

Oxford Handbooks Online

Hand and Facial Gestures in Conversational Interaction

Janet Bavelas, Jennifer Gerwing, and Sara Healing

The Oxford Handbook of Language and Social Psychology (*Forthcoming*)

Edited by Thomas M. Holtgraves

Online Publication Date: Mar
2014

Subject: Psychology, Personality and Social Psychology

DOI: 10.1093/oxfordhb/9780199838639.013.008

[–] Abstract and Keywords

Conversational hand and facial gestures are an integral part of language use in face-to-face dialogue. Extensive research shows that conversational hand gestures are tightly synchronized with words to demonstrate anything that can be represented (directly or metaphorically) as size, shape, position, or action and that they are highly sensitive to the immediate communicative context. Although research on conversational (nonemotional) facial gestures is much more limited, they, too, are precisely timed with words and context to demonstrate anything that can be represented by a facial configuration—whether in the past, present, or future and even hypothetical or metaphorical. Both hand and facial gestures can also function as collateral communication (meta-communication). This chapter includes theoretical, methodological, and technical requirements for studying these gestures in conversational interaction.

Keywords: Conversational hand gestures, conversational facial gestures, language use, dialogue, conversational interaction, collateral communication

In face-to-face dialogue, interlocutors spontaneously combine words, prosody, hand gestures, and facial gestures to produce rapidly changing integrated messages (Bavelas & Chovil, 2000, 2006). This chapter presents a framework for studying these integrated messages as social communication and then an overview of what experimental research is discovering about how the visible elements of communication contribute to language use in conversational interaction.

Looking Closely at Conversational Hand and Facial Gestures

To appreciate the dynamic, fleeting quality of conversational hand and facial gestures, it is helpful to see (rather than to read descriptions of) as many examples as possible. This chapter uses sequential frame shots from our research videos, a format that may help the reader imagine how these actions actually looked (and may lead to noticing them more often in everyday life). In his advice on how to get the most out of examples presented in print form, McNeill (1985, p. 352) urged readers to act out the examples in order to experience the single coordinated action of motor–speech synchrony that the examples are intended to show.

Figure 1. presents three frames in a 6.75-second video excerpt (from Bavelas, Gerwing, & Healing, 2014). The speaker's task was to watch several scenes from the animated movie *Shrek 2* and then retell them to an addressee. She had just finished describing an attack on Shrek by the cat character (Puss in Boots). In frame 1, she began to describe the next scene, in which Shrek picked up the defeated cat by the back of the neck, lifted him close to his face, and the cat started begging for his life. By frame 2, she had raised her hand up in front of her own face, pinching her first two fingers and thumb as if lifting and then suspending the cat, presenting an image of Shrek holding the cat in front of his face. This hand gesture demonstrated information that was either left ambiguous in her words (e.g., the shape of Shrek's hand) or was missing from her words (e.g., that Shrek

suspended the cat in the air at face level).



Click to view larger

Figure 1. While describing a scene from the movie *Shrek 2*, the speaker made the hand gesture in frame 2 and the facial gestures in frames 2 and 3. The three frame shots cover a 6.75-second period. (The face of the addressee, who was sitting across a table from the speaker, is inset at the upper right.)

The speaker's face was also active. In frame 2, while saying "neck," she demonstrated Shrek's slightly gloating expression as he held the cat at his mercy. In frame 3, she began to quote the cat ("Oh I'm so sorry!") while wrinkling her forehead and pursing her lips, creating a worried and pleading look. Each of these facial gestures portrayed one of the characters and illustrated her narrative for the addressee. Like her hand gestures, each added relevant details that her words did not convey (e.g., Shrek's expression as he held the cat, then the cat's apprehension).

Some Definitions and Assumptions

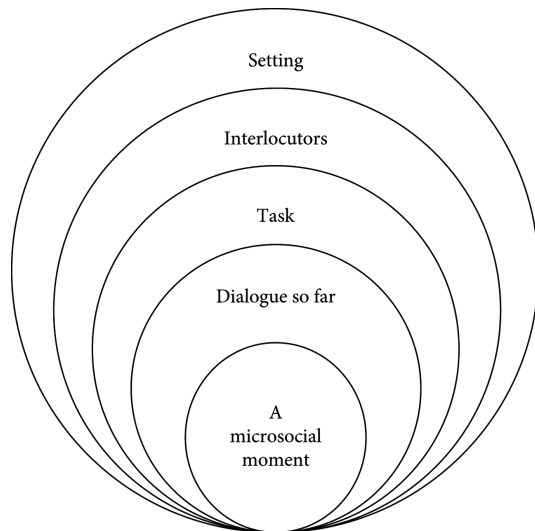
The close analysis of the three frames in Figure 1. illustrates the defining characteristics of conversational hand and facial gestures. Each gesture complemented the speaker's words and was timed precisely with the relevant words. Both hand and facial gestures occur as part of a spontaneous stream of speech, and they are tightly synchronized in both *meaning* and *timing* with the accompanying words. The speaker's hand gesture in frame 2 began and ended exactly with the words it was related to. The face is highly mobile, so speakers can make facial gestures even more quickly than hand gestures—at the rate of a single word or even syllable. In fact, Figure 1. includes only two of the speaker's facial gestures in the excerpt; there were two more (one before and one after frame 3). Altogether, in 4 seconds, she made four different facial gestures related to what she was saying.

The synchrony between visible and audible elements of communication results in efficiencies that could not be matched by words alone. Even if words could convey exactly the same information, the speaker would be limited to presenting it in a sequential, linear manner (e.g., "Shrek picks him up by the nape of his neck *with his thumb and two fingers and holds him up at eye level in front of his face, and looks at him in sort of a gloating way.*") The additional words needed to convey information that had been conveyed by the speaker's gestures would require a much longer description than the original integrated message.

This chapter treats hand and facial gestures—but not all nonverbal behaviors—as part of language use in face-to-face dialogue. There are several types of nonverbal behaviors that are *not* part of language use, including those that are involuntary or reflexive (e.g., blinking), instrumental (e.g., turning a door key), static postures (e.g., sitting with crossed arms), or behaviors that are just as likely to occur when no one else is present (e.g., squinting in bright light). In addition, the hand and facial gestures discussed in detail in this chapter are those that occur within the context of conversation, not those occurring in nonspeaking contexts such as hand signals (e.g., the hitch-hiking sign) or still photos.

The Role of Context in Meaning

In Figure 1. the meaning of each of the speaker's hand and facial gestures depended on the set of nested contexts in which they occurred. Figure 2. presents one way to visualize these contexts. The setting was a university psychology experiment for class credit. The interlocutors were unacquainted students. Their task was for the speaker to tell the addressee the scenes from *Shrek 2*. Their dialogue so far consisted of their conversation up to this point, which culminated in a particular *microsocial moment* in the dialogue (Bavelas, 2007), for example, demonstrating Shrek picking up the cat. In a different set of nested contexts, a hand and facial gesture combination similar to the one in frame 2 would have an entirely different meaning: imagine the meaning of these gestures if she had been in a pub talking with friends about how much fun it would be to drop a water balloon on her boss.



Click to view larger

Figure 2 . Image of the nested contexts that successively determine the meaning of words, as well as hand and facial gestures in face-to-face dialogue.

Notice that the same layers of context also determine the meaning of any word or phrase in a dialogue. For example, in this particular context, “the cat” unambiguously meant the Puss in Boots character rather than any other feline. A meaning-based analysis of hand gestures, facial gestures, and words must always take into account these contextual levels—which is what everyone does every day, in every dialogue.

Hand and Facial Gestures as Demonstrations

We have proposed (Bavelas, Gerwing, Sutton, & Prevost, 2008; Bavelas et al., 2014) that both conversational hand and facial gestures fit into a larger view of language, communication, and the general theory of signs initiated by Charles Peirce (e.g., Buchler, 1940). Peirce distinguished between symbols (e.g., words), indices (e.g., a weather vane), and icons (e.g., a picture). Clark and Gerrig (1990; Clark, 1996, chapter 6) put these three categorical distinctions into a social context by emphasizing how they function in everyday language use: with symbols, a speaker can describe something to an addressee by using words in their conventional, arbitrarily established meanings. Using indices, a speaker can also indicate something for the addressee by pointing at it, either physically or with a demonstrative pronoun (e.g., “I want *that* one”). Using icons, a speaker can demonstrate something by presenting a visual or auditory version of it. Demonstrations include direct quotations (Clark & Gerrig, 1990), metaphors, conversational facial gestures, and conversational hand gestures (Bavelas et al., 2014).

Demonstrations share several defining features. Rather than having standardized meanings, they “work by enabling others to experience what it is like to perceive the things depicted” (Clark & Gerrig, 1990, p. 765). They do so by resembling their referent in some way. Although a demonstration resembles its referent, it is also a selective transformation of the literal or actual properties of the referent; it selects some features and leaves out others. The gestural demonstration in frame 2 of Figure 1. presented the speaker’s version of the scene she had seen. Her hand resembled Shrek’s hand holding the cat, but it was not an exact replication. She selected (and emphasized) the image of Shrek holding the cat between his fingers in front of his face (but did not include Shrek bouncing his hand slightly, as shown in the movie). In the same frame, her face resembled how Shrek looked, but she selected a mildly gloating and victorious version and left out the anger that was clearly present in his face in the actual movie scene.

A Functional, Communicative Approach

This chapter concerns the communicative functions of hand and facial gestures in social interaction. First, the focus is on *function* rather than *type*. That is, we will be asking “What is this gesture doing in this dialogue?” rather than “What category does this gesture go in?” A functional approach leads to terms such as *demonstrated*, *portrayed*, *presented*, *complemented*, and *marked*, all of which direct attention to the function of the gesture in its

microsocial context. Although hand gestures can be classified as *iconics* or *illustrators*, such categorization abstracts them from their interactional context. Second, our focus on communicative function steers away from questions such as “What does this facial gesture tell us about what the individual is feeling?” Inferences about the cognitive or emotional processes associated with a hand or facial gesture are distinct from examining what it is doing in the interaction. Functions are not exclusive: actions that are communicative may have cognitive or emotional functions as well. Although there is no logical reason to privilege one function over another (much less to set them in opposition), they do require quite different research methods. In any case, contemporary research has moved past the adversarial debate about whether gestures communicate to the more interesting question of how they do so, which opens up entirely new possibilities for research and theorizing.

Methodological Issues

As implied in the previous paragraph, experiments on the communicative functions of conversational hand and facial gestures require a research design in which one condition is a face-to-face dialogue. The next sections explicate this requirement and critically examine some conventional barriers to meeting it.

Conversational Interaction as Face-to-face Dialogue

Conversational hand and facial gestures serve their communicative functions within face-to-face dialogues. Clark (1996, pp. 9–10) specified the defining characteristics of face-to-face dialogue, which stress two key principles: first, a face-to-face dialogue is *unmediated*, which means that the interlocutors are in the same physical environment, able to see and hear each other, and able to produce and receive at once and simultaneously. Second, the interlocutors can freely interact and collaborate with each other, moment by moment. In Clark’s terms, they are acting as themselves (e.g., not as a confederate), determining what actions to take when (e.g., without a script), and acting extemporaneously, in real time. These features provide essential criteria for research designs that will yield data relevant to face-to-face dialogue.

True Dialogues Versus Quasi-Dialogues

Traditionally, experimental social psychologists have shown a methodological preference for what Bavelas and Healing (2013) called *quasi-dialogues*, in which the participant interacts with someone who is acting under instructions that constrain his or her actions (e.g., a confederate, the experimenter, or another participant). These interactions do not fit Clark’s (1996) definition of true dialogues because one of the participants is not acting as him- or herself, is not determining his or her own actions, and is not able to extemporize. This section explores the reasons underlying the choice of quasi-dialogues. (See also Kuhlen & Brennan, 2012.)

First, many scholars have pointed out that social psychology has remained largely the study of the individual rather than the dyad (e.g., Bavelas, 2005, 2007; Clark, 1985; Danziger, 1990; Sears, 1951; Thibaut & Kelley, 1959). As Danziger pointed out, the line between the person and the environment has historically been drawn tightly around the individual, so that the interlocutor in a dialogue is “outside” this line—just another part of the environment. However, when the interest is in how the speaker’s actions function in dialogue, then the unit of study must include the dialogue itself.

Even social psychologists with an interest in communication during social interaction often use quasi-dialogues, perhaps because of an erroneous application of the principle of *reductionism*. Luria (2004, p. 537) described reductionism as the assumption that “the basic goal of science is to reduce complex phenomena to separate simple parts, and that such reduction provides significant explanations of phenomena.” By this principle, a true dialogue is a complex phenomena that can be better studied by reducing it to one individual in a controlled interaction. However, Luria went on to question this assumption:

Although the philosophy of reductionism was accepted as a general principle in the natural sciences and psychology, there are grounds to suppose it may be false. To study a phenomenon, or an event, and to explain it, one has to preserve all its basic features: one must be able to describe their rules and their mechanisms without the loss of any [characteristics of the phenomenon]. It can easily be seen that reductionism may very soon conflict with this goal. (p. 537)

In their critical examination of reductionism, Reber, Allen, and Reber (2009, p. 663) pointed out that the problem is not with reductionism as a scientific principle per se but with its overgeneralization to all topics in all fields. Reducing a complex phenomenon to its smallest possible units makes sense if the smaller units are the topic of interest. But studying the smaller units will not be informative if a more complex phenomenon of interest is lost in the process. For example, if one is interested in learning about salt (sodium chloride), then studying the properties of sodium (a metal so reactive that it almost explodes in water) and chlorine (a poison gas) would not be very enlightening. If the phenomenon of interest is social interaction in dialogue, then studying individuals is similarly unenlightening.

For experimental researchers, reductionism also underlies a second justification for the use of quasi-dialogues, namely, the assumption that the complexity of a real dialogue precludes experimental control. The fear is that if one interlocutor acts spontaneously and the other is also free to react spontaneously, experimental control will be lost and the ensuing variability will overwhelm any differences between conditions. However, face-to-face dialogue is inherently well organized. In everyday life, even strangers regularly accomplish complex tasks together in dialogue (e.g., giving directions, discussing purchases, exchanging technical information, making a diagnosis, or conducting an interview). The numerous experiments reviewed in this chapter demonstrate that allowing two individuals to respond extemporaneously to each other within an assigned task does not lead to unanalyzable data. Far from being chaotic, dialogues are intrinsically orderly; those of us who microanalyze dialogue are constantly awed by its precision and orderliness.

Some Technical Issues

As a practical matter, observing the precision of face-to-face dialogue requires digitized video with both interlocutors on screen at all times (e.g., using split-screen) plus software for frame-by-frame analysis and annotation (e.g., ELAN, <http://tla.mpi.nl/tools/tla-tools/elan/>; Wittenburg, Brugman, Russel, Klassman, & Sloetjes, 2006). Using comprehensive operational definitions, analysts can capture the details of dialogue with high interanalyst agreement. (Gullberg's, 2010, informative outline of methodological concerns in gesture research in second language acquisition is relevant to gesture research more widely, especially pp. 86–91.)

Conversational Hand Gestures

Imagine you are quietly observing two people having a conversation in a restaurant. You can't hear their speech, but you can see that they make a variety of hand movements. Not all of these movements are conversational hand gestures. First, some are instrumental actions involving objects (e.g., cutting their food with a knife and fork) or their own body (e.g., pushing hair behind an ear). These actions are often called *object-* and *self-oriented adaptors*, and when they serve purely instrumental functions, they have no communicative meaning. Second, you may see one person point to a corner of the restaurant, prompting the other diner to look in that direction as well. These *pointing (indicating or deictic)* gestures function to establish a joint focus of attention. Third, one person at the table catches the server's eye and holds one palm up while pretending to write on it with the other hand. If you are familiar with the cultural norms of the gesturer, you understand this movement as a request for the bill. If the server nods and then comes to the table with the bill, the gesture was effective. These hand signals, which are usually called *emblems*, have stereotypic forms and culturally specified meanings so they can function as messages outside of conversation. Aside from instrumental actions, pointing, and emblems, you see both diners moving their hands as they talk. These improvised movements seem to mean something, but you won't know their meaning without hearing the diners' words because these are likely to be *conversational (or co-speech) hand gestures*, as defined in the opening paragraphs of this chapter. (For a thorough review of the history and variety of gestures, see Kendon, 2004.)

Conversational hand gestures have at least three phases (Kendon, 1980). Most gestures begin with a *preparation phase*, lifting or moving the hands into position. The meaningful depictive action that follows is the actual *gesture stroke*. After the stroke, there is often some *retraction*, as the hands move back to a resting position or prepare for the next gesture. Immediately before or after the stroke, there may also be a brief *hold*, in which the hands pause in position. These sequences occur smoothly, in a matter of seconds or less, and they are tightly synchronized with the stream of speech.

The following sections focus, first, on what kinds of information the gesture stroke can contribute to a dialogue and, second, on how social factors affect them in a dialogue. The studies referred to in this section were experiments in which at least one condition fulfilled all of the criteria for a face-to-face dialogue.

What Can Hand Gestures Demonstrate?

Hand gestures are suited to depicting some kinds of information (e.g., shape, size, or direction) and not others (e.g., color, flavor, or names). Still, as will be seen in this section, the variety of possibilities is surprising. Most of the examples are from spontaneous conversational hand gestures made during laboratory experiments using dialogic tasks designed to elicit gestures (e.g., retelling a narrative, describing an object, or explaining to an addressee how to do something). They are highly varied because each was a demonstration specific to a particular micromoment in a dialogue. Recall that demonstrations do not have standardized conventional meanings. Instead, they resemble their referent and require the addressee to recognize this particular hand action, in this context, as a selective version of the referent.

Hand Actions

Gestures are well suited to demonstrating hand actions that are real or hypothetical; literal or metaphorical; done in the past, present, or future; and representing one's own or anyone's hands. In an experiment by Gerwing and Bavelas (2004), speakers first used a set of objects and later demonstrated their actions to an addressee (without the object). One of these objects was a woven tube-like toy that trapped the users' index fingers when they tried to pull them out. When telling their addressee about it, speakers usually gestured their actions, often by holding their index fingers in a straight, horizontal line with the tips touching, so the addressee could imagine a tube holding the fingers in this position. Sometimes they pulled their fingers back and forth to demonstrate how the fingers were (or could have been) trapped in the tube. Another object was a propeller-like toy that could be launched by spinning the upright stick between vertical palms. Speakers later gestured how they had launched the toy by moving one vertical palm rapidly past the other. In a study by Furuyama (2000), speakers taught addressees how they should fold an origami figure—without any paper. The teachers used gestures to demonstrate making the various folds. Thus, the addressee had to see the hands as though they were manipulating a piece of paper that was not actually there.

Speakers also demonstrate hand actions other than their own, as when the speaker in Figure 1. used her hand to represent Shrek's hand. The addressee had to see the speaker's hand as if it were Shrek's. In another experiment (Bavelas et al., 2008), speakers had to describe a picture of a woman holding a fan in front of her chest. Speakers often included a gestural demonstration of the fan, as if holding a fan themselves.

Finally, gestures of hand actions can be metaphorical. In an experiment by Rowbotham, Holler, Lloyd, and Wearden (2011), interviewers asked participants to describe a recent pain experience. Although pain is an intangible, subjective experience, the participants often used gestures to demonstrate various aspects of their pain. One person said, "It feels like *they're [sic] just sat there with like a hammer, hitting me*, that's how it feels." During the italicized words, the speaker gestured as if holding a hammer and hitting herself repeatedly on her temple (p. 8). Her demonstration required the addressee not only to "see" this action but also to appreciate that the headache felt as if the hammer were pounding her temple.

Spatial Features

Hands are not limited to portraying hand actions. Speakers can use their gestures to demonstrate physical features of an object such as its shape, size, or orientation. For example, some speakers in the Gerwing and Bavelas (2004) experiment used their hands to depict the propeller toy itself. Figure 3. presents one such pair of gestures: in frame 1, precisely as the speaker said "a rod," she began to bring her upright index finger straight down, drawing a vertical line for the rod. As she said "then two" (in frame 2), she drew a horizontal line to represent the propellers at the top of the rod. In frame 3, she held her gesture at the end of the horizontal line while she finished saying "propellers."



[Click to view larger](#)

Figure 3 . While describing a propeller toy to her addressee, the speaker made three related hand gestures. The three frame shots cover a 1.4-second period. (The addressee was sitting directly across from the speaker and filmed in split screen, indicated here by a dotted line.)

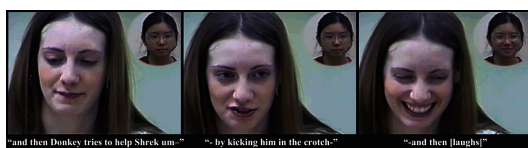
In other experiments, speakers have used gestures to demonstrate the shape of knotted pipes, dome-shaped roofs, or bridges (Holler & Stevens, 2007; see their figures 1–3, pp. 17–19), the evolving shape of an origami figure (Furuyama, 2000), and the shape and size of distinctive features of an unusual dress (Bavelas et al., 2008; see their figures 2–4, pp. 504–507).

Spatial gestures can also represent *virtual spaces*, that is, when the referent does not actually exist. Bavelas, Gerwing, Allison, and Sutton (2011) asked participants to design a floor plan together without pencil and paper. They “drew” their plans on the table with their gestures, demonstrating the location, size, and shape of virtual rooms that did not actually exist.

A single gesture can be *dense* with information, that is, demonstrating more than one feature of the referent (Gerwing & Allison, 2011). In several of the examples described in this section, gestures depicting shape also conveyed accurate information about the referent’s size (e.g., Holler & Stevens, 2007; Bavelas et al., 2008) or other features such as location (Gerwing & Allison, 2009, see their Figure 1. p. 313) or orientation.

Motion

Hands can move in almost any direction and are well suited to demonstrating aspects of motion. In an experiment by Parrill (2008), participants retold a cartoon in which a cat had swallowed a bowling ball, which then began rolling down a hill, transporting the helpless cat along with it. Speakers often abstracted and selected aspects of this motion to demonstrate with their gestures. For example, Parrill’s figures illustrate gestures that depicted the direction the cat rolled (moving the hand diagonally downward; her Figure 1. p. 287), the rolling manner of the movement (tracing a circular motion in one place; her Figure 4. p. 289), as well as both direction and manner simultaneously (tracing a circular motion moving downwards; her Figure 5. p. 290). In Özyürek’s (2002) experiment, participants used gestures to show the direction of motion as characters went into or out of buildings or across a street (Özyürek, appendix, p. 703). Gestures depicting motion can also be metaphorical; for example, representing the lessening of pain by moving the hand downward (Rowbotham et al., 2011, table 4).



[Click to view larger](#)

Figure 4 . The speaker’s and addressee’s faces in frame 3 showed their personal reactions as the speaker was telling about a scene from *Shrek 2*. The three frame shots cover a 6.2-second period. (The face of the addressee, who was sitting across a table from the speaker, is inset at the upper right.)



[Click to view larger](#)

Figure 5 . The addressee (inset at the upper left) was sitting directly across from the speaker, listening to the speaker’s close-call story. In frame 2, his facial gesture conveyed a personal reaction to a sudden danger in the story. The three frame shots cover 2 seconds.

Relative Positions

Some gestures demonstrate both the referent and its relationship to something else, such as another gesture, the gesturer's body, or a nearby object. In discussions about floor plans, Gerwing and Allison (2009) found two-handed gestures that demonstrated the relative positions of two or more rooms. For example, one participant gestured and held the location of the entrance with her right index finger while showing the proposed location of the kitchen with her left hand. She then held her left hand in position while moving her right hand over to the proposed location of the living room (Gerwing & Allison, 2009, their Figure 1. p. 313).

The rest of a speaker's body provides an additional resource: speakers can use the relationship between their hand(s) and body to demonstrate relative positions. In Figure 1. of this chapter, the speaker used the relationship between her gesture and her face to depict where Shrek suspended the cat while listening to his pleading. In another experiment (Bavelas et al., 2008), speakers demonstrated the distinctive features of a dress. Men as well as women often placed their gestures in relation to their own body. Some outlined the bodice of the dress as a deep "V" over their own chest (Bavelas et al., 2008, their Figure 4. p. 507) or showed the unusually wide shape of the skirt by extending their hands out from their own waist (Gerwing & Allison, 2011, their Figure 2. p. 311). In descriptions of their pain experiences, Rowbotham et al.'s (2011) participants also placed their gestures in relation to their body; for example, one participant put her palms over her lower stomach while leaning forward to show how she positioned herself as a reaction to the pain (Rowbotham et al., 2011, table 4, p. 17). Gestures can also build on the location of previous gestures, as if the previous gesture left a "residue" behind that could be used by later gestures. For example, in Figure 3. of this chapter, notice that the participant gestured the propellers directly over where she had previously placed the top of the rod. Healing and Gerwing (2012) analyzed the gestures of speakers who described a picture of a maze-like route. They found that speakers often linked contiguous sections of the route by gesturing one section, pausing, and then beginning the next gesture from the same position.

Gestures and Words

The previous examples show that gestures are particularly suited for many—but not all—kinds of information; for example, although they can demonstrate something, they cannot name it. Words and gestures are useful for communicating different kinds of information, and Bavelas and Chovil (2000, 2006) proposed that they work together to form integrated (but not necessarily redundant) messages. A semantic feature approach (e.g., Beattie & Shovelton, 1999) can analyze how information is distributed between words and gestures, and it yields reliable quantitative evidence of how their integration works. One application of this approach examined how dyads used gestures and words when designing floor plans (Gerwing & Allison, 2009). Gestures were significantly more likely to contribute spatial information (i.e., size, shape, and relative location of the rooms) whereas words contributed categorical information (e.g., the name of a room). One speaker said "there's a kitchen" while pointing to a space adjacent to where she had located the entrance. The words named the room, and the gestures specified its location. To understand the speaker's message, the addressee had to integrate speech and gesture. A follow-up study on dress descriptions (Gerwing & Allison, 2011) reported a similar pattern of integration: in face-to-face dialogues, speakers' gestures were significantly more likely than their words to convey spatial information such as the size and shape of the skirt. When describing their pain experiences, participants in Rowbotham et al. (2011) were more likely to put information about the location and size of their pain in their gestures rather than in their words, while putting information about the cause, the effect, and their awareness of the pain significantly more often in their words.

Conversational Factors that Affect Gestures

The physical capacities of hands to demonstrate referents are by no means the only determinants of what gestures actually do in a dialogue. The form of a gesture is also determined by its function in the dialogue at the precise moment it occurs. Most of the experiments cited in the preceding sections will reappear here because their primary purpose had been to show how factors in the social interaction can affect how and when the interlocutors use gestures.

Mutual Visibility

A frequent experimental manipulation in gesture research consists of varying whether or not the speaker and addressee are mutually visible (e.g., speaking face-to-face vs. through a partition). Since Cohen and Harrison's

initial experiment (1973), researchers have hypothesized that, if the rate of gesturing decreases when the addressee is not visible, then gestures serve a communicative function; if not, then they serve a cognitive function (e.g., helping the speaker formulate utterances).

However, there are several reasons to doubt the utility of this design for such a broad hypothesis. First, the results of numerous experiments on the effect of mutual visibility on overall gesture rate have been mixed, with half finding that mutual visibility leads to a significantly higher rate (Alibali, Heath, & Myers, 2001; Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss, Dushay, Chen, & Rauscher, 1995; Mol, Krahmer, Maes, & Swerts 2009a, 2009b) and the other half finding no significant difference in overall rate (Bavelas et al., 1992, 2008, 2014; De Ruiter, Bangerter, & Dings, 2012; Holler, Tutton, & Wilkin, 2011; Pine, Burney, & Fletcher, 2010; Rimé, 1982). Two of the latter studies found both results, depending on the kind of gesture studied (Bavelas et al., 1992; De Ruiter et al., 2012). The second problem is that all of the experiments reporting a significantly higher rate of gesture in the visibility condition were quasi-dialogues, whereas the other group were all true dialogues. Clearly, the results of the former do not generalize to the latter. (See Bavelas & Healing 2013, for a discussion of possible reasons.)

Finally, several experiments have shown that overall rate measures do not capture the effects of mutual visibility on gesture use: speakers make qualitatively different gestures when addressees can see them than when they cannot; see Bavelas and Healing (2013, pp. 76-79, and Table 5). For example, mutual visibility affects whether or not gestures convey essential information. When their addressee will see their gestures, speakers are significantly more likely to convey information in them that does not appear in words. When their addressee will not see their gestures, speakers' gestures are less likely to add any additional information; that is, they tend to be redundant with the words (e.g., Bavelas et al., 2008; de Ruiter et al., 2012). Bavelas et al. (2008) also found that speakers' gestures in face-to-face conversations were significantly larger than those on the telephone. Healing and Gerwing (2012) found that, when describing the maze-like route to a visible addressee, speakers were significantly more likely to link sequential gestures into a cohesive whole, maintaining their relative positions. When their addressee could not see them, gestures for sequential features were not linked but instead tended to pile up in one area. Other research has shown that some specific gesture functions require visibility: interlocutors who could see each other were more likely to mimic the form and meaning of each other's gestures (Holler & Wilkin, 2011).

Shared Social Space

Just as speakers can use their own body as a resource for contributing to gestures' meaning, they can also use the space that they share with their addressee. Özyürek (2000, 2002) compared experimental conditions in which speakers were seated in different spatial configurations with their addressees. For example, they might be seated directly across from the addressee or talking to two addressees in a triangular configuration (see examples in Özyürek's, 2002, Figure 1. p. 692, and Figure 2. p. 698). When demonstrating how a cartoon character moved from one place to another, the speakers displayed sensitivity to how the relative seating positions of their addressees and themselves would influence the meaning of their gestures. For example, they changed the direction of their gestures to match whether the gesture was moving *into* or *out* of the shared space. Interlocutors can even share their gestural space. Furuyama's (2000) study on origami instruction reported that instructors and learners often collaborated with their gestures in the same physical space, pointing to or acting on a single "virtual" origami figure that was not actually in that space.

Common Ground

It is well established for words that speakers are sensitive to whether or not they share common ground with their addressee (Haviland & Clark, 1974). For example, they provide fuller, more informative verbal descriptions when they know the addressee is unfamiliar with the information. Gerwing and Bavelas (2004) showed a similar effect for gestures by manipulating common ground between participants. Speakers described the same object twice: once to an addressee who they knew had just manipulated the same object and once to an addressee who they knew had not. The gestures that speakers used to identify the object for the addressee who did not share common ground were judged to convey more information, be more complex, or be more precise than those the speaker used to identify the object for the addressee who did share common ground (e.g., their figures 3 and 4, pp. 169–170). Holler and Stevens (2007) found an effect of common ground on the relationship between the gestures and words that speakers used to convey the size of several disproportionately large objects. They found that speakers who described these objects to addressees who had not seen them used gestures significantly more often than

words. Furthermore, their gestures depicted the objects' sizes more accurately than gestures used in the common ground condition. Conversely, when participants shared common ground, speakers used words to refer to the objects significantly more often than they used gestures.

As their dialogue unfolds, the interlocutors accumulate more common ground. Speakers can introduce *new* information that, when they refer to it later, becomes *given* information and part of their common ground. Gerwing and Bavelas (2004) found that gestures for an object would change as the information depicted went from new to given. For example, the gestures would change from large and precise to smaller and vaguer. These changes in gestures parallel what happens with words: new words that become given change from noun to pronoun or from a distinct pronunciation to a shorter and less distinct pronunciation. Fowler (1988) pointed out (for words) that such attenuation is not sloppiness but rather serves the communicative function of marking the status of the information in the dialogue. We propose that attenuation serves the same function for gestures.

Grounding with Gestures

The accumulation and coordination of common ground (*grounding*) is an interactive process during which participants give each other evidence that they have understood each other "well enough for current purposes" (Schober & Clark, 1989, p. 228). Several researchers have examined the role gestures play in grounding by analyzing sequences of immediately contiguous utterances that involved gestures. Clark and Krych (2004) varied visibility to examine the role of gestures in addressees' responses to their speaker's instructions. In each pair of participants, the speaker had a model constructed from Lego blocks and instructed the addressee (who could not see the model) how to build it. In the pairs who could see each other, addressees often responded with gestures that demonstrated their current understanding of the instructions. Speakers monitored these gestural signals of understanding, often following up by modifying their instructions, sometimes in mid-utterance. In their study of pairs designing floor plans together, Bavelas et al. (2011) focused on responses to the 552 gestures that had contributed information that was missing from the words (i.e., were nonredundant). The vast majority of addressees' responses to these gestures indicated understanding of the information. Moreover, the majority of speakers' follow-up responses accepted that the information in the gesture had become common ground (e.g., speakers presented new information that presupposed understanding of the previously presented information). These and other studies that examined the relationship between interlocutors' behaviors (e.g., Furuyama, 2000; Holler & Wilkin, 2011) represent necessary next steps in understanding how gestures are integrated sequentially into dialogue.

Dialogue Versus Monologue

The use of conversational hand gestures seems to be tied to dialogue, regardless of visibility. Bavelas et al. (2008, 2014) showed that a major determinant of the overall rate of gesturing is whether the context is a dialogue or a monologue. They created three experimental conditions: a face-to-face dialogue, a dialogue on the telephone, and a monologue into a microphone. The telephone dialogues met the criteria for a dialogue (summarized earlier) except that the interlocutors were not in the same room or able to see each other. In both experiments, the two dialogue conditions elicited a significantly higher rate of gestures than the monologue condition, and these rates were not significantly different from each other. Speaking in a monologue suppressed the gesture rate to about one-third of the two dialogues.

Gestures as Collateral Communication

Clark (1996, pp. 241–243 and 255–257) proposed that dialogue has two parallel tracks. The primary track is the topic of the conversation; the secondary track is about the process of having the dialogue (e.g., "OK, I understand"), which is meta-communication or *collateral communication*. We propose that hand gestures can also function as collateral communication. For example, when speakers change the form of their gesture to mark it as common ground, their depiction of the referent is the primary track, and its attenuated form is the secondary track (in this case, marking the "given" status of this information in the dialogue). *Entrainment* is another form of collateral communication. Clark and Wilkes-Gibbs (1986) and Schober and Clark (1989) have illustrated verbal entrainment: when describing Tangram figures, interlocutors who could not see each other entrained on the same arbitrary verbal descriptions of the figures they were working with. Using the same task with interlocutors who could see each other, Holler and Wilkin (2011) found gestural entrainment. For example, addressees often mimicked the speaker's gesture to show that they had understood which Tangram the speaker was referring to

(e.g., their figures 2 and 3, p. 145). Their mimicry was not conveying information about the topic (i.e., identifying the figure); it was instead collateral communication, showing mutual understanding of that information.

Some gestures operate entirely as collateral communication. Bavelas et al. (1992) and Bavelas, Chovil, Coates, and Roe (1995) identified a subset of gestures with *interactive* functions. These gestures do not convey information about the topic of the conversation. Instead, they refer directly to the addressee, in two senses: (a) The gesture is physically oriented directly at the addressee with no other depictive features, and (b) the gesture's meaning refers to the addressee and to the process of dialogue itself. Examples include moving the fingers or palms toward the addressee as if metaphorically delivering information to the addressee or flicking the index finger toward the addressee to cite a previous contribution (equivalent to "as you just said..."). (See sketches of some interactive gestures in Bavelas et al., 1992, their Figure 1. pp. 474–475, and in Bavelas et al., 1995, their Figure 1. p. 396.) A series of experiments designed to confirm their interactive function showed that these gestures occurred at a significantly higher rate in the following conditions: (a) when speakers were talking to an addressee rather than narrating alone to a camera (Bavelas et al., 1992), (b) when speakers were talking to a visible addressee rather than to one who could not see them (Bavelas et al., 1992), and (c) when interlocutors who could see each other were working together in a back-and-forth dialogue rather than contributing in alternating monologues (Bavelas et al., 1995). Two microanalyses further confirmed the hypothesized functions of interactive gestures (Bavelas et al., 1992, pp. 483–486; Bavelas et al., 1995, study 2).

Summary of Hand Gestures

An examination of the dates of the publications just sampled reveals that experimental research on hand gestures in dialogue is growing rapidly. In addition, the International Society for Gesture Studies with its regular conferences, the journal *Gesture*, and other publications in a wide range of refereed outlets all suggest that the field is a vital one, with a wide variety of research foci and established methods. Even better, the research is opening up new questions about hand gestures in conversational interaction.

Conversational Facial Gestures

Let's go back to people-watching in a restaurant. If you are paying close attention to the diners' faces, the first thing you would probably notice is that faces bear little resemblance to still photographs. With their numerous independent muscle groups, human faces are extraordinarily fast and flexible. Ekman and Friesen (1978) identified more than forty different *action units* corresponding to activity in distinct facial muscles or muscle groups, which account for the ability of the face to move flexibly and rapidly. You may also notice that most of these movements seem to fit in the same three broad functions as hand gestures.

First, all individuals will have facial movements that are *self-oriented adaptors* such as blinking, chewing, or swallowing; these actions serve instrumental functions (e.g., keeping the eyes moist or eating). Second, you may also notice one person directing a facial action toward someone at a distance, even though the two of them are not engaged in speaking with each other. For example, a diner may stare steadily across the room until the server looks up and makes eye contact; then the diner might smile slightly and raise his or her eyebrows expectantly to confirm a request for attention. These facial actions in nonspeaking contexts are (like hand gestures in the same context) called *emblems* (Ekman, 1977).

Finally, when you focus on two diners who are talking with each other, you will probably see a much wider range of facial actions: smiling, frowning, tilting and moving their heads, furrowing their eyebrows, widening their eyes, wrinkling their noses, grimacing, and so forth. All of these actions seem to mean something, but if you can't hear what they are saying, you won't know what these facial actions mean. These are likely to be *conversational* (or *co-speech*) *facial gestures*. As defined and illustrated in the opening paragraphs of this chapter, conversational facial gestures are tightly synchronized with speech in both timing and meaning. If you've previously read a lot about facial expressions, you might assume that these are facial expressions of emotions. However, some caution is in order because, as explained in the next section, conversational facial gestures are not likely to be facial expressions of emotion.

Facial Expressions Versus Facial Gestures

Kraut and Johnston (1979) introduced a distinction between two kinds of facial actions: some that express emotion and others that display information to an interlocutor. *Facial expressions* are part of an emotional, individual process, whereas *conversational facial gestures*¹ are part of a social, interactive process. We propose that this distinction leads to at least four differences: first, conversational facial gestures are precisely timed with speech, so they appear and change within seconds (as illustrated in Figure 1. of this chapter). Ekman (1997) pointed out that an emotion would run its own course and would not go on and off with a single word or phrase.

Second, most theories of facial expressions limit them to a set of stereotypic forms corresponding to a small number of basic emotions (anger, happiness, disgust, etc.). Conversational facial gestures, on the other hand, are part of whatever the person is saying at the moment, so they have innumerable forms, none of which is fixed or stereotypic. For example, in Figure 1. there is no emotion that corresponds to the speaker's demonstration of how Shrek or the cat looked in the movie.

The third distinction is methodological and affects the choice of research procedures: because facial expressions are often seen as stereotypic configurations that correspond to individual emotions, a typical experiment would focus on the participant's ability to recognize these facial expressions in still photographs. In contrast, an experiment on facial gestures in dialogue requires videorecording interlocutors in face-to-face dialogue, then microanalyzing their facial gestures, second by second, with the accompanying words and context.

Fourth, facial gestures include more aspects of the face than do facial expressions of emotion. The analysis of emotional expressions has been limited to configurations of facial muscles, such as Ekman and Friesen's (1978) action units. However, facial gestures often incorporate other aspects of the facial area, especially eye movements and head positions. For example, in Figure 1. frame 3, the tilt of the speaker's head and the upward direction of her gaze portrayed the cat directing his worried expression up at Shrek. Her head and eye positions were part of her facial gesture as a whole.

Conversational Facial Gestures as Demonstrations

These distinctions make it clear that, like hand gestures, conversational facial gestures are demonstrations. Bavelas and Chovil (1997) first proposed that conversational facial gestures meet the criteria of being nonliteral, selective depictions of a referent that they resemble. In Figure 1. the speaker was briefly depicting some selected features of a movie character who was gloating and then another character who was acting worried. Although her features resembled gloating or being worried, she was not literally gloating in frame 2 or worried in frame 3. Instead, she was creating an integrated and tightly synchronized demonstration consisting of her words, facial gestures, and hand gestures, all timed together within fractions of seconds. The result allowed her addressee to perceive directly what these characters looked like at this point in the movie—or at least the speaker's selected version of it. It is highly unlikely that an addressee would mistake these facial gestures for the speaker's own emotional expressions.

Although the vast majority of research on faces has been concerned with emotional expressions, these expressions seem to be infrequent in conversational interactions. Fridlund, Ekman, and Oster (1987, p. 160) reported unpublished data by Ekman and Friesen on nearly 6,000 facial actions from interviews in which psychiatric patients with affective disorders were mostly talking about their feelings. Even in this context, fewer than one-third of these facial actions were classifiable as emotional expressions. In Chovil's (1989, 1991/1992) data from ordinary conversations, the closest category to emotional expressions was "personal reactions" (see Personal Reactions below), which were 24 percent of the facial gestures she found. However, even these were timed with speech and therefore do not fit emotion theory. Given that facial actions during dialogue are primarily conversational facial gestures rather than facial expressions of emotion, it is striking that there are literally thousands of published studies on facial expressions of emotion but only a handful of studies on conversational facial gestures. This topic is wide open for original research.

This chapter relies on the seminal findings of Chovil (1989, 1991/1992, 2005; see also Bavelas & Chovil, 1997), which was the first and still the only extensive description of facial gestures in spontaneous face-to-face dialogues. To discover inductively what contributions facial gestures can make to conversation, Chovil videotaped twelve dyads (four male, four female, four mixed) talking about a variety of assigned topics (e.g., a close-call, an interpersonal conflict). The videos captured close-ups of both interlocutors' faces in split screen. Using the nested

context approach illustrated in Figure 2. Chovil focused on the *function* of each facial configuration at that particular moment in the dialogue and developed a highly reliable systematic analysis of facial gestures (e.g., Chovil, 2005). Smiles were excluded for practical reasons, because they are both frequent and highly varied (Ekman, 1985, p. 150), so that work has yet to be done. Even so, she identified more than a thousand facial actions: 880 were conversational facial gestures, 301 were adaptors, and only 3 were not analyzable. We have maintained Chovil's specific functions (including applying them to some smiles) but have introduced some slightly different groupings and terms and have illustrated them with examples from our own data (primarily from Bavelas et al., 2014).

The overview here follows the same format as the gesture section, first illustrating what facial gestures are able to demonstrate and then summarizing the small experimental literature on their conversational influences.

What Can Facial Gestures Demonstrate?

Faces present different possibilities and constraints than do hand gestures (or, indeed, words) and therefore demonstrate different kinds of referents. Facial gestures cannot portray objects, spatial relationships, motion, and the like, but they can depict what virtually anyone looked like at any particular moment. Conversations are often populated with oneself, someone else, or even imaginary characters, whether in the past, present, or future, so the capacity to depict people's faces is very useful.

Personal Reactions

Both speakers and addressees can demonstrate how they are reacting in the present situation, that is, their evaluation of or feeling about something in their current interaction. For example, in Figure 4. the speaker was relating a scene in *Shrek 2* during which the Donkey accidentally kicked Shrek while trying to help him repel the cat's attack. The speaker said "and then Donkey tries to help Shrek, um, by kicking him in the crotch." In frame 3, as soon as she had finished saying the word "crotch," she smiled and laughed to demonstrate to the addressee that she thought that this was funny. We do not infer from her smile or the smile of the addressee that they were feeling the emotion of happiness at precisely that moment. What we can observe is that their facial reactions demonstrated that the scene was humorous.

Figure 5.(from our archives) shows a different personal reaction by an addressee. The speaker was telling a personal close-call story about walking on a wet log to cross a flooding river. The speaker said "and my foot slips off and I fell into the river." Just as he finished saying the word "river" (in frame 1), the addressee (in frame 2) made a shocked face with pursed lips and said "Oooh." The addressee's face demonstrated specifically and vividly his understanding of the speaker's sudden danger at this point in the story. This facial gesture did not mean that he was actually shocked and concerned during the. 7-second it lasted; instead, it was timed to fit the final, culminating point of the speaker's sentence.

Facial Portrayals

Speakers and addressees can move beyond their present reactions in the dialogue and instead portray themselves or others in the past, the future, or even a hypothetical situation. Frames 2 and 3 of Figure 1. show two such portrayals in which the speaker portrayed the face of Shrek and then the cat. Another example comes from Chovil's (1989, 1991/1992) data. When describing a minor conflict in the past, speakers often portrayed how they had reacted to the other person—and how the other person had reacted to them (e.g., as indignant, angry, or appeasing). One speaker said "But I was like *errrrr!*" Precisely at "*errrrr!*" she tightened her eyes and produced an angry face (1991/1992, p. 182). Chovil (1989) also observed the special case of addressees' facial *motor mimicry*; for example, when the speaker tells about an injury that happened to him or her, the addressee's facial portrayal (e.g., a wince) would be appropriate to the situation the speaker is describing.

Facial Metaphors

Bavelas and Chovil (2000) proposed that facial gestures vary on a continuum of abstraction. In Figures 4 and 5, the facial gestures literally represented the speaker's own face. In Figure 1. however, the facial gestures were somewhat more abstract because they represented someone else's face; that is, the addressee had to interpret the speaker's face *as if* it were Shrek's or the cat's face looking that way. At an even more abstract level, "*facial*

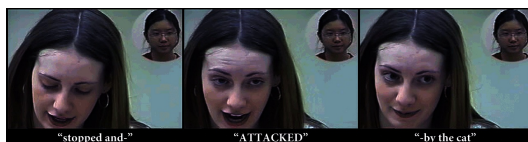
metaphors convert the literal meaning of a facial reaction to an abstract one by using it in a metaphor” (Bavelas & Chovil, 2000, pp. 174–175). For example, when an addressee begins to suspect that the speaker is exaggerating or teasing, he may tilt his head and squint, as if—metaphorically—looking more closely at the story. Or an addressee who is hearing a friend describe a social blunder may show a pained expression even though the friend’s “pain” was purely social, not physical. Ultimately, facial gestures may become even more abstract, as in many of those described in the next section.

Facial Gestures as Collateral Communication

Many facial gestures function as *collateral communication* (Clark, 1996, pp. 241–243, 255–257) because they demonstrate information about the status of the ongoing talk rather than the content of the talk. For example, speakers use various facial gestures to stress specific words or phrases, to explain pauses for a word search, to indicate a question, and to demarcate the beginning, continuation, or end of a story. Addressees use backchannel facial gestures to demonstrate that they are attending and following. Chovil (1989, 1991/1992) found that almost half of the speakers’ facial gestures in her data did not convey topical information such as a personal reaction or a portrayal. Instead, they marked some aspect of the discourse itself. Instead of using Chovil’s terms, we are grouping these facial gestures under collateral communication in order to emphasize their functional similarity to collateral communication by words or gestures.

Emphasizers, Underliners, and Question Markers

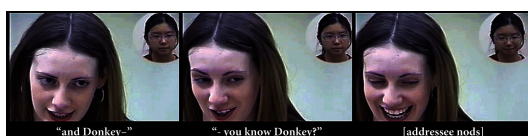
In Chovil’s data, the most common collateral communication with the face was the *emphasizer* (e.g., Ekman, 1979), which occurred synchronously with a stressed (e.g., prosodically marked) word. These were typically brow movements; the facial muscles that move the brows are rapid enough to go up and down with a single word or even syllable. Figure 6. shows an *emphasizer* in frame 2. The speaker was talking about a scene from *Shrek 2* in which Shrek and Donkey were suddenly ambushed by Puss in Boots. The speaker said “and they are stopped and **ATTACKED** by the cat”; precisely as she vocally stressed the word “**ATTACKED**,” she raised and lowered her eyebrows. Together, the forceful prosody and the raised eyebrows emphasized the key word in her sentence. Chovil also found what Ekman (1979) had called *underliners*, which mark a whole phrase rather than a single word. Bavelas and Chovil (2000, p. 175) later speculated that the raised eyebrows characteristic of *emphasizers* and *underliners* are a metaphor based on the startle reaction, although the speaker in Figure 6. was obviously not demonstrating that she was literally startled by this word.



[Click to view larger](#)

Figure 6 . In frame 2, the speaker’s raised eyebrows emphasized the main word in her sentence (“ATTACKED”). The three frame shots cover a 3.4-second period. (The face of the addressee, who was sitting across a table from the speaker, is inset at the upper right.)

The next most frequent collateral communication in Chovil’s (1989, 1991/1992) data had a grammatical function, the *question marker*. Speakers frequently raised or lowered their eyebrows when asking a question, including questions marked only with rising intonation. In Figure 7. the speaker started retelling the movie by introducing the characters. When she came to the Donkey character, she said, “You know, Donkey?” and raised her brows exactly with “know Don-” to mark her brief phrase as a question. The addressee nodded and smiled slightly, then the speaker continued. Chovil also found other, less frequent facial gestures that served as collateral communication, including ones that marked the beginning, continuation, or ending of a topic. For example, immediately after frame 3 in Figure 1. the speaker smiled to close that part of her narrative.



[Click to view larger](#)

Figure 7 . The speaker's raised eyebrows marked her words in frame 2 as a question. The three frame shots cover a 2.3-second period. (The face of the addressee, who was sitting across a table from the speaker, is inset at the upper right.)

Thinking Faces and Facial Shrugs

The *thinking* (or *remembering*) face is a collateral communication indicating that the speaker is thinking about what to say, trying to remember something, or searching for a word. Typically, the speaker breaks eye contact briefly by looking away (or closing both eyes) and might also suddenly make either a "blank" or an "effortful" expression. In Figure 8, the speaker had announced that she was about to describe the scene in which Shrek and Donkey met Puss-in-Boots; instead, she paused, shook her head slightly, and said (as if asking herself) "What happens?" while tilting her head and looking up. It took only 1.82 seconds to move into and hold this gesture, which ended as soon as she began to describe the scene. Whether or not the gesture actually helped the speaker to think or remember, it did serve the important social function of explaining her temporary speaking pause to the addressee. (Similarly, in Figure 1, as the speaker changed from being Shrek to being the cat during the fraction of a second between frames 2 and 3, she paused at "like—" and made a brief thinking face, looking away with a blank face.)



[Click to view larger](#)

Figure 8 . While recalling a scene from *Shrek 2*, the speaker made an extended (1.7-second) thinking face in frames 2 and 3. (The face of the addressee, who was sitting across a table from the speaker, is inset at the upper right.)

Another collateral communication is the *facial shrug*, which (like a shoulder shrug) conveys either not knowing or not caring about something at that moment in the dialogue (e.g., having said enough on the topic, having said it well enough, conceding or resigning a point). Both Ekman (1985) and Chovil (1989, 1991/1992) observed facial shrugs, which typically involve a quick eyebrow flash and the retraction of a corner of the mouth. Figure 9, shows the speaker shrugging with her face by pulling the corner of her mouth inward. This facial shrug occurred as she came to the end of a scene in the *Shrek* movie and paused between two parts of her utterance. She said "and then," made the facial shrug, and concluded with "I guess... they take him." The shrug functioned to demonstrate that the speaker was not certain whether her conclusion was actually what had happened, but it was good enough. Thinking faces and facial shrugs function as collateral communication (rather than personal reactions) because they comment on what the speaker is saying rather than providing topical content.



[Click to view larger](#)

Figure 9 . The speaker was completing her description of a scene from *Shrek 2* when, instead of continuing on from frame 1, she made a facial shrug (in frame 2). Then she said "I guess" (in frame 3) and went on. The three frame shots cover a 1.3-second period. (The face of the addressee, who was sitting across a table from the speaker, is inset at the upper right.)

Facial Backchannels

Addressees as well as speakers make facial gestures that function as collateral communication. The majority of addressees' facial actions in Chovil's (1989, 1991/1992) data were *backchannel responses*. Chovil found that facial backchannels typically consisted of brow raises, mouth corners turned down, or lips pressed together. Recall that Chovil did not study smiles, but Brunner's (1979) intensive analysis of backchannels revealed that

[Addressees'] smile beginnings showed a strong tendency to occur at the same kinds of location as "back channel" responses (such as "yeah," "uh-huh," and head nods). This finding indicates that the smile can function as a type of back channel. It is argued that smiles, like other forms of back channel, make communication more efficient by providing the speaker with feedback on a number of levels

simultaneously. (p. 728)

We would add that addressees' facial gestures are efficient for another reason as well: they provide rapid feedback to the speaker without interrupting or taking up a turn.

Facial Gestures and Words

As many of these examples show, conversational facial gestures are meaningfully related to the words they occur with, but they are not always redundant with them. Instead, faces can contribute unique information to the integrated message. Within one of her major groups, in which the speaker's face conveyed information about the topic, Chovil (1989, 1991/1992; see also Bavelas & Chovil, 1997) analyzed how often the face provided nonredundant information (i.e., not at all in the accompanying words) versus how many were at least somewhat redundant with words. A surprising 40 percent of this group were completely nonredundant; that is, they conveyed information that complemented but was not conveyed in the accompanying words.

Conversational Factors that Affect Facial Gestures

As shown in the preceding paragraphs, the interlocutors' faces are quite active during face-to-face dialogue, conveying both topical content and collateral communication. In view of this ubiquity, the lack of experimental research on facial gestures in dialogue can hardly be overemphasized (see review in Bavelas & Chovil, 2006). To our knowledge, the published reports of dialogic data consist of the three systematic analyses cited earlier (Brunner, 1979; Chovil, 1989, 1992/1992; Ekman, 1979) and two experiments (Chovil, 1989, 1991; Bavelas et al. 2014).

Effects of Visibility on Addressees' Motor Mimicry

The first controlled experiment on facial gestures in dialogue was Chovil's (1989, 1991) study of addressees' motor mimicry. As described earlier, motor mimicry occurs when someone's reaction is not appropriate to his or her own situation; instead, it is appropriate to the situation of the other person. Bavelas, Black, Lemery, and Mullett (1986) had shown that facial motor mimicry depended on eye contact with the injured person, but their experimental setting was not a dialogue. Chovil created four experimental conditions: the addressee was hearing a close-call story either in a face-to-face dialogue, in a dialogue on the telephone, in a dialogue through a partition, or when alone (listening to an answering machine). The speakers told their own stories in the first three conditions; the story on the answering machine was chosen from earlier data as a particularly alarming one. The rate of mimicry was significantly higher in the face-to-face dialogue than in the other three conditions, where an addressee would not see it. The rate was also significantly lower when the addressee was alone, compared to the three dialogue conditions.

Altogether, the series of experiments on motor mimicry including Bavelas et al. (1986), Chovil (1989, 1991), Bavelas, Black, Chovil, Lemery, & Mullett (1988), and Bavelas, Coates, and Johnson (2000) led to the conclusion that motor mimicry is not a reflexive or emotional reaction; it functions communicatively in dialogue, specifically, as a display of understanding. (See Bavelas, 2007, for a critical review of subsequent interpretations of this work.)

Effects of Visibility and Dialogue on Speakers' Facial Portrayals

To our knowledge, Bavelas et al. (2014) is the first formal experiment on speakers' facial gestures, specifically, speakers' facial portrayals of characters in the movie, *Shrek 2* (as illustrated throughout this chapter). The three experimental conditions were face-to-face dialogue, telephone dialogue, and monologue into a microphone, making it possible to assess the effects of both visibility and dialogue separately. Visibility increased the rate of facial portrayals (i.e., it was significantly higher in the face-to-face condition than in the telephone and tape recorder conditions combined). Dialogue also increased the rate of facial portrayals (i.e., it was higher in the combined face-to-face and telephone dialogues than in monologue). Note that Chovil's (1989, 1991) experiment on addressees' motor mimicry also showed both visibility and dialogue effects.

Summary of Facial Gestures

In contrast to hand gestures, experimental research on facial gestures is rare, probably overwhelmed by research on emotional expressions. The major descriptive studies were published decades ago, but, as shown in this chapter, they fit new data just as well. Given that faces in dialogue are lively, eloquent, efficient, infinitely varied, and ubiquitous, it is well past time to start investigating them in more detail.

Conclusion

Face-to-face dialogue is the basic form of language use (e.g., Clark, 1996; Goodwin, 1981; Levinson, 1983). The processes by which interlocutors in dialogue coordinate their actions to create meaning have begun to interest scholars from many disciplines, including linguistics, anthropology, and psychology. Research on how interlocutors use visible resources, such as hand and facial gestures, is intrinsic to understanding language use in face-to-face dialogue. Regardless of the discipline, investigations of language use in dialogue must meet, at minimum, two methodological imperatives. First, the interactions themselves must maintain the essential features of dialogue, including interlocutors' ability to be spontaneous and to extemporize. The studies reviewed here demonstrate that this imperative can easily be met in an experimental setting when necessary control is achieved through careful research design rather than through reduction to a quasi-dialogue or, as we put it elsewhere, committing "dialogicide" (Gerwing & Bavelas, 2013).

The second methodological imperative is that the original data must not only be recorded full-face on video, but the analysis must also be conducted directly from video. Hand and facial gestures coordinate with and complement speech, and they often provide essential details that are not in the words. Digitized video and suitable software (e.g., ELAN, <http://tla.mpi.nl/tools/tla-tools/elan/>; Wittenburg et al., 2006) permit the close observation of dialogue as a lively synchrony of words, gestures, and faces. With frame-by-frame, repeated viewing, analysts can observe interlocutors' behaviors carefully and systematically. Rigorous procedures for analysis and detailed operational definitions, plus regular tests for interanalyst reliability are essential for ensuring interanalyst agreement.

The overview presented in this chapter leads to several broad conclusions: hand gestures are not simply redundant movements nor signs of dysfluency. They contribute a wide variety of information ranging from the concrete and physical to the metaphorical to meta-communicative collateral communication. The evidence suggests that this field has matured from a debate over *whether* gestures are communicative into solid research on *how* they communicate. Nor are facial gestures a limited, static set of emotional configurations. They are a constantly changing accompaniment to words, adding new and nuanced details, again ranging from the concrete to the abstract to collateral communication.

An abundant and increasing number of quantitative experiments on hand gestures are exploring subtle relationships and unexpected topics. In short, this field is healthy and growing rapidly. However, the study of facial gestures needs to emerge from the dominance of emotion theory in order to produce a body of research of its own. Research on the contributions that interlocutors make with their faces cannot advance until face-to-face dialogues are routinely recorded and microanalyzed. For young researchers, this field is wide open for systematic description and especially for experimental research. Every area touched on in this chapter could expand many times over. Recent research on hand gesture provides models and methods that might apply to studies of facial gesture as well.

One reservation is that experimental work has too often treated gestures, faces, and words as separable phenomena. In face-to-face dialogue, these (and other) communicative resources work together, and it is essential to explore the details of their division of labor. The semantic features approach (e.g., Beattie & Shovelton, 1999; Gerwing & Allison, 2009) is one possible method for quantifying the contributions of each modality by providing a systematic approach for exploring how variations in social context (e.g., mutual visibility, common ground) might influence the distribution of information among hand gestures, facial gestures, and words. Another, mostly unexplored area is how all of these come together to present a unitary multimodal configuration, such as the one emphasized in Figure 1. Moreover, language use in face-to-face dialogue is more than words, hand gestures, and facial gestures. Prosody and gaze suggest themselves as new areas for research—provided that these studies always take place within true face-to-face dialogues. Scraps of scripted talk and recordings of gaze at inanimate targets fit the outmoded principle of reductionism that continues to threaten the study of conversations in face-to-face dialogue.

Other than a few exemplars reviewed here (Bavelas et al., 2011; Clark & Krych, 2004; Holler & Wilkin, 2011; Furuyama, 2000), researchers have tended to ignore the *sequential* relationships between interlocutors' behaviors as their dialogue unfolds. This hesitation has held the field back from understanding how visible behaviors are sequentially integrated into the dialogue and how they influence subsequent utterances. One well-established model for studying these relationships within an experimental setting is a referential communication task (e.g., Holler & Wilkin, 2011), in which pairs of participants must complete a series of steps together but one person has all of the necessary information and the other person does not. Although used extensively to study the sequential processes of verbal communication, this method is underused for exploring how interlocutors use their hand and facial gestures sequentially.

On a theoretical level, our research group is pursuing a theory of *demonstrations* (Bavelas et al., 2008, 2014), which include not only hand and facial gestures but also direct quotations and metaphors. We propose that demonstrations require an addressee and are therefore uniquely linked to interacting in dialogue rather than monologue. Underpinning this theory is the proposal that, rather than dividing the elements of dialogue along physical lines such as "verbal" and "nonverbal," it might be more useful to make functional distinctions, such as between *description* and *demonstration*. Exactly how demonstrations work alongside descriptive speech is another wide-open question.

Throughout this chapter, we have emphasized experimental research, not because it is superior to other methods, but because it is a useful and informative method that is most often used in language and social psychology. It would be remiss not to point out that most original insights come from everyday observations and qualitative analyses (e.g., Bavelas, 1987). These provide a springboard for research questions and hypotheses, which lead to designing experiments that might answer or test them. There is a world of unexplored dialogic phenomena out there, one that is not found in libraries or statistics.

References

- Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169–188.
- Bavelas, J. B. (1987). Permitting creativity in science. In D. N. Jackson & J. P. Rushton (Eds.), *Scientific excellence: Origins and assessment* (pp. 307–327). Beverly Hills: Sage.
- Bavelas, J. B. (2005). The two solitudes: Reconciling Social Psychology and Language and Social Interaction. In K. Fitch & R. Sanders (Eds.), *Handbook of language and social interaction* (pp. 179–200). Mahwah, NJ: Erlbaum.
- Bavelas, J. B. (2007). Face-to-face dialogue as a micro-social context. The example of motor mimicry. In S. Duncan, E. Levy, & J. Cassell (Eds.), *Gesture and the dynamic dimension of language* (pp. 127–146). 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, UK: Cambridge University Press.
- Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning. An integrated message model of language use in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163–194.
- Bavelas, J. B., & Chovil, N. (2006). Hand gestures and facial displays as part of language use in face-to-face dialogue. In V. Manusov & M. Patterson (Eds.), *Handbook of nonverbal communication* (pp. 97–115). Thousand Oaks, CA: Sage.
- Bavelas, J. B., Chovil, N., Coates, L., and Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social*

Psychology Bulletin, 21, 394–405.

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

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

Bavelas, J. B., Gerwing, G., Allison, M., & Sutton, C. (2011). Dyadic evidence for grounding with abstract deictic gestures. In G. Stam, M. Ishino, & R. Ashley (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 49–60). Amsterdam: Benjamins (Gesture Studies series).

Bavelas, J. B., Gerwing, J., & Healing, S. (2014). The effect of dialogue on demonstrations: Direct quotations, facial portrayals, hand gestures, and figurative references. *Discourse Processes*, in press.

Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, 495–520.

Bavelas, J. B. & Healing, S. (2013). Reconciling the effects of mutual visibility on gesturing. *Gesture* 13(1), 63–92.

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, 123, 1–30.

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

Buchler, J. (Ed.). (1940). *The philosophy of Peirce: Selected writings*. London: Routledge & Kegan Paul.

Chovil, N. (1989). *Communicative functions of facial displays in conversation*. Unpublished Ph.D. 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. (2005). Measuring conversational facial displays. In V. Manusov (Ed.), *The sourcebook of nonverbal measures: Going beyond words*. Hillsdale, NJ: Lawrence Erlbaum.

Clark, H. H. (1985). Language and language users. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., vol. 2, pp. 179–232). New York: Random House.

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

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(1), 62–81.

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 Communication*, 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(2), 276–279.

Danziger, K. (1990). *Constructing the subject. Historical origins of psychological research*. Cambridge (UK): Cambridge University Press.

de Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science* 4(2), 232–248.

Ekman, P. (1977). Biological and cultural contributions to body and facial movement. In J. Blacking (Ed.), *Association of Social Anthropologists monograph 15, The anthropology of the body*. London: Academic Press.

Hand and Facial Gestures in Conversational Interaction

- Ekman, P. (1979). About brows: Emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology* (pp. 169–249). Cambridge, UK: Cambridge University Press.
- Ekman, P. (1997). Should we call it expression or communication? *European Journal of Social Sciences*, *10*, 333–359.
- Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P. E. (1985). *Telling lies*. New York: Berkley Books.
- Emmorey, K., & Casey, S. (2001). Gesture, thought and spatial language? *Gesture*, *1*, 35–50.
- Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, *31*(4), 307–319.
- Fridlund, A. J., Ekman, P., & Oster, H. (1987). Facial expressions of emotion: Review of literature, 1970–1983. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (2nd ed., pp. 143–224). Hillsdale, NJ: Erlbaum.
- 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., & Allison, M. (2009). The relationship between verbal and gestural contributions in conversation: A comparison of three methods. *Gesture*, *9*(3), 313–336.
- Gerwing, J., & Allison, M. (2011). The flexible semantic integration of gestures and words. Comparing face-to-face and telephone dialogues. *Gesture*, *11*(3), 308–329.
- Gerwing, J., & Bavelas, J. B. (2004). Linguistic influences on gesture's form. *Gesture*, *4*(2), 157–195.
- Gerwing, J., & Bavelas, J. (2013). The social interactive nature of gestures: Theory, assumptions, methods, and findings. In C. Müller, A. Cienki, E. Fricke, S.H. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body-language-communication, Volume 1, Contemporary approaches* (Ch. 51, pp. 821-836). Berlin: Mouton De Gruyter.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Gullberg, M. (2010). Methodological reflections on gesture analysis in SLA and bilingualism research. *Second Language Research*, *26*(1), 75-102.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512–521.
- Healing, S., & Gerwing, J. (2012 July). *Spatial cohesiveness in gesture sequences. Effects of mutual visibility*. Paper presented at the meeting of the International Society for Gesture Studies, Lund (Sweden).
- Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, *26*(1), 4–27.
- Holler, J., Tutton, M., & Wilkin, K. (2011). Co-speech gestures in the process of meaning coordination. In *Proceedings of the 2nd GESPIN—Gesture in Speech and Interaction Conference*, Bielefeld, 5–7, September 2011.
- Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, *35*, 133–153.
- 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.
- Kendon, A. (2004). *Gesture. Visible action as utterance*. Cambridge, UK: Cambridge University Press.

- Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, 31(6), 533–552.
- 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.
- Kuhlen, A. K., & Brennan, S. E. (2012). Language in dialogue: When confederates might be hazardous to your data. Accepted for publication at *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-012-0341-8. The final publication is available at www.springerlink.com.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Luria, A. R. (2004). Reductionism in psychology. In R. L. Gregory (Ed.), *The Oxford companion to the mind* (2nd ed., pp. 675–676). Oxford, England: Oxford University Press.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92, 350–371.
- Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009a). Communicative gestures and memory load. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society, Amsterdam, Netherlands*, 1569–1574.
- Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009b). The communicative import of gestures. Evidence from a comparative analysis of human-human and human-machine interactions. *Gesture*, 9(1), 97–126.
- 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, UK: 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(4), 688–704.
- Parrill, F. (2008). Subjects in the hands of the speakers: An experimental study of syntactic subject and speech-gesture integration. *Cognitive Linguistics*, 19, 283–299.
- Pine, K. J., Burney, D. J., & Fletcher, B. (2010). The semantic specificity hypothesis: When gestures do not depend upon the presence of a listener. *Journal of Nonverbal Behavior*, 34, 169–178.
- Reber, A. S., Allen, R., & Reber, E. S. (2009). *Penguin dictionary of psychology* (4th ed.). London, UK: Penguin.
- Rimé, B. (1982). The elimination of visible behavior from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology*, 12, 113–129.
- Rowbotham, S., Holler, J., Lloyd, D., & Wearden, A. (2011). How do we communicate about pain? A systematic analysis of the semantic contribution of co-speech gestures in pain-focused conversations. *Journal of Nonverbal Behavior*, 36(1), 1–21.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211–232.
- Sears, R. R. (1951). A theoretical framework for personality and social psychology. *American Psychologist*, 6, 476–483.
- Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.
- Wittenburg, P., Brugman, H., Russel, A., Klassman, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In: *Proceedings of LREC 2006, Fifth International conference on language resources and evaluation*, Genoa.

Notes:

(¹) Kraut and Johnston used the term *facial display*, borrowed from the study of animal behavior, where the overt

behavior acts as a social display to others. Chovil also adopted this term in her extensive work. We propose a change to *facial gesture*, which emphasizes the many similarities to hand gestures.

Janet Bavelas

Janet Bavelas, Department of Psychology, University of Victoria

Jennifer Gerwing

Jennifer Gerwing, Department of Psychology, University of Victoria

Sara Healing

Sara Healing, Department of Psychology, University of Victoria

