

# John Benjamins Publishing Company



This is a contribution from *Gesture 13:1*

© 2013. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

Tables of Contents, abstracts and guidelines are available at [www.benjamins.com](http://www.benjamins.com)

# Reconciling the effects of mutual visibility on gesturing

## A review

Janet Bavelas and Sara Healing  
University of Victoria

Fourteen visibility experiments, which compared the overall rate of gesturing when participants could or could not see each other, have produced perfectly contradictory results: Seven found a significantly higher overall gesture rate in the visible condition, and seven found no significant difference. Experiments that used *quasi-dialogues* in which the addressees' responses were experimentally constrained (e.g., using a confederate) found a significant difference; the experiments that used *free dialogues* did not. This review examined three possible explanations and found that (1) the use of quasi-dialogues did not ensure better experimental control, (2) the constrained addressees may have introduced a confound that could account for the significant difference, and (3) although mutual visibility did not affect the overall gesture rate in free dialogues, it significantly increased more specific features of gestures that are useful to addressees. These findings raise several issues about the utility of conventional visibility designs for understanding conversational gestures.

**Keywords:** conversational hand gestures, visibility effects, confederates, gesture measures, experimental design and interpretation, visible communicative acts, unit of analysis

## Introduction

Experiments that use a visibility design compare gesture rates in face-to-face dialogues versus dialogues that lack mutual visibility. In the past 40 years, there have been 14 controlled experiments using this design: Cohen & Harrison (1973); Cohen (1977); Rimé (1982); Bavelas, Chovil, Lawrie, & Wade (Exp. 2, 1992); Krauss, Dushay, Chen, & Rauscher (1995, Exps. 1 & 2); Alibali, Health, & Myers (2001); Emmorey & Casey (2001); Bavelas, Gerwing, Sutton, & Prevost (2008);

Mol, Kraemer, Maes, & Swerts (2009a, 2009b); Pine, Gurney, & Fletcher (2010); Holler, Tutton, & Wilkin (2011); de Ruiter, Bangerter, & Dings (2012); Bavelas, Gerwing, & Healing (2013b).

The conventional hypothesis in a visibility design is that if the speakers' overall rate of gesturing decreases significantly when their addressee cannot see them, then gesturing must be communicative. The counter-hypothesis is that, if mutual visibility has *no* significant effect on gesturing, then gestures are not communicative. In the latter case, it was common in the above studies to conclude that the function of unseen gestures — and perhaps all gestures — must therefore be individual rather than social or communicative. Others have gone further and interpreted null results more specifically, as evidence that gestures aid cognition (e.g., word-finding; Krauss, Chen, & Gottesman, 2000). This article focuses on unrecognized problems in the design and interpretation of visibility experiments and, in the process, raises questions about whether simple visibility designs can support either hypothesis.

### *Contradictory results*

One would expect that so many replications of the same design would produce a clear set of findings, but the results split exactly in half. Table 1 (after the References) summarizes seven experiments that found a significant difference in the amount of gesturing between the visible and not-visible conditions. Table 2 summarizes another seven that found no significant difference. A comparison of the various features of these experiments eliminates several initial explanations for this difference. For example, the pattern of discrepant results does not match how each experiment manipulated visibility (whether by screen, intercom, or telephone). Nor is there any obvious difference in the gender of participants, the sample sizes, or the diversity of the tasks used. The gesture measures do vary considerably in clarity and specificity, but that is true for the experiments that found a difference (Table 1) as well as for those that did not (Table 2). Note that both tables include only the *overall gesture rate*, which was the broadest category of gestures in each experiment (either all gestures or all representational, iconic, or illustrator gestures). These omnibus gesture rates were present in every study, and they are the usual basis for concluding whether or not mutual visibility affects gesturing.<sup>1</sup>

One pattern that does seem promising is chronological. In recent years, as the number of visibility studies has increased, there is a trend toward finding no significant effect of mutual visibility. Of the seven studies since 2008, five found no significant difference between their visible and not-visible experimental conditions. A closer examination of this trend reveals a pattern that fits even better: The recent studies are more likely to have used *free dialogues*, in which both the

speaker and addressee were real participants, interacting spontaneously within their assigned task. Most earlier studies used what we will call *quasi-dialogues*, in which the addressee was either a confederate, the experimenter, or another participant — all of whom were interacting under experimental guidelines that constrained their interaction with the speaker. In fact, this single methodological difference corresponds exactly to the difference in findings. As shown in row 2 of each table, the seven experiments with a significant difference between their visible and not-visible conditions all used quasi-dialogues, and the seven experiments with no significant difference between conditions all used free dialogues.

These two sets of well-replicated results are clearly incompatible, presenting the problem of how to decide which set of results to accept and which to reject. This article will raise and critically evaluate a sequence of three relevant questions, with the goal of arriving at an empirically supported explanation that will account for all of the data.

*Question 1:* Were the experiments using quasi-dialogues more tightly controlled than those with free dialogues and therefore better able to detect real differences between conditions — differences that the presumably less well-controlled free dialogues missed?

*Question 2:* If not, then did the experiments using quasi-dialogues inadvertently create a confound that could account for the difference between the experimental conditions at least as well as visibility does, which would make the non-significant difference in the free dialogues more credible?

*Question 3:* If so, then is the overall rate of gesturing the right measure for testing the effects of mutual visibility on gesturing, or are more specific measures required?

The close examination of these questions is methodological and often technical, focusing on the logic of experimental design and interpretation. However, we propose that these technical matters arise directly from — and can inform — a more fundamental theoretical issue, the choice of unit of analysis. Therefore, before examining the three questions, we will outline a theoretical choice that forms the context for everything that follows.

### *Choosing a unit of analysis*

It is very likely that the use of free versus quasi-dialogues follows from broader theoretical assumptions about the minimum unit of analysis for studying dialogue. As Kuhlen and Brennan (2013) pointed out, researchers' use of confederates "reflects their theories, implicit or explicit, on the nature of dialogue" (p. 58). We propose that constraining an interactional partner in a quasi-dialogue (whether

this person is a confederate, the experimenter, or another participant) is the methodological expression of a paradigm in which the basic unit of analysis in a dialogue is the *individual*. If so, then the other person in the dialogue is an external or environmental factor, and any spontaneous influence of this other person on the target individual is a source of noise or error variance. Danziger (1990) pointed out that this “Robinson Crusoe myth” is prevalent in individualistic societies and found a firm place early in the history of psychology. The focus on the individual as the unit of analysis was soon revealed by

the way in which the regularities emerging out of investigative situations were re-attributed. They were attributed, of course, to the individuals who participated in the investigation as subjects. However, the regularities in question were not taken as conveying information about an individual-in-a-situation but about an individual in isolation whose characteristics existed independently of any social involvement. (p. 186)

In this view, using a quasi-addressee is equivalent to holding the situation constant.

The alternative paradigm treats the freely interacting *dyad* as the unit of analysis (e.g., Bavelas, 2005, 2007; Clark, 1985, 1996, Ch. 1; Kuhlén & Brennan, 2013; Sears, 1951; Thibaut & Kelley, 1959, Ch. 1). For example, Clark’s (1996) definition of dialogue as the basic setting of language use includes three essential characteristics:

- Extemporaneity: The participants formulate and execute their actions extemporaneously, in real time.
- Self-determination: The participants determine for themselves what actions to take and when.
- Self-expression: The participants take actions as themselves [e.g., one of them is not a confederate acting as if he or she is a real participant]. (p. 10)

In other words, providing someone for the speaker to talk to can only be considered a free dialogue if both participants are free to interact spontaneously, without prior constraints about how they talk with each other.

Implicit in the dyadic paradigm is the assumption that spontaneous dialogue is neither noise nor error variance. Spontaneous dialogue is systematic, orderly, and shaped to the context in which it occurs. Indeed, the goal of research on free dialogues is to reveal and understand this orderliness rather than to learn about how single individuals behave in a dialogue. Research on dyads as the unit of analysis can be qualitative or quantitative, experimental or non-experimental. When dyadic researchers choose to do experiments, they follow the same principles of classical experimental design and control as the researchers who focus on the

individual. However, the dyadic researchers interpret these principles differently and apply control at a different level. Before contrasting their choices, it is useful to outline the principles that both kinds of experimenters do agree on.

### *The rationale for experimental control*

The principle of experimental control is traceable to one of John Stuart Mill's (1882/2008) "Methods of Experimental Inquiry ... comparing instances in which the phenomenon does occur, with instances in other respects similar in which it does not" (pp. 478–479). In other words, keep everything the same except the one thing you want to contrast. Experimenters want to know the effect of their independent variable on behavior, that is, whether one condition (e.g., mutual visibility) leads to more gesturing than another condition (e.g., lack of mutual visibility) does. To establish such a difference requires tight control of extraneous or unwanted variables, for two reasons:

First, extraneous variance tends to produce non-significant results (Type II error). The random effects of other, unwanted variables are "noise" that gets in the way of detecting statistical differences between conditions. For example, it is obviously better to have all participants, in both conditions, do the same task rather than to give each participant a different task. Different tasks could lead to widely different amounts of gesturing, obscuring any effects of visibility. Using one standardized task would minimize these unnecessary differences and make the experiment more likely to detect any differences caused by the independent variable.

Second, good experimental control makes the interpretation of significant results clearer (increasing internal validity). For example, if the visible condition did one task and the not-visible condition did another task, this would create an unwanted *confound*. The worst kind of confound is one that can explain the results just as well as the experimenter's hypothesis; it is an equally plausible alternative explanation. If the task in one condition was one that elicited more gestures than the task in the other condition, then any difference between the conditions could as easily be due to the task as due to visibility. The risks of both Type II error and of confounds will be relevant in this article.

### *Methodological consequences of the unit of analysis*

When researchers assume that the unit of analysis is the individual, then a spontaneous dialogue is, by definition, not under the control of the experimenter. The recommended solution has been to control or hold constant the actions of the person who is not the target of the research by constraining his or her behavior (e.g., by using a confederate; Aronson & Carlsmith, 1968). Those of us trained in

experimental social psychology, as well as in some areas of communication research and psycholinguistics, would have been taught this principle as an absolute. It was unquestioned that standardizing and holding constant the other person's behavior always created a better experiment by maximizing internal validity. We also learned that this was an acceptable trade-off for ecological validity (the degree to which our findings would apply outside the lab). The experiments in Table 1 followed this principle, using referential communication tasks that focused on a one-way flow of information from the speaker. The quasi-addressee's understanding of this information was not an issue. Instead, the quasi-addressee's behavior was standardized, essentially serving as a screen on which the speaker's description and gestures were projected (see Kuhlen & Brennan, 2013, p. 58).

For researchers who assume that the unit of analysis is the dyad, experimental control of one of the individuals would destroy the very object of the study. (In fact, most dyadic researchers would question whether it is even possible to control one of the participants in the rapid complexity of a face-to-face dialogue.) However, experiments with free dialogues are not uncontrolled; they simply apply control at the level of the dyad. Control at this level is most likely to involve a careful choice of the task and of the reciprocal roles for the participants, as was the case for the visibility studies in the free dialogues in Table 2:

- Rimé (1982) used a task that ensured both participants would contribute freely, instructing them “to explain to each other their opinions on movies and to express what they liked to find in the cinema” (p. 117).
- Both Holler et al. (2011) and de Ruyter et al. (2012) used the Tangram task, in which the Director and Matcher collaborated freely in order to ensure that the addressee (the Matcher) understood the Director's information correctly.
- The task in Pine et al. (2010) was also a referential communication task in which “the participants took turns ... describing the item on [a card] to their partner without naming the object .... The describer could not move on to the next trial until the correct answer had been given” (pp. 173–174).
- Similarly, in Bavelas et al. (2008, 2013b), the speaker and addressee knew that the addressee would later have to choose the picture that the speaker had described from a set of four similar pictures.
- Bavelas et al. (1992) asked speakers to tell a personal close-call story to their addressee, who did not know the story. Although this task would seem to be a one-way transmission, Bavelas, Coates, and Johnson (2000) showed that addressees who were “just listening” to such stories made a wide variety of spontaneous responses and did so at a mean rate of 3.5 responses per second (p. 946). When addressees in a contrasting experimental condition had to do a

task that distracted them, they responded significantly less, and their narrators told their stories significantly more poorly.

All of these experiments introduced control on the dyad as a pair, not on one individual within the dialogue; a freely interacting addressee was an essential part of the dialogue. It may be that control at the level of the dyad has more claim to ecological validity for phenomena such as conversational gestures, because the gestures are occurring in a truly conversational context. However, the questions below focus on internal validity.

### **Question 1: Do quasi-dialogues ensure better experimental control?**

Researchers operating within the individual paradigm would argue that control at the level of the individual addressee's responses (rather than at the level of the dyad) is more precise and creates a more advantageous experimental context in which to focus on the speaker's gestures. By this reasoning, the seven experiments in Table 1, which used quasi-dialogues, would have had good experimental control, whereas the experiments in Table 2 with free dialogues would be more poorly controlled. As outlined above, achieving better experimental control by standardizing the addressees' responses would mean less random variation from these addressees, which would in turn reduce Type II error, making the experiment more sensitive to detecting real differences statistically. The corollary of this view would be that, because the addressees in the free dialogues could respond as they wished, these seven experiments would have had more error variance, which would have made them less able, statistically, to detect differences. A first glance at the results in Table 1 versus Table 2 would seem to support this reasoning: It appears that the quasi-dialogues controlled unwanted variance from addressees and were therefore able to detect a difference between conditions — a difference that the free dialogues could not detect because of the spontaneous, uncontrolled actions of their addressees.

However, the principle of experimental control includes the necessity to show both how the control was implemented and that it was effective. In this spirit, we will ask two more specific questions about the control features of the experiments with quasi-dialogues: (a) Were the constraints on the quasi-addressees clearly and precisely defined so that the reader knows exactly which behaviors were controlled and how? (b) Did the experiment include a procedure for confirming that the intended control was carried out in every case, in both conditions?

*Definition of constraints*

Whether the quasi-addressee was the experimenter, a confederate acting as another participant, or a real participant, the experimental constraints on their behavior during the dialogue were the key control feature. These constraints should be precise and similar (across speakers and across experimental conditions) in what they permitted the addressee to do and not do. However, the descriptions of the constraints in these studies showed how difficult it was to define precisely what was being controlled. Table 3 includes the verbatim descriptions of the constraints on the quasi-addressee; we will choose illustrative examples for the text. As can be seen in Table 3, the constraints on the quasi-addressees ranged from remaining “as non-reactive as possible” to simply “not to ask any questions”. Within that range, some of the quasi-addressees could ask questions but not interrupt, and some could neither interrupt nor ask questions. It seems to have been common practice not to provide detailed operational definitions for the permitted and precluded addressee behaviors. For example, an addressee’s *question* that requests new information may use the same rising intonation as a *formulation* that is just checking on understanding; were both of these prohibited? *Interruptions* are not easily distinguished from rapid turn exchanges and overlapping speech. *Non-reactive* is simply the absence of unspecified actions. Perhaps most important, “nods” and “looking at the speaker” were the only references to the wide variety of non-interruptive visible communicative acts available to addressees who were face to face with their speakers. For example, none of these experiments mentioned whether the quasi-addressee was allowed to gesture or not.

The descriptions of constraints in Table 3 imply further ambiguity by including qualifying phrases and contingencies such as those italicized below:

- “remained as non-reactive *as possible*”
- “to interact with their partner in a natural fashion, *but not to ...*”
- “to provide limited feedback (‘uh-hum,’ ‘yeah,’ and nods) *when it seemed appropriate to do so, especially when the subject seemed to be trying to elicit such a response*”
- “only *occasionally* asked for clarification *if a subject’s instruction was particularly unclear*”
- “not to interrupt the speaker, but to act naturally *otherwise ... addressees ... gave occasional non-verbal feedback, but they tried to avoid speaking themselves*”
- “addressees were looking at the speaker and gave *occasional non-verbal feedback, but they tried to avoid speaking themselves*”

In the moment-by-moment pace of dialogue, how did the quasi-addressee decide when and whether it was appropriate “to provide limited feedback” or permissible to “ask for clarification”? In short, if the descriptions in Table 3 corresponded to the actual behaviors of the quasi-addressee in each experiment, then their diversity and imprecision offer no evidence of well-defined experimental control.

### *Verifying experimental control*

The above criticism would be less relevant if the constraints reported in the articles were only summaries of more detailed operational definitions that had been given to the quasi-addressees and then subsequently used to confirm that the quasi-addressees carried them out. The tasks themselves presented a particular challenge to implementation because the speakers were the ones providing information, so they had the moment-by-moment initiative in the task. Each quasi-addressee was supposed to respond within the assigned constraints no matter what the speaker said or did — and to do so consistently for every speaker and in both experimental conditions. Also, they could not introduce systematic, unintended differences in their behavior, especially differences across conditions. In some cases, there was only one experimenter or confederate for the whole experiment. In others, each speaker had a different quasi-addressee, so all of the above standards had to be met by each of them. It was particularly important that quasi-addressees who happened to be in one experimental condition met the same standards as the quasi-addressees in the other condition.

Verification of the above degree of experimental control requires a systematic check of recordings of what the quasi-addressees actually did. However, none of these experiments reported any verification of the control procedure. (Some quasi-addressees did not appear to have been video-recorded in the visibility condition.) Only two articles commented on what actually happened; italics indicate problematic issues:

- “All of the listeners sat quietly while listening to the narrator retell the story, and *some occasionally* provided back-channel feedback (e.g., nodding or saying ‘uh-huh’). *In one case*, a listener provided a word for which a narrator was searching. Listeners also *sometimes* smiled or laughed aloud, since the cartoon story can be quite funny, *depending on* the narrator’s skill at retelling. Of course, in the screen condition, narrators could hear audible back-channel responses and laughter but could not see nods or smiles.” (Alibali et al., 2001, p. 175)
- “*Occasionally, there was some* auditory feedback (laughs, occasional uh-huh’s). The face-to-face condition *was similar*.” (Mol et al., 2009a, p. 105)

In order to ensure that quasi-dialogues do in fact ensure well-controlled experiments, future studies would have to provide both more precise specification of the permitted actions of addressees and empirical verification that the experimental procedures did in fact achieve the intended level of control.

### *Experimenter bias*

As shown in Table 1, the quasi-addressee was either a confederate or the experimenter in at least four of the experiments. As Kuhlen and Brennan (2013) pointed out,

Confederates' own expectations about the outcome of a study may cause them to inadvertently bias participants in favor of the experimental hypothesis. The power of confederates to bias behavior is a particular danger when the confederates know what type of behavior is expected from the subjects. It is even more problematic when the confederates know which experimental condition they are participating in. ... Even the most conscientious confederates are at risk of inadvertently shaping participants' behavior by giving verbal backchannels or nonverbal cues such as facial expressions, body posture, tone of voice, pauses, or eye gaze patterns. (p.60)

Clearly, these potential problems apply even more when the experimenter is the quasi-addressee, because he or she would certainly know what type of behavior is expected from the speakers, as would a confederate from the same research lab. Obviously, they would also know which experimental condition they were participating in. Therefore, eliminating the possibility of experimenter or confederate bias should be added to the control requirements outlined above. It is necessary to ensure, for example, that the back-channels of these quasi-addressees were not encouraging more gesturing in one condition than the other. (For a more comprehensive discussion of standards for research with confederates, see Kuhlen & Brennan, 2013.)

### *Summary*

Each experiment in Table 1 used a method that assumed the individual was the basic unit of analysis in a dialogue, namely, treating the actions of the other person as an external factor to be controlled. The issue here was whether this method should, by definition, be considered sounder than the use of free dialogues, on the basis that it provided tighter experimental control and was therefore better able to detect a significant difference between conditions. Our examination of the reported procedures shows that there were no clear definitions of the control procedures,

no verification that the procedures were implemented consistently, and no controls for possible experimenter bias. Therefore, our answer to the first question must be that there is no objective evidence that constraining the addressees created a better controlled experiment. As a result, we cannot accept that their findings are more credible than those of the experiments with free dialogues. Instead, we are left still needing an explanation that will indicate which of the two sets of results we should accept.

### **Question 2: Did the constraints in the experiments with quasi-dialogues create a potential confound?**

This section takes a dyadic view of dialogue, which leads to an entirely different question about the difference in the results from the two groups of visibility experiments. Far from being better controlled, could the constraints on quasi-addressees have inadvertently introduced a confound, namely, an unintended difference in the addressees' responsiveness in the visible versus not-visible conditions? This question arises from two observations. First, in the experiments with quasi-addressees, there was no mention of controlling their spontaneous visible responses. It is therefore possible that the quasi-addressees in the visible condition seemed more responsive because their speakers could see the uncontrolled visible responses; without these, the quasi-addressees in the not-visible condition could have seemed to their speakers to be much less responsive. Second, there is evidence that the lack of a responsive addressee itself can lower the speakers' rate of gesturing. The following sections will offer evidence for these two possibilities.

#### *Addressees' visible contributions to dialogue*

In the dyadic view, even when one person has the major speaking role, addressees are an active and influential part of the dialogue — both audibly and visibly. In particular, they have an essential role in the process of grounding, through which speakers and addressees “collaborate with each other moment by moment to try to ensure that what is said is also understood” (Schober & Clark, 1989, p.211). Speakers present information, but addressees must display that they have understood (Bavelas et al., 2000). The addressee's resources for moment-by-moment collaboration include much more than occasional words and prosody in the audible “back channel” (Yngve, 1970).

Because visibility designs manipulate *mutual* visibility, addressees in one condition have a rich array of visible responses that provide constant feedback to the speaker without interrupting the ongoing flow of conversation. For example,

*nodding* is often a silent alternative to “mhm” or “yeah”. The addressee’s *gaze pattern* is also important. In western cultures, gaze asymmetry is correlated with speaking roles; addressees look fairly constantly at their speaker while speakers tend to look around (e.g., Kendon, 1967). Two variations on this pattern are also communicative: Even within their turn, speakers occasionally look at their addressee; this creates a brief period of mutual gaze that, with high probability, elicits a back-channel response from the addressee (Bavelas, Coates, & Johnson, 2002). Also, addressees as well as speakers can use gaze as deixis, looking at something in order to “point” at it.

Addressees as well as speakers use *hand gestures* as part of the grounding process. Holler and Wilkin (2011a) reviewed the limited but consistent literature on gestures as addressee feedback. For example, Holler and Wilkin’s (2011b) experiment showed explicitly that addressees may use gestural mimicry (with or without words) to indicate when and how well they are understanding the speaker. In Bavelas, Gerwing, Allison, and Sutton’s (2011) analysis of grounding with gestures, speakers and addressees often coordinated their understanding with gestures.

Less well known are the rapid and precisely timed *conversational facial gestures* that both speakers and addressees use to convey information that is directly relevant to their dialogue (and not indicative of their emotional state; see review in Bavelas, Gerwing, & Healing, 2013a). Chovil’s (1989, 1991/1992) early studies of facial actions in dialogue included the faces of both participants. Addressees’ facial gestures can provide non-interruptive back-channel information such as comprehension, confusion, skepticism, amusement, and surprise. Even a “blank expression” is informative to a speaker who is trying to explain something. Smiles often function as back-channel responses (Brunner, 1979), and facial mimicry such as wincing can convey precise understanding of the speaker’s narrative point (Bavelas et al., 2000; Bavelas, 2007).

### *Interaction of addressee constraints with visibility conditions*

The visible communicative acts outlined above would interact with the two key features of the visibility experiments analyzed here: the presence or absence of mutual visibility and the presence or absence of constraints on the addressees; see Table 4. The following analysis assumes that the quasi-addressees were adhering to some degree with their constraints and were not responding freely, at least verbally. If so, it is useful to compare what the speakers would be likely to see and hear from the addressees in the visible and not-visible conditions of the free dialogues versus quasi-dialogues.

In the visible conditions (first row of Table 4), the addressees in the free dialogues could use any of the visible co-speech acts described above in addition

to unconstrained speech. The verbally constrained addressees in a quasi-dialogue had the same visible resources available and could, to some extent, use them to compensate for their verbal constraints. For example, instead of asking questions, the quasi-addressees could elicit further explanation by a confused or quizzical facial gesture, by shaking their heads, or by simply ceasing to nod. Notice that these and other visible acts would also not have violated an injunction against interrupting, although they typically overlap the speaker's speech.

In the not-visible conditions (second row of Table 4), more substantial differences between the free and quasi-dialogues could have occurred. Even without visible resources, the addressees in free dialogues still had an unconstrained range of audible alternatives. They were free to intervene verbally as much as they wished and in any way that they found helpful; they could interrupt, ask questions, describe their current state of understanding, formulate what they had understood so far, and even ask the speaker to proceed differently. In contrast, the constraints on the quasi-addressees were likely to have limited their ability to adapt to the lack of visibility. For example, as shown in Table 3, a common constraint was to prohibit interruption. In a study of the effects of mutual visibility on speech (using free dialogues), Boyle, Anderson, and Newlands (1994) found a significantly higher rate of interruption in the not-visible compared to the visible condition. Without gaze to regulate their turns, interrupting is a practical option for addressees who wish to gain a speaking turn. If both gaze and interruptions were precluded, it seems that the quasi-addressees could only sit in silence until the speaker paused or asked for a response. The other common constraint precluded asking questions, so they would have been less able to speak up in this way as well.

### *The effects of nonresponsive addressees on speakers' gesturing*

If the above analysis is plausible, then the addressees in the free dialogues would have been able to remain fairly responsive, even in the not-visible condition, by adapting as addressees often do in everyday life (e.g., when speaking on the phone). It is also possible to infer that the quasi-addressees in the not-visible condition would have been less able to adapt and would therefore seem much less responsive. If so, they would have provided minimal feedback to their speakers, thereby creating something closer to a monologue than to a dialogue. Lack of response from the addressee has been shown to affect speakers' rates of gesturing: Speakers gestured at a significantly lower rate to a nonresponsive addressee or no addressee at all than to a responsive addressee (Beattie & Aboudan, 1994). Speakers also gestured significantly less when their addressees were less attentive (Jacobs & Garnham, 2007). Finally, speakers gestured significantly less in a mono-

logue with no addressee than when talking to an (unconstrained) addressee either face-to-face or on the phone (Bavelas et al., 2008, 2013b).

### *Summary*

A variety of studies combine to suggest that the answer to Question 2 is yes, the use of quasi-dialogues could have confounded visibility with addressee responsiveness. It is possible that the significant difference in speakers' overall rate of gesturing between the two quasi-dialogue conditions could have been caused primarily by differences in their addressees' responsiveness and only incidentally and indirectly by mutual visibility. By the same reasoning, this potential confound would have affected free dialogues less because those addressees would have been able to compensate verbally in the not-visible condition. If so, the difference in the responsiveness of quasi-addressees in visible and not-visible conditions is a confound that provides a plausible alternative explanation for the findings of the quasi-dialogues. Because the free dialogues could minimize this confound, their findings of no significant visibility effect would have stronger internal validity.

Note that even if one rejects our answers to both Questions 1 and 2 and believes, instead, that quasi-dialogues were well-controlled experiments and that there was no confound, it is still necessary to accept that these two methods created two different contexts for gesture production. The quasi-dialogue experiments showed the effect of mutual visibility on the overall rate of gesturing by an individual who was talking to a minimally responsive addressee. The free dialogue experiments showed the effect of mutual visibility on overall gesture rate by a speaker who is talking to a free and unconstrained addressee. For those interested in generalizing to free dialogues, it would seem that there is no visibility effect. However, even that is not the full picture.

### **Question 3: Is overall gesture rate the best way to assess visibility effects?**

Throughout this article, we have been careful to refer to an "overall rate of gesturing", which is the omnibus measure used in all of the visibility experiments reviewed here. However, this is not the only measure used in some visibility designs. Table 5 summarizes experiments with free dialogues (including several of the studies in Table 2) that included more specific measures in order to capture a particular use or function of a gesture in a particular context.

Some experiments showed that speakers in free dialogues with a visible addressee used different kinds of gestures depending on whether their addressee was visible or not. As shown in Table 5, speakers in the visible condition used more

gestures that served *interactive functions* directly related to their dialogue rather than to the topic of the dialogue (Bavelas et al., 1992, 2008). In an experiment using the classic Tangram task, speakers in the visible condition *pointed* at the Tangram figures, which they did not do in the not-visible condition (de Ruiter et al., 2012). In another Tangram experiment, both the speakers and the addressees were more likely to use gestures that *mimicked* each other's gestures only in the visible condition (Holler & Wilkin, 2011b).

The form and location of gestures by speakers in free dialogues differed depending on whether the addressee was visible or not. As shown in Table 5, their gestures varied in *size*: Speakers in the visible condition made gestures describing an elaborate dress that were life-sized. That is, they made their gestures at the scale of (and often placed them on or around) their own bodies, whereas the speakers in the not-visible condition described the dress with gestures at the much smaller scale of the picture they were describing (Bavelas et al., 2008). Speakers in the visible condition placed their gestures in front of their upper body, whereas speakers in the not-visible condition made their gestures lower, where they would not be visible to an addressee if there were one (Holler et al., 2011). Participants who were following directions for building a Lego model in the visible condition often *exhibited* or *poised* a particular block for their director, that is, they brought it “into a conspicuous location and manifestly [held] it there for inspection” (Clark & Krych, 2004, p. 71). Builders in the not-visible condition never did this. Speakers in the visible condition who were describing a connected sequence of geometric figures made gestures that were *sequentially cohesive*, preserving their spatial relationships. In the not-visible condition, the speakers tended to pile their gestures into the same space, without preserving the spatial relationships between the figures that they were depicting (Healing & Gerwing, 2012).

Finally, speakers in free dialogues whose addressees could see them changed the relationship between their gestures and their words. In the visible conditions, speakers were more likely to use verbal deixis to refer to their topic gestures (Bavelas et al., 2008) or to their pointing gestures (Bangerter & Chevalley, 2007). Speakers in the not-visible conditions seldom used such deictics. Speakers in the visible condition were also more likely to use *non-redundant gestures*, which conveyed information that was not at all in the accompanying words, than did speakers in the not-visible condition (Bavelas et al., 2008). Similarly, de Ruiter et al. (2012) distinguished between *obligatory* gestures, which conveyed “disambiguating information that [was] not in the accompanying speech but [was] nevertheless essential for understanding” (p. 238) versus gestures that were not obligatory (i.e., the accompanying speech was sufficient without the gesture). In the visible condition, speakers were more likely to use obligatory gestures; in the not-visible condition, speakers produced almost no obligatory gestures. There was no effect

of mutual visibility on the non-obligatory gestures (de Ruiter et al., 2012). Finally, a semantic features analysis of the Bavelas et al. (2008) data compared the information conveyed in words versus gestures and found that visibility affected the *distribution* of information across gestures and words: Speakers conveyed more semantic information in gestures than in words in the visible condition and the reverse for the non-visible condition (Gerwing & Allison, 2011).

The common characteristic of these nine specific differences between gestures in the visible versus not-visible conditions in free dialogues is that the gestures in the visible conditions were arguably useful or even necessary for the addressee: Interactive gestures seek and evoke specific responses from the addressee (Bavelas et al., 1995, Study 2). Mimicry is an efficient way to display understanding, but it is only possible when the participants can see each other. Pointing works well when both the pointing gesture and the object are mutually visible. When depicting the features of a dress, gestures scaled to the speaker's body are more informative to the addressee than gestures that are much smaller. Gestures made in front of the speaker's upper body are visually more available to the addressee. Exhibiting or poising the object under consideration explicitly demonstrates to the addressee what the gesturer is proposing to do with it. Gestures that accurately represent the spatial relationships in a path give the addressee essential information about the picture as a whole. Finally, gestures with information that is not in the accompanying words (non-redundant or obligatory gestures) are an efficient way to convey information that the addressee needs, as is changing the distribution of information between words and gestures according to visibility condition.

It is noteworthy that all of the above ways in which speakers made their gestures helpful to their addressee involved choice and even effort by the speaker: The speakers in the visible conditions of the above experiments had to attend to the visual availability of their addressee and choose to include (or not to include) interactive gestures, mimicry, or pointing as part of their dialogue. They chose to make larger, life-sized gestures (rather than smaller ones); to bring their gestures up where the addressee could see them (rather than down in their laps); to make their actions into gestures by poising or exhibiting them for the addressee (rather than simply carrying on with the action); and to portray the relative locations of their gesture accurately (rather than ignoring their spatial relationship). Finally, they chose to draw attention to their gestures with deictic references (or not); to make gestures that were obligatory (versus simply redundant with their words); and to systematically shift information between gestures and words. Thus, mutual visibility led to choices and effort that had significant effects on specific features of the gestures. These are the documented visibility effects.

### Summary

The answer to Question 3 is no, overall gesture rate is not the best way to assess visibility effects; overall rate may even be misleading if it seems to indicate that visibility has no effect. A closer look at how speakers use their gestures reveals that visibility affects many aspects of gestures including the kinds of gestures, their size, location, and relationship to words. All of these differences seem to be done for the addressee's benefit.

### Discussion

The answers to our three main questions strongly imply that the conventional hypothesis and counter-hypothesis — namely, that the effects of visibility on overall gesture rate will confirm whether or not speakers make gestures to communicate — are simplistic and clearly inadequate for such a broad conclusion. Experiments that vary visibility with free dialogues show no effect on speakers' overall rate of gesturing, yet experiments that examine specific features of gestures show that visibility does affect *how* speakers gesture.

One way to reconcile these effects is to distinguish between social and communicative effects. As described above, speakers in a free dialogue gesture at a significantly higher overall rate than speakers in a monologue, regardless of visibility, which is a *social* effect. In a dialogue without visibility, these gestures are not beneficial to the addressee. In contrast, speakers in a free dialogue whose addressees can see them make gestures with specific features that are beneficial to their addressees; this is a *communicative* effect. Bavelas et al. (2008) separated these two effects by varying both dialogue/monologue and visibility/no-visibility and using multiple dependent measures. The overall rate of gesturing was significantly lower in the monologue condition than in either the face-to-face or telephone dialogue conditions, and this rate was not significantly different between the two free dialogue conditions (i.e., no visibility effect).<sup>2</sup> However, the same data showed strong visibility effects on more specific features of gestures that, as argued above, were communicative features designed for addressees.

More broadly, the findings reviewed here are congruent with two well-known theories of gesturing. First, the effects of mutual visibility summarized in Table 5 are clear evidence of Kendon's principle that a speaker

will select a model of formulation, not only in the light of a comparison between its adequacy of representation and the image that it is intended to convey, but also in the light of *what the current communication conditions are*. These include transmission conditions, ... but they also include an appreciation of the kind of

impact a gestural formulation may have on a recipient as compared to a verbal formulation (1987, p. 90; emphasis added).

When the current communication conditions included mutual visibility and the recipient needed the information, speakers shaped their gestures in ways that were beneficial or even necessary to their addressee. Without mutual visibility, they did not use their gestures in these ways.

Second, it makes sense that speakers in a not-visible condition would avoid making gestures that would be necessary for their addressee's understanding and would instead shift the information into words. But that still leaves another puzzle: If the information is in the words and the gesture would be useless to an addressee who will not see it, then why bother to gesture at all? The finding that speakers do make gestures that no addressee will see fits McNeill's theory of *growth points*, which are "the initial form of thinking out of which speech-gesture organization emerges" (McNeill, 2000, p. 314). In this view (as in Kendon's, 1980), the process of utterance formulation includes both speech and gesture from the beginning. Gestures are not an optional add-on later in the process. In a later version, McNeill (2005) added the "social, interactive context" to the growth point:

A growth point is always connected to the discourse context, including any social interactive aspects. ... The context of speaking, including the social-dynamic context, affects this process throughout, including the growth point itself. (p. 82)

Combining this more social view of the growth point with the Kendon (1987) principle above, we can propose that gesturing is part of speaking in a social context, and shaping gestures to fit the social interactive context is part of this process as well.

One nagging question is how gesturers spontaneously formulate and enact these gesture-speech relationships in the second-by-second rush of dialogue, while continuously adapting to the needs of the interaction. The combined complexity and rapidity with which speakers make these precise selections may seem to be contradictory, but there is a model from another field that we could borrow. Although Kahneman (2011) has not, to our knowledge, applied his dual process model of judgment and decision making to language, it seems to fit language use in dialogue: *System 1* is fast and automatic "with little or no effort and no sense of voluntary control," although it generates "surprisingly complex patterns" (Kahneman, 2011, pp. 20–21). *System 1* fits the complex choices listed in Table 5. *System 2* is a slower, conscious process "associated with the subjective experience of agency, choice, and concentration" (p. 21). This description fits examples such as thinking about the best way to phrase something or correcting oneself.

## Implications for future research

We cannot assume that quasi-dialogues will produce the same findings as free dialogues. In psycholinguistics, Lockridge and Brennan's (2002) results for speakers talking to real addressees in free dialogues were strikingly different from those in a previous experiment with a confederate addressee (who heard the same story repeatedly). Brown-Schmidt and Tanenhaus (2008) found significant differences in addressees' reactions to the same stimulus words from an experimenter who was following a script versus an interacting partner. The 14 gesture experiments expand this literature in two ways: (a) There are a large number of experiments using the same method for the same topic. (b) The difference holds for a broader class of constrained addressees, including confederates, experimenters, and other participants. The failure of the results of quasi-dialogues to generalize to free dialogues is particularly problematic for conversational gestures which, by definition, occur in free dialogues. More broadly, in any area where the phenomenon of interest naturally occurs in dialogue, it would be wise to conduct the experiments with free dialogues.

Similarly, visibility designs may have outlived their usefulness. We propose that the evidence reviewed in this article strongly supports the conclusion that speakers' gestures are social or communicative in more precise and subtle ways than conventional visibility designs have captured. Visibility does not significantly affect speakers' overall rate of gesturing in free dialogues. However, having an unresponsive addressee or being in a monologue does reduce the overall rate. Speakers also make choices about *how* to gesture according to what is useful and available to their addressees (which implies a need for more and varied studies about addressees; e.g., Kendon, 1994). Moreover, conventional visibility experiments also serve cognitive theories poorly, because these theories are often little more than a vague default alternative. More well-developed cognitive theories deserve designs of their own, with more closely relevant experimental conditions. In short, increasing the number and variety of designs that include free dialogues as well as more specialized independent and dependent variables should lead to a better understanding of conversational gestures.

Finally, the choice of an individual or dyadic unit of analysis must become an explicit, open matter for discussion, not just a latent assumption behind a methodological preference. Gesture research could contribute to the ongoing debate about whether dialogue consists of the acts of autonomous individuals or is intrinsically collaborative (Clark, 1996; Schober & Clark, 1989). The limited experimental evidence supports a collaborative theory of gestures: Bavelas et al. (2011), Clark & Krych (2004), Holler & Wilkin (2011b). However, with the exception of these and a few other studies (e.g., Rimé, 1982), most gesture researchers with a dyadic

perspective still use tasks in which one person has the primary speaking role, and the focus is almost exclusively on the gestures of that person, ignoring the gestures of the addressee in the same dialogue. More research is needed on what goes on between the participants, especially on how they affect each other's gesturing (e.g., entrainment).

## Acknowledgements

Special thanks to our colleague Dr. Jennifer Gerwing for her valued contributions during the review, analysis, and editing process. We acknowledge the support of the Social Sciences and Humanities Research Council of Canada through a grant to the first author. An earlier version of this material was presented at the 2012 ISGS conference in Lund, Sweden.

## Notes

1. Some of these studies also included measures of more specific features of gestures, which showed different effects of visibility. We will return to these measures in a later section.
2. Bavelas et al. (2008) initially reported a regression for overall rate of hand gestures that showed a significant visibility effect. However, this regression did not control for dialogue. A later re-analysis added a control for dialogue and found *no* significant visibility effect. This result is congruent with the post-hoc finding (2008, p. 512) of no significant difference between the face-to-face and telephone dialogues, and both findings were replicated in Bavelas et al. (2013b).

## References

- Alibali, Martha W., Dana C. Heath, & Heather J. Myers (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.
- Aronson, Eliot & J. Merrill Carlsmith (1968). Experimentation in social psychology. In Gardner Lindzey & Elliot Aronson (Eds.), *The handbook of social psychology* (2nd ed., Vol. 2, pp. 1–79). Reading, MA: Addison-Wesley.
- Bangerter, Adrian & Eric Chevalley (2007). Pointing and describing in referential communication: When are pointing gestures used to communicate? In *CTIT Proceedings of the MOG 2007 Workshop on Multimodal Output Generation*. Retrieved from <http://www.utwente.nl/ctit/library/proceedings/MOG2007.pdf>
- Bavelas, Janet B. (2005). The two solitudes: Reconciling social psychology and language and social interaction. In Kristine Fitch & Robert Sanders (Eds.), *Handbook of language and social interaction* (pp. 179–200). Mahwah, NJ: Erlbaum.

- Bavelas, Janet B. (2007). Face-to-face dialogue as a micro-social context. In Susan D. Duncan, Justine Cassell, & Elena T. Levy (Eds.), *Gesture and the dynamic dimension of language. Essays in honor of David McNeill* (pp. 127–146). Amsterdam: John Benjamins.
- Bavelas, Janet B., Nicole Chovil, Linda Coates, & Lori Roe (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, 21, 394–405.
- Bavelas, Janet B., Nicole Chovil, Douglas A. Lawrie, & Allan Wade (1992). Interactive gestures. *Discourse Processes*, 15, 469–489.
- Bavelas, Janet B., Linda Coates, & Trudy Johnson (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79, 941–952.
- Bavelas, Janet B., Linda Coates, & Trudy Johnson (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, 52, 566–580.
- Bavelas, Janet B., Jennifer Gerwing, Meredith Allison, & Chantelle Sutton (2011). Dyadic evidence for grounding with abstract deictic gestures. In Gale Stam & Mika Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 49–60). Amsterdam: John Benjamins (*Gesture Studies Series*).
- Bavelas, Janet B., Jennifer Gerwing, & Sara Healing (2013a). Hand gestures and facial displays in conversational interaction. In Thomas Holtgraves (Ed.), *Handbook of language and social psychology*. Oxford: Oxford University Press.
- Bavelas, Janet B., Jennifer Gerwing, & Sara Healing (2013b). The effect of dialogue on demonstrations: Direct quotations, facial portrayals, hand gestures, and figurative references. Manuscript submitted for publication.
- Bavelas, Janet B., Jennifer Gerwing, Chantelle Sutton, & Danielle Prevost (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, 495–520.
- Beattie, Geoffrey & Rima Aboudan (1994). Gestures, pauses and speech: An experimental investigation of the effects of changing social context on their precise temporal relationships. *Semiotica*, 99 (3/4), 239–272.
- Boyle, Elizabeth, Anne Anderson & Alison Newlands (1994). The effects of visibility on dialogue and performance in a cooperative problem solving task. *Language and speech*, 37 (1), 1–20.
- Brown-Schmidt, Sarah & Michael K. Tanenhaus (2008). Real-time investigation of referential domains in unscripted conversations: A targeted language game approach. *Journal of Cognitive Science*, 32, 643–684.
- Brunner, Lawrence J. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology*, 37, 728–734.
- Chovil, Nicole (1989). Communicative functions of facial displays in conversation. Unpublished PhD dissertation, University of Victoria, Victoria, B.C.
- Chovil, Nicole (1991/1992). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, 25, 163–192.
- Clark, Herbert H. (1985). Language use and language users. In Gardner Lindzey & Elliott Aronson (Eds.), *Handbook of social psychology* (3rd ed., pp. 179–231). New York: Harper and Row.
- Clark, Herbert H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H. & Meridith Krych (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, 62–81.
- Cohen, Akiba A. (1977). The communicative functions of hand illustrators. *Journal of Communication*, 27, 54–63.

- Cohen, Akiba A. & Randall P. Harrison (1972). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology*, 28 (2), 276–279.
- Danziger, Kurt (1990). *Constructing the subject: Historical origins of psychological research*. Cambridge, UK: Cambridge University Press.
- de Ruiter, Jan P., Adrian Bangerter, & Paula Dings (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.
- Emmorey, Karen & Shannon Casey (2001). Gesture, thought and spatial language? *Gesture*, 1, 35–50.
- Gerwing, Jennifer & Meredith Allison (2011). The flexible semantic integration of gestures and words: Comparing face-to-face and telephone dialogues. *Gesture*, 11 (3), 308–329.
- Healing, Sara & Jennifer Gerwing (2012). Spatial cohesiveness in gesture sequences: Effects of mutual visibility. Paper presented at the *Fifth Conference of the International Society for Gesture Studies*, Lund, Sweden. Abstract retrieved from [http://www.gesturestudies.com/isgs2012/ISGS5\\_BOA\\_120718.pdf](http://www.gesturestudies.com/isgs2012/ISGS5_BOA_120718.pdf)
- Holler, Judith, Mark Tutton, & Katie Wilkin (2011). Co-speech gestures in the process of meaning coordination. In *Proceedings of the 2nd GESPIN- Gesture in Speech and Interaction Conference*, Bielefeld, September 5–7, 2011.
- Holler, Judith & Katie Wilkin (2011a). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics*, 43, 3522–3536.
- Holler, Judith & Katie Wilkin (2011b). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, 35, 133–153.
- Jacobs, Naomi & Alan Garnham (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, 56, 291–303.
- Kahneman, Daniel (2011). *Thinking, fast and slow*. Toronto: Doubleday Canada.
- Kendon, Adam (1967). Some functions of gaze direction in social interaction. *Acta Psychologica*, 26, 22–63.
- Kendon, Adam (1980). Gesticulation and speech: Two aspects of the process of utterance. In Mary Ritchie Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague: Mouton.
- Kendon, Adam (1987). On gesture: Its complementary relationship with speech. In Aron Siegman & Stanley Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 65–97). Hillsdale, NJ: Lawrence Erlbaum.
- Kendon, Adam (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, 27 (3), 175–200.
- Krauss, Robert M., Yihsiu Chen, & Rebecca Gottesman (2000). Lexical gestures and lexical access: A process model. In David McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge, UK: Cambridge University Press.
- Krauss, Robert M., Robert Dushay, Yihsiu Chen, & Frances Rauscher (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, 31 (6), 533–552.
- Kuhlen, Anna K. & Susan E. Brennan (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, 20 (1), 54–72.
- Lockridge, Calion B. & Susan E. Brennan (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, 9 (3), 550–557.

- McNeill, David (2000). Catchments and contexts: Non-modular factors in speech and gesture production. In David McNeill (Ed.), *Language and gesture* (pp.312–328). Cambridge: Cambridge University Press.
- McNeill, David (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- Mill, John Stuart (2008). *A system of logic, ratiocinative and inductive, being a connected view of the principles of evidence, and the methods of scientific investigation* (8th ed., Vol. 1, pp. 478–502). Retrieved from <http://www.cddc.vt.edu/gutenberg/2/6/4/9/26495/26495-pdf.pdf> (Original work published in 1882).
- Mol, Lisette, Emiel Krahmer, Alfons Maes, & Marc Swerts (2009a). Communicative gestures and memory load. In Niels A. Taatgen & Hedderik van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, July 29–August 1, Amsterdam (pp.1569–1574). Austin, TX: Cognitive Science Society.
- Mol, Lisette., Emiel Krahmer, Alfons Maes, & Marc Swerts (2009b). The communicative import of gestures. Evidence from a comparative analysis of human-human and human-machine interactions. *Gesture*, 9 (1), 97–126.
- Pine, Karen. J., Daniel J. Burney, & Ben Fletcher (2010). The semantic specificity hypothesis: When gestures do not depend upon the presence of a listener. *Journal of Nonverbal Behavior*, 34, 169–178.
- Rimé, Bernard (1982). The elimination of visible behavior from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology*, 12, 113–129.
- Schober, Michael F. & Herbert H. Clark (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211–232.
- Sears, Robert R. (1951). A theoretical framework for personality and social psychology. *American Psychologist*, 6, 476–483.
- Thibault, John W. & Harold H. Kelley (1959). *The social psychology of groups*. New York: Wiley
- Yngve, Victor 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.

Table 1. Experiments that found a significant visibility effect on overall gesture rate

		Experiments				
Summary of Experiment	Cohen (1973)	Krauss et al., Exps. 1 & 2 (1995)	Alibali et al. (2001)	Emmorey & Casey (2001)	Mol et al. (2009a)	Mol et al. (2009b)
<i>Participants:</i>	24 males	18 males & 18 females	8 males & 8 females	10 males & 10 females <sup>b</sup>	19; genders unspecified	20 <sup>c</sup> ; genders unspecified
<i>Speakers</i>						
<i>Participants:</i>	A female	A same-sex confederate	8 males & 8 females	The experimenter	Unspecified number and genders of confederates	20 randomly assigned with constraints; genders unspecified
<i>Addressees</i>						
<i>Visibility</i>	Face to face vs. intercom	Face to face vs. intercom	Face to face vs. screen	Face to face vs. partial screen	Face to face vs. screen	Face to face vs. screen
<i>Manipulation</i>	(Within subjects)	(Between subjects)	(Within subjects)	(Between subjects)	(Between subjects)	(Between subjects)
<i>Task</i>	Give directions to locations	Describe sounds & designs	Retell a cartoon	Give directions re placing blocks	Retell a cartoon	Retell short excerpts from a cartoon
<i>Operational Definition of Gestures</i>	Speakers' illustrators: number of "hand gestures that were related...to what the subject was saying while making the gestures." (p.278)	Speakers' "gestures that accompanied each... description." (p.541)	Speakers' "representational gestures [depicting] semantic content via the shape, placement, and/or motion trajectory of the hands." (p.175)	No operational definitions. The examples of speakers' gestures include gestures of shape and motion as well as pointing. (pp.39-45)	Speakers' gestures, i.e., not including "other movements such as self-adjustment." (p.106)	Speakers' gestures, i.e., not including "other movements such as self-adjustment." (p.1571)
<i>Quantitative Measure</i>	Number of illustrators	The gesture rate per minute	Rate of representational gestures per 100 words and per minute	Mean number of gestures	Rates of all gestures per 100 words and per second	Rate of all gestures per 100 words and per second

Table 1. (continued)

		Experiments					
Summary of Experiment	Cohen & Harrison (1973)	Cohen (1979)	Krauss et al., Exps. 1 & 2 (1995)	Alibali et al. (2001)	Emmorey & Casey (2001)	Mol et al. (2009a)	Mol et al. (2009b)
Visible vs.	The mean number of illustrators was significantly higher in the face-to-face condition than in the intercom condition. (p.278)	The means for the number of illustrators and for the rate of illustrators per second were both higher in the face-to-face condition than in the intercom condition. (p.59). No significance test reported. <sup>d</sup>	The mean gesture rate per minute was significantly higher in the face-to-face intercom condition. (p.541)	The mean rates of gestures per 100 words and gestures per minute were significantly higher in the face-to-face condition than in the screen condition. (pp.177-178)	The mean percentage of orientation instructions accompanied by a gesture was significantly higher in the face-to-face condition than in the screen condition. (p.42)	The mean rates of all gestures per second were significantly higher in the face-to-face condition than in the screen condition. (p.107)	The mean rates of all gestures per 100 words and all gestures per second were significantly higher in the face-to-face condition than in the screen condition. (p.1572)

N.B. This table includes only the broadest and most comparable gesture measures across studies. It also includes Ns, measures, and statistical tests only for the visibility manipulation and not for any other variables or conditions.

<sup>a</sup> These Ns do not include an "alone" condition in which speakers "had the opportunity to 'practice' giving the directions before they actually spoke to the experimenter [face-to-face]" (p.56).

<sup>b</sup> These Ns do not include two "gesture not permitted" conditions in which speakers were sitting on their hands.

<sup>c</sup> These Ns do not include a condition reported in this article that was from a previous study by the same authors.

<sup>d</sup> The author reported a main effect for a "form of interaction" variable, which included the face to face condition, the intercom condition, and the practicing-alone condition. There was no direct statistical comparison of the face-to-face and intercom conditions.

Table 2. Experiments that did not find a significant visibility effect on overall gesture rate

		Experiments				
Summary of Experiment	Rimé (1982)	Bavelas et al., (Exp. 2) (1992)	Pine et al. (2010)	Holler et al. (2011)	de Ruiter et al. (2012)	Bavelas et al. (2013) <sup>a</sup>
Participants: <i>Speakers</i>	20 males	20 females	42 females and 26 males. Pairs alternated in the roles of speaker and addressee	16; genders unspecified	48; genders unspecified	20; genders unspecified
Participants: <i>Addressees</i>	20 males	20 females	(See above.)	16; genders unspecified	48; genders unspecified	20; genders unspecified
Visibility Manipulation	Face to face vs. screen (Between subjects)	Face to face vs. screen (Between subjects)	Face to face vs. screen (Within subjects)	Face to face vs. screen (Between subjects)	Face to face vs. screen (Between subjects)	Face to face vs. telephone (Between subjects)
Task	Discuss movies that they like	Tell a close call story of an 18th century dress	Describe a picture	Describe a set of praxic and non-praxic objects	Describe Tangram figures	Describe an abstract line drawing
Operational Definition of Gestures	Both participants' "communicative gestures, that is, any gesture accompanying or paralleling the content or rhythm of the verbal flow." (p. 119)	Speakers' topic gestures, i.e., gestures depicting "information related to the topic at hand (e.g., details of the story being told)." (p. 473)	Speakers' total gestures and also topic gestures, i.e., depicting "some aspect of the current topic of conversation, in this case, the dress." (pp. 507-508)	Speakers' gestures, including iconic, metaphorical, deictic, and interactive. (Section 2.4.2)	Speakers' "iconic gestures [that] illustrated a particular feature of the target (e.g., shape)." (p. 238)	Speakers' "purpose and meaningful hand movements that were synchronized with speech and that illustrated the speaker's words." (pp. 26-27)

Table 2. (continued)

Experiments							
Summary of Experiment	Rimé (1982)	Bavelas et al., (Exp. 2) (1992)	Bavelas et al. (2008)	Pine et al. (2010)	Holler et al. (2011)	de Ruiter et al. (2012)	Bavelas et al. (2013) <sup>a</sup>
Quantitative Measure	Frequency of communicative gestures and percentage of time spent gesturing	The rate of topic gestures per minute	All gestures' rate per minute and per 100 words. Topic gestures' rate per minute and per 100 words	For the whole sample: mean number of gestures per object. For a subset of 40: mean number of gestures per minute	Number of gestures per trial divided by the number of words produced by that speaker during the same trial	The rate of iconic gestures per 100 words	The rate of gestures per 100 words and per minute
Visible vs. Not-visible	There was no significant difference between the face-to-face condition and the screen condition for either the mean number of communicative gestures or the percentage of time spent gesturing (p. 121)	There was no significant difference between the face-to-face condition and the screen condition for the mean rates of topic gestures per minute. (pp. 482–483)	There was no significant difference between the face-to-face condition and the telephone condition for the rates of all gestures or topic gestures, either per minute or per 100 words. (pp. 511–512)	There was no significant difference between the face-to-face condition and the screen condition for the mean rates of object or gestures per minute. (p. 174)	There was no significant difference between the face-to-face condition and the screen condition for the mean rates of gestures per word. (Section 3.2)	There was no significant difference between the face-to-face condition and the telephone screen condition for the mean rates of iconic gestures per 100 words. <sup>b</sup>	There was no significant difference between the face-to-face condition and the telephone condition for the mean rates of gestures per 100 words minute.(Table 5)

N.B. This table includes only the broadest and most comparable gesture measures across studies. It also includes Ns, measures, and statistical tests only for the visibility manipulation and not for any other variables or conditions.

<sup>a</sup> These data were gathered but not analyzed as part of the Bavelas et al. (2008) experiment.

<sup>b</sup> Jan P. de Ruiter (Personal communication, July 13, 2012). This information is not in the published article.

**Table 3.** Descriptions of the constraints on addressees in the quasi-dialogues

- In the visible condition, the confederate “stood in the door and performed a nonreactive role.”<sup>a</sup> The subject was told “not to ask her if she understood his directions”. (Cohen & Harrison, 1973, p.277)
- The experimenter “stood facing the respondent and remained as non-reactive as possible”<sup>a</sup> (Cohen, 1977, p. 58)
- “Confederates were instructed to interact with their partner in a natural fashion, but not to interrupt the descriptions or to ask questions. They were allowed to provide limited feedback (‘uh-hum,’ ‘yeah,’ and nods) when it seemed appropriate to do so, especially when the subject seemed to be trying to elicit such a response.” (Krauss et al., 1995, p. 541)
- “Listeners were instructed to listen carefully and were told that they would later be asked to retell the story to the experimenter. Listeners were also instructed not to ask any questions.” (Alibali et al., 2001, p. 174)
- “The experimenter said very little during the task, and only occasionally asked for clarification if a subject’s instruction was particularly unclear.” (Emmorey & Casey, 2001, p. 38)
- “We prevented true dialogue from happening by instructing addressees not to interrupt the speaker, but to act naturally otherwise. Thus, addressees were looking at the speaker and gave occasional non-verbal feedback, but they tried to avoid speaking themselves.” (Mol et al., 2009a, p. 102)
- “Addressees were instructed not to interrupt the speaker.” (Mol et al., 2009b, p. 1571)

<sup>a</sup> There was no information about the not-visible condition.

**Table 4.** Differences in addressees’ communicative resources as a function of free vs. quasi-dialogues and of visible vs. not-visible conditions.

Experimental Condition	Free dialogues	Quasi-dialogues
Mutually Visible	Unconstrained speech with full use of visible resources: - nods, head shakes - gaze - hand gestures - facial gestures	Constrained speech but still able to use visible resources: - nods, head shakes - gaze - hand gestures - facial gestures
Not Mutually Visible	Unconstrained speech with - no nods, head shakes - no gaze - no hand gestures - no facial gestures but free to adapt verbally	Constrained speech with - no nods, head shakes - no gaze - no hand gestures - no facial gestures but less free to adapt verbally

**Table 5.** Effects of visibility on specific features of gestures

Specific Gesture Feature	Experiments	Significant Effect of Visibility
Rate of Gestures with Interactive Functions	Bavelas et al. (1992) <sup>a</sup> Bavelas et al. (2008) <sup>a</sup>	Lower rate of interactive gestures in the not-visible condition
Mimicry of Interlocutor's Gestures	Holler & Wilkin (2011b)	Less mimicry of each other's gestures in the not-visible condition
Pointing	de Ruiter et al. (2012) <sup>a</sup>	Lower rate of pointing in the not-visible condition
Size of Gestures	Bavelas et al. (2008) <sup>a</sup>	Smaller gestures in the not-visible condition
Position of Gestures	Holler et al. (2011) <sup>a</sup>  Clark & Krych (2004)	Gesturing less likely to be in the gesture space in the not-visible condition  Gestures displaying an object or its placement to addressee only occurred in the visible condition
Sequential Cohesiveness	Healing & Gerwing (2012)	Gestures about a path were not sequentially connected in the not-visible condition
Verbal Deictic References to Gestures	Bavelas et al. (2008) <sup>a</sup>  Bangerter & Chevalley (2007)	Lower rates of gestures with a deictic expression in the not-visible condition  Lower rate of deictic references in the not-visible condition
Redundancy of Gestures with Words	Bavelas et al. (1992, 2008) <sup>a</sup> de Ruiter et al. (2012) <sup>a</sup>	Lower rate of non-redundant (obligatory) gestures in the not-visible condition
Shift in Distribution of Information between Words and Gestures		Less information in gestures and more in words in the not-visible condition

<sup>a</sup> As shown in Table 2, these experiments also found no significant difference between visible and not-visible conditions for the overall rate of gesturing.

*Authors' address*

Janet Bavelas and Sara Healing  
Department of Psychology  
University of Victoria  
P.O. Box 1700  
Victoria, B.C., V8W 2Y2  
Canada  
Correspondence to Janet Bavelas  
bavelas@uvic.ca, shealing@uvic.ca

*About the authors*

**Janet Bavelas** is an active Emeritus Professor of Psychology at the University of Victoria (Canada). Her experimental research focuses primarily on the unique features of face-to-face dialogue, which include moment-by-moment collaboration between speaker and addressee as well as visible co-speech acts such as hand and facial gestures. She also applies this research to non-lab settings.

**Sara Healing** recently completed her M.Sc. in Psychology at the University of Victoria. She is co-author of the chapter on hand and facial gestures in the *Oxford handbook of language and social psychology*, as well as an experimental article on dialogue and demonstrations and other articles on medical and psychotherapeutic dialogues.