities. It is clear that, while the cognitive and social approaches to gesture are ultimately complementary and

TABLE 10.1 Contrasting approaches to hand and facial actions

	Social Communication	Lexical Access Theory	Growth-Point Theory	Emotional Expression
Actions of interest	Hand gestures and facial displays	Hand gestures	Hand gestures	Facial expressions
Social setting	Must be spontaneous face-to-face dialogue with a visible recipient (who is not a confederate or experimenter)	An individual speaking in monologue or dialogue; no visible recipient required	An individual speaking in monologue or dialogue, no visible recipient required	Preferably a solitary individual, no visible recipient required
Timing	Gesture synchronized with words or with specific moments in dialogue	Gesture precedes its lexical affiliate	Gesture synchronized with words	Not necessarily timed to the flow of talk; may be unrelated to speech (e.g., in a photograph)
Meaning	Depends on immediate communicative context, polysemous	Closely related to a particular lexical affiliate, usually a word	Depends on meaning, concept	Universal rather than context-specific; fixed set of global meanings
	Tightly integrated with words and prosody, but may be complementary rather than redundant	Same meaning as the lexical affiliate (redundant with word)	Components linguistic components not necessarily redundant	Does not require words
Function	Social: illustrates or regulates face-to-face dialogue	Individual/cognitive, prompts a missing word	Individual/cognitive, language production	Individual/emotional: expresses emotional state

definitely not mutually exclusive, they are as different as they can be while sharing a common interest in the broad class of hand movements related to speech. Similarly, the study of facial expression of emotion and the approach we are describing here differ so much in their units of analysis, methods, and ultimate goals as to constitute independent fields of study.

EXPERIMENTAL INVESTIGATIONS

Social psychologists face at least two obstacles when considering the study of face-to-face dialogue. First, from early in its development, experimental social psychology has eschewed the use of real dyads in favor of individuals or confederates (e.g., Aronson & Carlsmith, 1968). Similarly, the goal of objective measurement would seem to preclude analyzing non-verbal acts at the level of meaning (rather than counting them as physical movements). For these reasons, while acknowledging a historical and ongoing debt to qualitative studies in other fields (see Kendon, 2004, for a comprehensive review), we will focus on experimental studies. The available experimental research illustrates that both of the above apparent barriers can be overcome. It is possible to study the use and meaning of hand gestures or facial displays in real dyads without sacrificing either experimental control or objective measurement. What follows is not a complete literature review, although we believe we have located most of the studies that meet our four criteria.[3]

Experimental Research on Conversational Hand Gestures

It would be premature to organize the limited experimental research on conversational hand gestures within a specific theory, because these studies were conducted for different reasons at different times, often in isolation from each other. However, these studies do demonstrate the viability of at least two preliminary conclusions: First, the participants in face-to-face dialogue often use *each other's gestures* to coordinate their conversation or complete a shared task (Bavelas et al., 1995; Bavelas, Chovil, Lawrie, & Wade, 1992; Clark & Krych, 2004; Furuyama, 2000). Second, social factors such as common ground, shared perspective, or perceptual co-presence have a significant effect on the form of an individual participant's gestures (Bangerter, 2004; Gerwing, 2003; Gerwing & Bavelas, 2004; Özyürek, 2000, 2002; Woods, 2005). We will summarize each of the above ten studies in some detail, in order to illustrate both the substantive findings and the variety of methods already being used in the field.

Interactive Gestures Close observers of gestures have long noticed that not all hand gestures in conversation are about the topic of discussion. There is a small group of simpler gestures, often mere flicks of the hand, that do not depict anything the speaker is overtly talking about. Bavelas and colleagues (Bavelas et al., 1992; Bavelas et al., 1995) isolated this group, which were about 15% of gestures in their face-to-face dialogue data, and noticed two common features: First, the physical form of the gesture always included an orientation of the hand toward the

addressee; for example, a quick point with one or several fingers, an exposed palm, or briefly "offering" open hands to the addressee. Second, when interpreted in its immediate communicative context, the meaning of the gesture included some reference to the addressee or to the speaker–addressee interaction. For example, flicking the hand toward the addressee often accompanies a verbal reference to what the addressee had said earlier – a metaphorical citation.

> *Example 3:* In a getting-acquainted conversation (Bavelas et al., 1992, p. 471), one participant had suggested that working for Canada Customs would be a good summer job. Later, the other participant listed several other possibilities and then added the other person's suggestion:
>
> Speaker: ". . . and Customs is DEFINITELY, is career or-oriented."

As he said "Customs", the speaker moved his hand up and toward the listener, almost as if tossing something to him. We propose that this gesture conveys a meaning akin to "which YOU suggested," that is, the speaker is citing or crediting the listener as the source of the idea.

Or, while searching for a word, the speaker may hold the palm out as if to receive something from the addressee. Because all of these various functions are related to the social interaction independently of topic or context, Bavelas et al. (1992) called them *interactive gestures*.

Having developed a reliable operational definition, we were able to test the hypothesis that this subset of hand gestures functions to aid the maintenance of conversation as a social system. Several experiments demonstrated that these forms were much more likely to occur in spontaneous face-to-face dialogue. Their rate was significantly reduced when the speaker was (1) alone, talking in monologue, (2) talking to an addressee who could not see them (Bavelas et al., 1992), or (3) talking to a visible addressee but in alternating monologues rather than the give-and-take of dialogue (Bavelas et al., 1995). Moreover, although they were virtually never redundant with the words they accompanied, the interactive gestures themselves had a statistically significant effect on the addressees who, for example, provided a word when prompted solely by the interactive gesture. The existence of these "gestures specialized for dialogue" (Bavelas et al., 1995) elucidated how the speaker, besides delivering content, must attend to a parallel social requirement to include and coordinate with the addressee on a moment-by-moment scale. These quick, non-redundant gestures seem to be one efficient way to do so.

Collaborative Gestures Furuyama (2000) pointed out that many gesture studies investigate the intrapersonal influences on the gestures. His interest focused instead on possible interpersonal influences on gestures, specifically, how the gestures of one participant in a dialogue might influence how gestures are formed by the other participant. Furuyama videotaped pairs of participants while they engaged in a collaborative task: He had taught one participant (the Instructor) in each pair how to make a complex origami figure and then asked that

person to instruct the other participant (the Learner) how to make it. He did not give the participants any origami paper, so they often used gestures to depict the paper and their own actions with it.

Furuyama's (2000) data revealed the highly social phenomenon of *collaborative gestures*. He was able to show, reliably, that instead of gesturing in their own gesture space (near their own body), Learners often placed their gestures directly in front of the Instructor, right in the space created by the Instructor's previous gesture and building on the Instructor's gesture. For example, the Instructor might gesture a particular fold, perhaps by outlining a virtual piece of paper and folding one side over another, and the Learner would reach over to point at the corner of the virtual paper (which was, in fact, empty space) to ask a question. It was as though there were invisible paper, and the Instructor's gestures, which created and maintained the imaginary paper, could be manipulated by both participants. These collaborative gestures were almost 18% of the 400 Learners' gestures analyzed (calculated from Furuyama, 2000, p. 108, figure 5.2), and they were significantly more likely than non-collaborative (individual) gestures to accompany speech – often the Teacher's speech rather than the Learner's own speech. Moreover, the Learners only made collaborative gestures when their Teacher had faced the focal point of their gesture toward the Learner.

Collaborative gestures are a dialogic event, created by the *joint actions* (Clark, 1996) of the participants. They reveal an intricate coordination of space, gesture, and speech to complete a complex spatial task with little difficulty. To appreciate the mutual importance of dialogue and gesture in this study, one need only imagine doing such a task on the phone or with a videotaped Instructor whose gestures were not tangibly available to the Learner for collaboration.

Monitoring Understanding via Gestures Clark and Krych (2004) expanded our understanding both of what gestures are (by elevating actions with objects from a simply instrumental function to a symbolic one) and of the influence of addressees' gestural contributions on speakers' utterances. Their dyads consisted of Directors, who had a model constructed from Lego blocks, and Builders, who were to build the same model as instructed by the Director. In addition to several other analyses, Clark and Krych conducted a reliable, in-depth analysis of the gestures in a subset of their data. They found that when the pair could interact and the Director could see the Builder's workspace, they did the task more quickly. Analysis of the way they worked revealed the probable reason for the advantage they had. While Directors gave instructions, Builders often responded with actions such as pointing to a particular block, picking it up and exhibiting it to the Director, or poising a block over a possible position. These actions, which provided updates on the Builders' current state of understanding, immediately influenced Director's utterances, often mid-sentence. For example, when the Builder gesturally proposed a correct action, the Director broke off further instruction about that action and moved on to the next step. Actions that proposed an erroneous step were met by a precisely timed correction.

By examining the participants' audible and visible contributions to a conversation while collaborating to accomplish a joint task, Clark and Krych (2004) made a

strong case for taking into account all of the participants' resources when analyzing language use. In their terms, the participants' contributions to the conversation were the "artful orchestration" of vocal signals, gestural signals, nodding or shaking heads, directing eye gaze, and other mutually visible events, and "models of language use that are limited to only part of this process are necessarily incomplete and, for many purposes, incorrect" (p. 33).

Common Ground: Gestural Space Common ground (also called *common knowledge, mutual knowledge* or *belief*, or *joint knowledge*; cf. Clark, 1996, Ch. 4) is information that the participants in a conversation share, for example, because they share a culture or a past experience or are simply present in the same surroundings. The latter can be called *perceptual co-presence* and has been shown to affect the form of gestures. Özyürek (2000, 2002) examined the effects of a dyad's spatial relationship to each other on the form of the speaker's gestures. The speakers had watched a cartoon and then narrated it to addressees who had not seen it. Within the cartoon were several situations where characters or objects moved from one place to another (e.g., running into a hotel or climbing up a drainpipe), and speakers often used gestures to describe these movements. The 2000 study used a within-design in which speakers narrated the story to two addressees who were seated at either side of the speaker in a triangular formation, and also to one addressee who was sitting on one side of the speaker. In the 2002 study, the speakers narrated the story to only one addressee, who sat either directly across from the speaker or off to one side. Özyürek located all of the speakers' gestures depicting movement and, using a transparency on the monitor, she was able to trace the trajectory of each gesture. The direction of movement gestures that depicted "into" and "out of" differed according to how the participants were seated. Speakers represented the direction as into or out of *the space the participants shared*, which differed by experimental condition. For example, what was "in" or "out" changed when the addressee was sitting to the side rather than facing the speaker, and speakers accommodated to this difference, presumably so that the meaning of their gestures would be clear to their addressees. Thus, the location of the shared physical space between participants influenced the direction and orientation of the speaker's gestures. These adjustments in gesture direction were not related to changes in the speaker's speech; speakers did not change their verbal descriptions when the shared space changed. Nor was there a change in gestures for movements, such as "up," the meaning of which would not be affected by these particular seating conditions (i.e., a gesture indicating "up" would look the same whether participants were sitting face to face or side by side). It was the relationship between the configuration of shared space and the meaning that the speakers were conveying that determined differences in the speakers' gestural representations of the same movements in the cartoon. Özyürek (2002) claimed that her findings supported the view that "speakers use representational gestures accompanying their speech to communicate their intended message to their addressees" (p. 701), stating that the findings provide evidence against the view that communicative functions of the gestures are incidental.

Common Ground: Visual Perspective Bangerter (2004) investigated more closely the relationship between words and gestures, showing how participants can elect to contribute information either way in order to minimize the collaborative effort required to establish mutual understanding. Pairs of participants did a referential communication task in which a Director, who had an array of photos of faces that were arranged in a particular order, helped the Matcher construct the same arrangement from a larger collage of photos. The larger set was on a board that both participants could see. The participants were free to use whatever means of communication they chose at the moment to do the task, including words (i.e., descriptions of the photos) or gestures (i.e., pointing).

Bangerter (2004) varied whether Directors and Matchers could see each other (and thus use pointing to refer to the photos) as well as their shared visual perspective, that is, the distance between the board and the participants. When the participants could not see each other (but could both see the photo board), they used significantly more words to do the task. Presumably, this difference occurred because they could not use pointing, only words. For the visible pairs, the manipulation of distance changed the relative utility of words and pointing. At close distances, pointing is efficient and unambiguous; at farther distances, pointing would be more ambiguous. His reliable analysis revealed, first, that when pointing was an option (i.e., at close distances) both participants used it preferentially to indicate a particular photo, rather than verbally describing the location and features of the photo. They often combined pointing with verbal deixis (*this, that, here, there*), which suppressed or replaced full verbal descriptions. When the photos were far enough away that their shared visual perspective rendered pointing gestures ambiguous, participants used verbal descriptions. These findings strongly suggest that the participants were systematic, flexible, and opportunistic in their choice of modality for referring to the photos. They loaded information into their words or gestures in ways that established reference as quickly and accurately as possible, given their shared perspective.

Common Ground: Shared Experience Common ground can also be experimentally manipulated by varying the experiences that the dyad do or do not share with each other. The effects of common ground on words are well known: When interlocutors can take information as mutually understood (as "given"), they often use fewer words to refer to it (e.g., Clark & Wilkes-Gibbs, 1986), use pronouns instead of full reference (e.g., Chafe, 1974; Kess, 1992), and even articulate the familiar words less precisely (e.g., Fowler, 1988; Fowler & Housum, 1987; Hunnicutt, 1985). Gerwing (2003; Gerwing & Bavelas, 2004) investigated whether conversational hand gestures respond to participants' common ground in a similar manner. We gave participants tasks to do alone and then asked them to discuss the tasks with each other. The tasks involved objects with no familiar name, so reference to them was likely to require gestures. The experimental manipulation was whether participants shared common ground or not, that is, whether they had had the same objects or not. Before they began their conversation, they were told whether they had done the same tasks (common ground condition) or not (no common ground condition). Common ground had a significant effect on the form

of the initial gestures that a speaker used to refer to the identity of the object. That is, when participants did not share common ground (i.e., when their addressee was unfamiliar with the object), the speakers' initial gestures were reliably judged to be more complex, precise, or informative than when they did share common ground. For the latter, they made sketchier, "sloppier" gestures, because that was all the addressee needed.

Common Ground as "Given" Information Common ground can also develop within a dialogue. After new information has been introduced and mutually understood, it becomes *given* (versus *new*) information. Using the same dyads as described above, Gerwing (2003) did a qualitative analysis of the effect of previous reference within their conversations on the form of successive gestures. As predicted, later gestures for the same referent became more schematic, while gestures for new information were sharper and clearer:

> *Example 4:* The speaker had already described a "whirligig" (a stick with a propeller on top) and gestured a launch by twirling the stick between the palms. Now he was describing how he caught it (Gerwing & Bavelas, 2004, p. 175):
>
> Speaker: "And you just twirl it and catch it, that's the idea."

His gesture depicted a very tiny twirling action, using only his fingers, followed immediately by a life-sized catching motion. The twirling was given information, and the catching was new.

Woods (2005) showed the same effect in a controlled experiment. He taught one participant (the Teacher) several novel dance moves, using his own actions and diagrams but no hand gestures. Then, over a series of trials, each Teacher showed a Learner various combinations of these moves. Because they were seated at a table with no diagrams, the Teachers gestured the dance moves with their hands. Several of the same moves were required repeatedly in successive trials (in different combinations and positions within the dance), and the dependent variable was the time the Teacher devoted to gesturing each move; interanalyst agreement was $r = .995$. There was a significant decline in the duration of the same move from first to second trial, then a stable, shorter time thereafter. Thus, when the move was new to the Learner, the Teacher took longer to depict it. When the move appeared in a later trial, it was given information, and the Teacher made a quicker version of it because the Learner was now familiar with it.

The effect of successive trials on the length of gestural reference replicates and extends a similar effect on verbal reference found by Clark and Wilkes-Gibbs (1986) and Schober and Clark (1989). More broadly, this experiment, like the others in this group of studies, showed that the form of the gesture or description was not fixed by its referent but was instead flexible and significantly influenced by variation in the participants' common ground. Gesturers were acutely sensitive to what their interlocutors knew or could see, and they shaped their gestures accordingly. Arguably, they met Grice's (1975) maxims of quality and quantity – their gestures were sufficient for their task, but no more than necessary.

Systematic Investigations of Conversational Facial Displays

A good observer will notice that, even in the absence of obvious emotion, the faces of individuals who are engaged in ordinary conversation are not blank or impassive. However, non-emotional facial displays are a rare topic in the research literature. For example, in 2005, *facial expression* as a default term in PsycInfo produced over 3000 peer-reviewed hits, but the exclusion of any variation on *emotion* eliminated two-thirds of these, and the application of our four criteria (cf. Table 10.1) reduced the remainder to a handful. Perhaps because of this virtual equation of facial expression with emotional expression, Ekman later clarified and restated the Ekman and Friesen (1969) category distinctions to explicitly include communicative facial actions. As noted earlier, Ekman (1997) confirmed that faces can produce not only emotional expressions but also *referential expressions* and *conversational signals* (analogous to gestural illustrators and regulators), which are "part of the structure of the conversation, part of the flow of talk, and governed by the rules which govern the production of speech" (Ekman, 1997, p. 340). However, our criteria that the individuals be in a spontaneous face-to-face dialogue and that their faces be studied in the context of the verbal part of that dialogue reduce a vast literature to a handful of studies, of which few are experimental. We will therefore begin with a strong set of systematic and often quantitative investigations which, although they lacked experimental manipulation, nonetheless drew the first maps of directions in which future experiments might proceed.

Communicative Eyebrow Actions Ekman (1979) applied his newly developed Facial Action Coding System (FACS; Ekman & Friesen, 1978) to eyebrow actions both in emotional expression and as conversational signals.[4] Drawing on a diverse but unspecified observational database, he described a wide range of meaningful conversational eyebrow movements by speakers and addressees. First, speakers often reinforce their vocal stress on a word with a *baton*, or quick eyebrow raising movement, or they may hold this position to *underline* several words as the emphasized unit. Second, Ekman found, as had observers back to Darwin (1872/1965), that speakers often use brow raises to indicate that what they are saying is a *question*, even though it may not be a question syntactically. (It is likely the brows and prosody work together in these cases; that is, prosodic stress would accompany a baton, and rising intonation would accompany a question.) Ekman speculated that brow movements also serve other punctuation functions, analogous to commas, exclamation marks, and so forth. Third, when they momentarily cannot find the word or phrase they want, speakers may mark their *word search* by squinting or by eyebrow movements with eyes looking up (as in Example 2, above). Fourth, Ekman's observation that these actions often occur with speakers' fillers, such as *uh*, is particularly noteworthy in view of Clark and Foxtree's (2002) subsequent evidence that such verbal fillers act as signals to the addressee about encoding difficulty. Finally, addressees' eyebrows are active as well. In addition to verbal back-channels and nods, addressees may move their brows to indicate that they are attending, not understanding (e.g., by squinting), or questioning the

speaker. By combining the latter two (i.e., simultaneously squinting and raising eyebrows), listeners can indicate incredulity. None of these functions is directly or even necessarily related to emotional expression. Instead, they are intimately related to the dialogue, providing information to the interlocutor about how to interpret what the other person is communicating at that precise moment.

Smiles as Back-Channels Brunner (1979) studied the conversational use of addressees' smiles in an intensive analysis of four conversations videotaped in the laboratory. He conducted detailed statistical analyses of these smiles and the linguistic and paralinguistic aspects of the conversation, with an emphasis on the relationship between the addressee's smiles and the speaker's turn signals. Like other back-channels (e.g., nods, "yeah," and "uh-huh": Yngve, 1970), the addressee's smiles had a significant tendency to occur at points where the speaker had completed a grammatical unit and turned his or her head toward the addressee. Also, like other back-channels (but in contrast to addressees' attempts to gain the speaking turn), the addressees' smiles were only loosely related to speech boundaries and were not repressed by gesticulation. Brunner's evidence suggested strongly that addressees' smiles act as back-channels, providing feedback to the speaker without beginning to speak themselves. This study is probably the first quantitative demonstration not only of the close relationship between facial actions and speech but also of the role of a facial action in conversation as coordinated social interaction.

Conversational Functions of Facial Displays Chovil (1989; also described in Chovil, 1991/1992, and Bavelas & Chovil, 1997) conducted the first systematic, quantitative description of a much wider range of facial displays in face-to-face dialogue. Twenty-four participants in 12 same- or mixed-gender dialogues first got acquainted and then talked freely within three tasks designed to elicit a variety of facial displays: telling about a close-call experience; retelling a conversation that involved a minor conflict with another person; and planning a nutritional meal of foods they both disliked. From the outset, Chovil excluded smiles, because they are so frequent and so varied (e.g., Ekman, 1985) that they would have dominated the lesser-known kinds of facial displays. Therefore her study asked: What do faces do in dialogue, other than smile?

Although Chovil (1989) described each facial movement physically, she focused on its communicative or linguistic *function* in the conversation, always in the immediate context of the words, the ongoing narrative, and other concomitant features such as prosody or hand gestures. Chovil identified several broad kinds of functions, each of which subsumed many specific variations. It is important to note that the inter-analyst reliability of such description was usually well above 90%, which compares very favorably with systems such as FACS. Chovil identified almost 1200 distinct movements that produced a change in one or more areas of the face. Of these, only 25% were adaptors, that is, non-communicative actions such as briefly licking or biting one's lips. It was clear that, even without adding in smiles, the participants were making conversational facial displays at a high rate.

In each dyad, the one who was speaking at any given moment made the greatest variety of facial displays. Chovil (1989) divided these *speaker displays* into syntactic and semantic functions. Her *syntactic* displays replicated all of Ekman's (1979) meaningful brow movements (described above), that is, *emphasizing* a single word, *underlining* a phrase, asking a *question*, and also added other movements analogous to the punctuation or organization of the dialogue (e.g., announcing the beginning of a new topic). These syntactic displays do not have a meaning specific to the topic of conversation, for example, an eyebrow emphasizer would look the same regardless of the word being emphasized. Instead of conveying topical information, they function to structure the content of the speaker's talk for the benefit of the addressee (analogous to punctuation and format in written communication). It was noteworthy that, although syntactic displays tended to be closely associated with prosody (e.g., when emphasizing or questioning), they were usually *not* redundant with words; for example, the speaker did not have to say explicitly "I want to emphasize this word (or point)" nor have to indicate syntactically that the utterance was a question. The syntactic facial displays served a useful function that was separate from but complementary to the spoken words.

Chovil (1989) also identified, for the first time, a large class of speakers' displays that she called *semantic*, that is, they illustrated some aspect of what the speaker was saying at the moment. One had been mentioned by Ekman (1979) as being part of a word search; Chovil called this a *thinking/remembering* face. However, the largest group of semantic displays depicted *personal reactions*:

> *Example 5:* In a close-call story, the speaker was describing a time when the family received an alarming phone call from her brother:
>
> Speaker: "The phone rings, [brief pause] my brother's on the phone, 'We're in an accident!' We're goin' [$_1$exaggerated intake of breath] "$_2$Ohhh my gawd!"

She began the description of the intense reaction of her family with a dramatic and extended intake of breath, accompanied by an alarmed and fearful face (display 1): her mouth rounded open, eyes wide open under raised brows, with pupils fixed straight ahead. This display ended just before she started to say "Ohhh my gawd!" and was replaced with a smile (display 2) during this phrase.[5] The fear display was a depiction of what her reaction might have been at the time of the accident, not of how she was feeling as she told the story later in the laboratory. As noted earlier, Ekman (1997) also distinguished these facial actions from affect displays because they are tightly synchronized with words. Such displays are timed to the structure of talk and not to the actual emotional state of the speaker. That is, it is highly unlikely that there was such a precise onset and offset of real-time fear and alarm, which turned equally precisely to happiness while she said "Ohhh my gawd!" Instead, as with many narrators of close-call stories, she was illustrating the fearful part while at the same time making fun of the fact that the fear turned out to be unjustified.

In addition to personal reactions, Chovil's (1989) speakers also displayed *portrayals* of other persons, for example, when describing the individual with whom

they were having the conflict. These facial displays were often part of what Clark and Gerrig (1990) called "quotations as demonstrations," in which the speaker demonstrates rather than simply describes what someone had said (or might say):

Example 6: While retelling a minor argument with her sister, the speaker portrayed her sister's reaction:

Speaker: "She's goin' ₁like "NO! ₂I don't want you to . . ."

At "like," the speaker screwed up her face, and precisely with "NO!" she shook her head and squinted her eyes in an angry expression (display 1), which was quickly replaced with a smile (display 2) when she continued with "I don't want you to . . .". Again, the facial displays were timed with words to match the rapidly changing moments and tenor of her story. As in Example 5, the meaning of the first ("angry") display was similar to the words being spoken at the moment; this redundancy served to add vividness and interest. However, the smiles in both examples were not redundant; they indicated that the speaker was presenting this information as humorous. Chovil found that 40% of semantic facial displays in her data (which excluded smiles) were non-redundant. These displays added closely timed, complementary information that did not appear at all in the words. As noted above, virtually all syntactic displays were non-redundant with verbal content.

The addressees' faces were not as active as the speakers', often remaining in an attentive configuration. Still, they often made relatively simple *back-channel* displays (in addition to smiles), such as raised eyebrows or pressing the lips together. They also depicted, more distinctly, their own personal reactions:

Example 7: In the meal planning task, the speaker named a food she disliked:

Speaker: "Thick slices of liver."

Addressee: "Yeah, oouu."

Precisely with "yeah," the addressee squinted his eyes and raised his upper lip in a disgust expression. We assume that the addressee meant that liver is disgusting.

The final major group of addressee displays in Chovil's (1989) data was *motor mimicry*, in which the addressee displayed a reaction that someone else (especially the speaker) might have had in the situation that the speaker was describing. We had studied motor mimicry earlier in social but not interactive settings (Bavelas, Black, Chovil, Lemery, & Mullett, 1988; Bavelas, Black, Lemery, & Mullett, 1986), but Chovil found its real home in the close-call dialogues, where addressees looked alarmed, fearful, in pain, etc. at precisely the moments when the narrator might have experienced those reactions. We will return to facial motor mimicry below.

We have described Chovil's (1989) research in some detail because it was the first and still the only intensive and extensive examination of what faces do, conversationally, in dialogue. She showed that such analysis, although in some sense

interpretive, could be just as reliable as physical descriptions. Finally, the functions she identified were clearly directed at the conversational partner, that is, they conveyed narrative information, structured it syntactically, or indicated how it was received. Often, these functions were only in the facial display, not in the words.

Although the three studies above were systematic, and in one or two instances were quantitative or even statistical, none were experiments with true independent variables. True experiments meeting the same criteria as the previously described gesture studies were rare: we found only two. Before describing those, we should point out again that we eliminated, for our purposes, otherwise interesting experiments that did not involve spontaneous dyadic face-to-face dialogue, studies of special populations, and studies focused on what Duncan (1969) called external variables, that is, external to the dialogue itself, such as hostility (Prkachin & Silverman, 2002) or power and gender (Hecht & LaFrance, 1998). Most of the studies we excluded examined the facial displays in isolation from words.

Visual Availability Chovil (1989; also described in Chovil 1991) investigated the effect of visibility (whether a conversational partner would see the facial display or not) in an experiment with addressees' facial displays as the dependent variable. There are a number of studies showing that, without a visible receiver to see them, facial displays are less likely to occur (cf. reviews by Chovil, 1997; Bavelas & Chovil, 2000), but most of these did not include a dialogue. Chovil (1989, 1991) analyzed the faces of participants who listened to the close-call story of another participant in one of four experimental conditions: (1) speaker and addressee were face-to-face; (2) speaker and addressee were sitting on either side of a partition; (3) speaker and addressee were in separate rooms, talking on the phone; (4) the addressee was alone, listening to a recorded close-call story, as if to an answering machine.

The dependent variable was the frequency (or log frequency) of motor mimicry displays. Motor mimicry was defined as an overt reaction by the addressee that was not appropriate to his or her own situation but would have been appropriate to the situation of the narrator or someone else in the story; for example, grimacing or displaying fear at a point in the story when the narrator might have done so:

Example 8: In one close-call story, the speaker was telling about being a passenger in a friend's truck:

Speaker: "He drove, he drove too close and we had one front wheel off the cliff, my side, and the truck was going over the edge."

From "my side" to "over the edge," the addressee displayed fear by raising her eyebrows and widening her eyes.

Analyst reliability for locating motor mimicry was high; the frequencies per story correlated highly, $r = .94$ for two independent analysts, one of whom was unfamiliar with the theory and predictions. The results supported the three major predictions: a significant linear drop in rate of facial displays over the four

conditions, as they became less and less social; a significant difference between the three dyadic conditions and the addressee listening alone, in which facial displays virtually disappeared; and significantly more displays in the face-to-face condition than the other three conditions, which lacked visual availability. Facial motor mimicry to the narrator's story occurred predominantly in face-to-face dialogue.

Effects of Preoccupied Addressees The other experiment was a broader study of addressees' responses to close-call stories, but because of the lack of other experiments and because it included motor mimicry we will describe it here. Bavelas et al. (2000) divided listeners' responses into *generic* responses, such as "mhm", nodding, or "yeah", and *specific* responses, which were tightly connected to the particular point of the story (e.g., Example 8). As in Chovil's (1989, 1991) experiment, most of the specific responses had a facial component, especially motor mimicry, although they also could involve gesturing or finishing sentences for the narrator. The analysts in this study were unaware of hypothesis or experimental condition; reliability checks revealed 95% agreement on whether a response was generic or specific.

The independent variable in Bavelas et al.'s (2000) two experiments was whether the addressee was attending naturally to the story or distracted by an assigned mental task. For example, when the listener was counting the narrator's words that began with the letter *t*, he or she was listening closely to the speaker's words but not to the narrative. When the listeners were thus distracted, their specific responses, including facial motor mimicry, virtually disappeared. One implication is that it is necessary for the listener to be tightly tuned in to the narrative to be able to make such responses. Further analysis revealed that the narrators who were telling their stories to distracted listeners told their stories more poorly, especially at what should have been the dramatic ending. The significant difference in quality of story endings is evidence of the importance of listeners' facial feedback for narrators.

The above two experiments on addressees' facial displays were the only ones that met our criteria. It is therefore striking that the literature lacks any experimental studies of what the face of the *speaker* does as part of dialogue – as well as more studies of addressees and of the relationship between the two participants. Although few in number, the studies available have shown that it is possible to analyze faces reliably, objectively, and quantitatively in terms of their conversational meaning (see also Chovil, 2005) and that experimental manipulation can reveal the functions of facial displays in dialogue. For those who are intrigued by facial actions that are timed to conversation rather than emotion, this is an unexplored and promising frontier.

TECHNOLOGICAL ADVANCES

Current and constantly developing technology greatly facilitates both dialogic experiments and analyses that were time-consuming, difficult, or even impossible to undertake in the past. Researchers do not have to depend on real-time observations of behaviors in face-to-face dialogue except, of course, as valuable sources for inspiration. In the laboratory, two or more tightly synchronized cameras permit a split-screen view of both participants and a permanent record of the exact timing of reciprocity between them. This level of synchrony in recording is essential because one of the impressive puzzles of face-to-face dialogue is that the participants often respond precisely to each other in less than simple reaction time (e.g., Bavelas & Coates, 1992).

Digitized video has replaced analogue: Analysts no longer have the fear and frustration of viewing fragile videotaped data repeatedly with a rewind or jog shuttle. Repeated frame-by-frame analysis is now easier, risk-free, and better at establishing the precise timing of the gestures, words, and other acts of both participants. For example, separating the stroke of a gesture (the purposive, meaningful part) from its preparatory and return phases is more straightforward and less labor-intensive than with earlier methods. The technological advantages are especially important for the analysis of conversational facial displays, which are so rapid that frame-by-frame analysis of both partners is essential for a full understanding of each event. We prefer software that does not impose a format for transcription and analysis, in order to be free to follow the data; we therefore rely on a playback system only (Broadway, http://www.b-way.com). However, there are free-ware systems that support both playback and annotation of video/audio data, such as ELAN (http://www.mpi.nl/tools/elan.html), ANVIL (http://www.dfki.de/~kipp/anvil/r45-7411), and Transana (http://www.transana.org), each with its advantages and disadvantages.

Presentation systems such as Power Point have made it much easier to demonstrate phenomena with video examples, so that the audience can see directly what the statistical summaries stand for. However, hard-copy publication formats are still a barrier. As in this volume, one can only describe illustrative excerpts in words; very few publication outlets have a supplementary CD or other video format. Indeed, one can speculate whether the relatively late development of an appreciation of conversational gestures and facial displays is, at least in part, because journals could present only still photos or drawings, which actually distort the phenomena by freezing them in time and out of context.

CONCLUSIONS

The focus of this chapter has been on the micro-social aspects of hand gestures and facial displays in face-to-face dialogue, especially the participants' use of these actions to reinforce, complement, or structure their verbal interaction. This aspect of social interaction is a recently growing area of research and, contrary to many preconceptions, is amenable to objective experimental investigation.

Our first obligation was to distinguish this particular subset from other interests in similar non-verbal behaviors (e.g., self-prompting hand gestures or facial expressions of emotion). To do so, we articulated four characteristics of conversational hand gestures and facial displays: (1) They occur in spontaneous face-to-face dialogue between real participants, where the dyad is the unit of

analysis, even when the actions of one person may be the dependent variable. (2) They are synchronous with co-occurring conversational elements, that is, they are closely timed to the words they accompany and to the specific moment of interaction at which they occur. (3) Like words, their meaning is highly context-dependent, in two senses: each has an immediate communicative function at a particular point in the conversation (rather than a stereotypic or universal meaning), and the meaning of each also depends on the integrated or composite message of which it is a part. (4) Our focus is on the social, interactive functions of these acts, as they shape and are shaped by the ongoing dialogue.

Research on hand gestures that meets the above criteria has been accelerating in the past decade. A rich group of experiments has documented the important role that hand gestures play in language as a social process. Moreover, these experiments illustrate the variety of tasks and variables that can be used to elucidate gestures' functions, as well as an equal variety of dependent variables to measure outcomes. Comparable research on facial displays is so far much more limited. However, knowledge from even the few systematic observational studies and experiments could be combined with the methodological lessons from gesture research to suggest new experiments just waiting to be done.

NOTES

1. Often the best way to understand gestures in transcription is to act them out with the words.
2. Reproduced as Figure 1 in Bavelas, Kenwood, Johnson, & Phillips (2002).
3. The review includes only studies involving adults, and only adults who were not from special populations, on the assumption that at this stage of research we need baseline data.
4. We will describe these facial actions with lay terms, see Ekman (1979) for descriptions in terms of FACS Action Units.
5. We have reanalyzed these examples from Chovil's original data to include smiles.

REFERENCES

Aronson, E., & Carlsmith, J. M. (1968). Experimentation in social psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., Vol. 2, pp. 1–79). Reading, MA: Addison-Wesley.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, *15*, 415–419.

Bavelas, J. B. (1990). Nonverbal and social aspects of discourse in face-to-face interaction. *Text*, *10*, 5–8.

Bavelas, J. B. (in press). The micro-social dimension of face-to-face dialogue. In S. Duncan, E. Levy, & J. Cassell, *Language in mind, body and context*. Amsterdam: John Benjamins.

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., 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. (1997). Faces in dialogue. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 334–346). Cambridge: Cambridge University Press.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163–194.

Bavelas, J. B., & Chovil, N. (2006). Nonverbal and verbal communication: Hand gestures and facial displays as part of language use in face-to-face dialogue. In V. Manusov & M. L. Patterson (Eds.) *The Sage handbook of nonverbal communication* (pp. 97–115). Thousand Oaks, CA: Sage.

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, *21*, 394–405.

Bavelas, J. B., Chovil, N., Lawrie, D., & 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., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*, 941–952.

Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, *52*, 566–580.

Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2006). Gesturing on the telephone: Independent effects of dialogue and visibility. Manuscript under review.

Bavelas, J. B., Hutchinson, S., Kenwood, C., & Matheson, D. H. (1997). Using face-to-face dialogue as a standard for other communication systems. *Canadian Journal of Communication*, *22*, 5–24.

Bavelas, J. B., Kenwood, C., Johnson, T., & Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, *2*, 1–18.

Beattie, G., & Aboudan, R. (1994). Gestures, pauses and speech: An experimental investigation of the effects of changing social context on their precise temporal relationships. *Semiotica*, *99*, 239–272.

Brunner, L. J. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology*, *37*, 728–734.

Buck, R. (1984). *The communication of emotion*. New York: Guilford Press.

Chafe, W. L. (1974). Language and consciousness. *Language*, *50*, 111–133.

Chovil, N. (1989). Communicative functions of facial displays in conversation. Unpublished PhD dissertation, University of Victoria, Victoria, BC.

Chovil, N. (1991). Social determinants of facial displays. *Journal of Nonverbal Behavior*, *15*, 141–153.

Chovil, N. (1991/1992). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, *25*, 163–194.

Chovil, N. (1997). Facing others: A social communicative perspective on facial displays. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 321–333). Cambridge: Cambridge University Press.

Chovil, N. (2005). Measuring conversational facial displays. In V. Manusov (Ed.), *The sourcebook of nonverbal measures: Going beyond words* (pp. 173–188). Hillsdale, NJ: Lawrence Erlbaum.

Clark, H. H. (1985). Language use and language users. In G. Lindzey & E. Aronson (Eds.),

The handbook of social psychology (3rd ed., Vol. 2, pp. 179–231). New York: Harper and Row.

Clark, H. H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Foxtree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, *84*, 73–111.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, *66*, 764–805

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62–81.

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

Darwin, C. (1965). *The expression of emotion in man and animals* Chicago. The University of Chicago Press. (Original work published 1872)

Duncan, S. J. (1969). Nonverbal communication *Psychological Bulletin*, *72*, 118–137

Ekman, P. (1979). About brows: Emotional and conversational signals. In J Aschool, M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology Claims and limits of a new discipline: Contributions to the Colloquium* (pp. 169–248). Cambridge: Cambridge University Press.

Ekman, P. (1985). *Telling lies*. New York: Berkeley Books.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.

Ekman, P. (1997). Should we call it expression or communication? *European Journal of Social Sciences*, *10*, 333–359.

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. (1978). *The Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press.

Fillmore, C. (1991). Pragmatics and the description of discourse. In P. Cole (Ed.), *Radical pragmatics* (pp. 143–166). New York: Academic Press.

Fowler, C A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, *31*, 307–319.

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, *26*, 489–504.

Fridlund, A. J (1997). The new ethology of human facial expressions. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 103–129). Cambridge. Cambridge University Press.

Furuyama, N. (2000). Gestural interaction between the instructor and the learner in origami instruction. In D. McNeill (Ed.), *Language and gesture* (pp. 99–117). Cambridge: Cambridge University Press.

Gerwing, J. (2003). The effect of immediate communicative function on the physical form of conversational hand gestures. Unpublished Masters thesis, University of Victoria, Victoria, BC.

Gerwing, J., & Bavelas, J. B. (2004). Linguistic influences on gesture's form. *Gesture*, *4*, 157–195.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgans (Eds.), *Syntax and Semantics: Speech Acts* (Vol. 3, pp. 225–242). New York: Academic Press.

Hecht, M. A., & LaFrance, M. (1998). License or obligation to smile: The effect of power and sex on amount and type of smiling. *Personality and Social Psychology Bulletin*, *24*, 1332–1342.

Hunnicutt, S. (1985). Intelligibility versus redundancy: Conditions of dependency. *Language and Speech*, *28*, 47–56.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp 207–227). The Hague: Mouton Publishers.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kess, J F. (1992). *Psycholinguistics: Psychology, linguistics and the study of natural language*. Philadelphia: John Benjamins Publishing Company.

Krauss, R M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication. What do conversational hand gestures tell us? In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 389–450). San Diego, CA: Academic Press.

Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, *37*, 1539–1553.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London: Routledge.

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

McNeill, D., & Duncan, S. D. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge: Cambridge University Press.

Nobe, S. (2000). Where do *most* spontaneous representational gestures actually occur with respect to speech? In D. McNeill (Ed.), *Language and gesture* (pp. 186–198). Cambridge: Cambridge University Press.

Özyürek, A. (2000). The influence of addressee location on spatial language and representational gestures of direction. In D. McNeill (Ed.), *Language and gesture* (pp 64–83). Cambridge: Cambridge University Press.

Özyürek, A. (2002). Do speakers design their co-speech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, *46*, 688–704.

Prkachin, K. M., & Silverman, B. E. (2002). Hostility and facial expression in young men and women: Is social regulation more important than negative affect? *Health Psychology*, *21*, 33–39.

Rimé, B., & Schiaratura, L. (1991). Gesture and Speech. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 239–281). New York, NY: Cambridge University Press.

Russell, J. A., & Fernandez-Dols, J. M. (1997). What does a facial expression mean? In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 3–30). Cambridge: Cambridge University Press.

Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 123–164). Mahwah, NJ: Lawrence Erlbaum.

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

Woods, J. (2005). New vs. given information: Do gestures dance to the same tune as words? Unpublished Honour's thesis, University of Victoria, Victoria, BC.

Yngve, V. H. (1970). On getting a word in edgewise. In *Papers from the sixth regional meeting of the Chicago Linguistic Society* (pp. 567–578). Chicago: Chicago Linguistic Society.