Available online at www.sciencedirect.com
ScienceDirect
Journal of Memory and Language 58 (2008) 495–520
Journal of Memory and Language
www.elsevier.com/locate/jml

Gesturing on the telephone: Independent effects of dialogue and visibility ☆

Janet Bavelas *, Jennifer Gerwing, Chantelle Sutton, Danielle Prevost

Department of Psychology, University of Victoria, P.O. Box 3050, Victoria, BC, Canada V8W 3P5

Received 7 November 2006; revision received 21 February 2007
Available online 1 May 2007

Abstract

Speakers often gesture in telephone conversations, even though they are not visible to their addressees. To test whether this effect is due to being in a dialogue, we separated visibility and dialogue with three conditions: face-to-face dialogue (10 dyads), telephone dialogue (10 dyads), and monologue to a tape recorder (10 individuals). For the rate of gesturing, both dialogue and visibility had significant, independent effects, with the telephone condition consistently higher than the tape recorder. Also, as predicted, visibility alone significantly affected *how* speakers gestured: face-to-face speakers were more likely to make life-size gestures, to put information in their gestures that was not in their words, to make verbal reference to their gestures, and to use more gestures referring to the interaction itself. We speculate that *demonstration*, as a modality, may underlie these findings and may be intimately tied to dialogue while being suppressed in monologue.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Gestures; Telephone; Face-to-face dialogue; Visibility; Demonstration

☆ The Social Sciences and Humanities Research Council of Canada supports this program of research on face-to-face dialogue with a grant to the first author. We also thank the many individuals who assisted in the analyses: Meredith Allison, Laura Fawcett, Trina Holt, Paula Romagosa, Erin Boone, Patricia Wallis, and Jon Woods. Herb Clark contributed insightful comments to the earliest version of this article. Mary Lesperance and Angelia Vanderlaan gave valuable statistical advice, and A.A.J. Marley helped us sort out some perception questions. We also acknowledge Susan Goldin-Meadow, Meredyth Krych Appelbaum, Jan Peter de Ruiter, and the other reviewers who remained anonymous; their comments, both encouraging and critical, substantially improved the manuscript. In previous years, several students did pilot studies investigating the puzzle of why speakers gesture on the phone, all of which moved us toward the dialogue hypothesis. These students included Julia Allain, Ross Fleming, and Sheryl Shermack.
* Corresponding author. Fax: +1 250 727 6573.
E-mail address: bavelas@uvic.ca (J. Bavelas).

Introduction

Conversational hand gestures are those that accompany and illustrate speech. Speakers improvise these gestures along with their words, so that words and gestures are coordinated in both timing and meaning. This spontaneity and synchrony distinguishes them from emblems, which are stereotypic hand signals typically used in the absence of speech, and from adaptors (e.g., scratching). Even casual observation reveals that speakers often gesture when talking on the telephone. This phenomenon is more than a curious oddity, because it has been central to the ongoing debate about the communicative nature of conversational hand gestures. One important method for investigating whether or not gestures have a communicative function has been to vary *visibility* between speaker and addressee (Alibali,

0749-596X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2007.02.004

Heath, & Myers, 2001; Bavelas, Chovil, Lawrie, & Wade, 1992, Exp 2; Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss, Dushay, Chen, & Rauscher, 1995; Rimé, 1982; see Table 1 for a summary of these seven studies). In all of these experiments, speakers gestured at a higher rate when they spoke to a visible addressee than when speaking to someone on an intercom or through a partition. One plausible interpretation is that these speakers used their gestures to convey information to an addressee who could see them.

However, although the mean was lower when the addressee was not visible, it did not drop to zero, which is difficult for a communicative theory to explain. The default hypothesis has been that these gestures must serve an individual (e.g., cognitive) function for the speaker him- or herself. Only a few authors have hinted at an alternative explanation. In addition to the possibility of an encoding function for the speaker's benefit, Cohen and Harrison (1973) speculated that the gestures no one would see might be "a habit" (p. 277), presumably left over from "the more 'natural' use of illustrators when there is someone present who can see them" (p. 279). Clark (1996, pp. 179–180) took a stronger position, proposing that if gestures, like intonation, are *integral* to the composite signals typical of face-to-face dialogue, then "it should be difficult to produce the speech without the gesture" (p. 179). We agree with this position and go a step further to examine more closely *which* characteristic of face-to-face dialogue might be so closely tied to gesture and whether such a characteristic might still be present on the telephone.

In the following, we will emphasize two salient characteristics of face-to-face dialogue: The participants can see each other (visibility), and they are speaking with each other (dialogue). We hypothesize that visibility and dialogue can have independent effects and that dialogue may be the variable that is eliciting gestures on the telephone, even in the absence of visibility. To test this proposal, we compared both whether the speaker was visible to the addressee or not and whether the speaker was in a dialogue or a monologue. That is, we investigated not only the traditional independent variable of visibility but also the less obvious variable of dialogue. The dependent variables included not only the rate of gestures but differences in how speakers used their gestures in these conditions.

Dialogue as a variable

The experimental studies of gestures that have manipulated visibility have usually overlooked the possibility that dialogue itself could be having an effect. Although talking through an intercom or a partition did differ from talking face to face on the intended variable of visibility, the two conditions were still similar in that both were dialogues rather than monologues, that is, they involved interaction with another person. We propose that dialogue itself was an unrecognized variable that may have elicited some gestures, a proposal that can only be tested by adding a monologue condition.

To understand why dialogue might have an effect, it is first necessary to consider the nature of language use in face-to-face dialogue. Several scholars have proposed that, in contrast to other formats (e.g., written text, telephone, or email), face-to-face dialogue is the basic or fundamental setting for language use (e.g., Bavelas & Chovil, 2000, 2006; Bavelas, Hutchinson, Kenwood, & Matheson, 1997; Clark, 1996; Fillmore, 1981; Garrod & Pickering, 2004; Goodwin, 1981; Levinson, 1983; Linell, 2005). Face-to-face dialogue is humans' first language in both an evolutionary and developmental sense, and it remains the most common form of language use in everyday life. It is therefore useful to note the features that make face-to-face dialogue unique. For example, Clark (1996, pp. 9–10) enumerated 10 features of face-to-face conversations, a combination that does not occur in any other form of language use; see Table 2. However, while no other format for language use has all 10 of these features, a spontaneous dialogue on the telephone comes very close. The participants can hear each other with no perceptible delay; their messages fade rapidly and leave no record; they both can act at once and (if they choose) simultaneously; they act extemporaneously and are not scripted or playing a role. The only differences are that they are not in the same physical environment and cannot see each other. In contrast, talking into a tape recorder has only three of the 10 features; see Table 2. If a dialogue on the telephone is similar in so many ways to a face-to-face dialogue, and if (as Clark, 1996, proposed) gesturing is integral to face-to-face dialogue, then speakers may continue to gesture in telephone dialogues. Even participants giving descriptions to a hypothetical addressee gestured frequently (Bavelas, Kenwood, Johnson, & Phillips, 2002), so it is plausible that speakers would also gesture when talking to a real and responsive addressee whom they cannot see.

In brief, the presence of dialogue may have an effect on speakers' gesturing, independent of visibility. There is some suggestive evidence for such an effect in two previous experiments. In addition to face-to-face and intercom conditions, Cohen (1977) had a tape-recorder condition in which participants "had an opportunity to 'practice' giving the directions before they actually spoke to the experimenter [face to face]" (p. 56). During these three practice trials, their gesture per second rate was lower than in the intercom condition, as we would predict. However, there are several limitations of these findings for our purposes: First, note that the instructions (quoted above) were different for the tape-recorder condition, which was practice, whereas the face-to-face

Table 1
Summary of previous visibility experiments

Experiment[a]	Experimental design			Gesturing in each condition			Not-visible/visible[d]
	Not-visible Condition	Speaker's task	Addressee status (and constraints)	Face-to-face condition	Not-visible condition	Difference	
Cohen and Harrison (1973)	Intercom	Giving directions to a location	Confederate ("a nonreactive role" p. 277)	Hand illustrators: *M* frequency = 8.65[b]	Hand illustrators: *M* frequency = 4.96[b]	$p < .005$	.57
Cohen (1977)	Intercom	Giving directions between two locations	Experimenter ("as non-reactive as possible" p. 58)	Hand illustrators: *M* rate = 0.20 per second[b]	Hand illustrators: *M* rate = 0.10 per second with addressee; *M* rate = .07 when alone[b]	$p < .01$ (for main effect)[c]	.50
Rimé (1982)	Partition	Giving opinions on movies	Other participants (with no constraints)	Communicative gestures: *M* frequency = 13.00 (*SD* = 11.80)	Communicative gestures: *M* frequency = 5.10 (*SD* = 6.14)	$p \leqslant .10$	.39
Bavelas et al. (1992, Exp. 2)	Partition	Telling a personal close-call story	Other participants (with no constraints)	Topic gestures: *M* rate = 20.75 per min (*SD* = 7.69)	Topic gestures: *M* rate = 18.43 per min (*SD* = 8.72)	Topic gestures: p = n.s.	.89
				Interactive gestures: *M* rate = 4.38 per min (*SD* = 2.38)	Interactive gestures: *M* rate = 2.09 per min (*SD* = 1.86)	Interactive gestures: $p = .03$	.48
Krauss et al. (1995, Exp. 1 and 2)	Intercom	Describing synthesized sounds and graphic designs	Confederates ("not to interrupt the descriptions or to ask questions," "giving limited feedback" p. 541)	Gestures: *M* rate = 14.13 per min (*SD* = 7.15)	Gestures: *M* rate = 11.53 per min (*SD* = 7.80)	$p < .0001$	.82
Alibali et al. (2001)	Partition	Retelling a "Tweety and Sylvester" cartoon	Other participants ("listen carefully" for later recall but "not to ask questions" p. 174)	Representational gestures: *M* rate = 14.82 per min (*SE* = 1.72)	Representational gestures: *M* rate = 8.37 per min (*SE* = 1.18)	Representational gestures: $p < .001$	.56
				Beats: *M* rate = 7.42 per min (*SE* = 1.27)	Beats: *M* rate = 5.91 per min (*SE* = 1.08)	Beats: p = n.s.	.80
Emmorey and Casey (2001)	Partition	Directing addressee where to place blocks in a spatial puzzle	The experimenter ("said very little"and "only occasionally asked for clarification" p. 38)	Gestures: *M* frequency = 45[b]	Gestures: *M* frequency = 18[b]	Difference not tested	.40

[a] We have not included two other experiments that manipulated visibility: Mahl (1961) reported a visibility effect with no quantitative documentation, and Gullberg's (2006) experiment was designed for other purposes and included only gestures referring to people.
[b] No *SD*s reported.
[c] Differences between pairs of conditions not tested statistically.
[d] Proportion derived from the rate of gesturing in the non-visible condition divided by the rate in the visible condition.

Table 2
Ten unique features of spontaneous face-to-face conversations

Feature	Description	Medium		
		Face-to-face	Telephone	Tape recorder
1. Co-presence	Both participants are in the same physical environment	+	−	−
2. Visibility	They can see each other	+	−	−
3. Audibility	They can hear each other	+	+	−
4. Instantaneity	They see and hear each other with no perceptible delay	+	+	−
5. Evanescence	The medium does not preserve their signals, which fade rapidly	+	+	−
6. Recordlessness	Their actions leave no record or artifact	+	+	−
7. Simultaneity	Both participants can produce and receive at once and simultaneously	+	+	−
8. Extemporaneity	They formulate and carry out their actions spontaneously, in real time	+	+	+
9. Self-determination	Each participant determines his or her own actions (vs. scripted)	+	+	+
10. Self-expression	The participants engage in actions as themselves (vs. roles)	+	+	+

Note. Adapted from Clark (1996, pp. 9–10).

and intercom conditions were actual test trials. Second, the tape-recorder participants were to keep the experimenter in mind; we will discuss the effect of an implicit audience below. Finally, there was no statistical comparison of the intercom to tape recorder condition, which differed in dialogue but not visibility.

In an experiment on listeners' conversational (nonemotional) facial displays, Chovil (1989, 1991) used more comparable conditions, which separated dialogue and visibility by having addressees listen to a close-call story face to face, on the telephone, through a partition, or from a tape recording. Listeners used significantly fewer facial displays in the tape-recorder condition than in any of the three dialogue conditions, even those in which the speaker in the dialogue was not visible (i.e., the telephone and partition conditions). In the present experiment, we separated visibility and dialogue by creating three conditions: (i) two participants talking face to face (dialogue/visibility); (ii) two participants talking on a telephone (dialogue/no visibility); (iii) one participant talking to a tape recorder with no addressee at all (monologue/no visibility). We proposed that the dialogue variable itself would elicit some speaker gestures and therefore predicted that the two dialogues (face-to-face and telephone) would have a higher rate of gesturing than the monologue condition.

This design required a more precise definition of what we mean by dialogue vs. monologue than is readily available in the literature. Our definition of dialogue follows Clark's, (summarized in Table 2), and we note that previous studies usually did not meet all of these criteria. That is, with two exceptions (Bavelas et al., 1992, Exp. 2; Rimé, 1982), the addressee was the experimenter or a confederate who was acting under instructions that restricted his or her behavior (Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss et al., 1995), or the addressee was another participant with similar constraints (Alibali et al., 2001). Arguably, such studies do not fit the criteria of extemporaneity, self-determination, and self-expression because the addressees were not acting spontaneously, not determining their own actions, nor (in the case of confederates) engaging in actions as themselves. In the present experiment, both speaker and addressee were participants, and both could interact freely and naturally within their experimental task.

Creating a monologue condition turned out to be more problematic, because there is no clear definition of what constitutes a monologue. Most approaches used public speaking as the prototype; for example, Garrod and Pickering (2004, p. 8) described monologue as "preparing and listening to speeches." A more specific criterion seems to be whether the listener responds or not; for example, "*monologue* [is] speech or writing by a single person, as in a lecture or commentary; opposed to *dialogue*, where two people are participants in the interaction (Crystal, 2001, pp. 220–221). Clark (1996, p. 4) used similar examples but was more explicit about the audience's passivity in a monologue: "one person speaks with little or no opportunity for interruption or turns by members of the audience."

However, speeches differ from face-to-face dialogue in many more ways than the responsiveness of the audience or addressee, and the criterion of a nonresponsive audience does not transfer easily to a smaller setting. To illustrate the problems, we will consider here a succession of possible manipulations: First, if the essential nature of a dialogue is a responsive addressee, then a monologue could be created by an addressee who is nonresponsive (face to face or on the phone). However, there is experimental evidence (Beattie & Aboudan, 1994; Bavelas, Coates, & Johnson, 2000) that a physi-

cally present but nonresponsive or minimally responsive addressee disrupts many aspects of the speaker's communication. Presumably these effects occur because an unresponsive addressee is unnatural or inexplicable for the speaker, and it would be difficult to disentangle any monologue effects from this confound.

To escape such confounds, a second possibility would be to leave the speaker physically alone, with no addressee. This would be, in Crystal's terms, a *soliloquy*, "a literary monologue uttered by a speaker who thinks no one else is present" (2001, p. 221). However, speakers can easily speak to the camera (e.g., Bavelas et al., 1992, Exp.1) or to an implicit audience. Bavelas et al. (2002) asked participants who were alone in the room to pretend they were in a TV game show and to describe pictures to a partner who would either see or hear their description. These speakers were highly sensitive to the experimentally manipulated viewing conditions of their non-existent partner and changed their gestures accordingly. That is, when they were to imagine that the partner would see their videotape (vs. only hear their audiotape), they gestured at a higher rate and used more gestures that were not redundant with their speech. Our pilot work for the present experiment suggested that speakers who were talking to an answering machine usually spontaneously visualized an addressee—and gestured a great deal. Finally, experiments on facial displays have also shown significant increases for implicit audiences (Fridlund, 1991; Fridlund et al., 1990).

This leads to the third possibility, which is to leave the speakers physically alone but to also minimize the possibility that they would talk to an implicit audience. However, the latter cannot be achieved by experimental instruction (e.g., "Don't think about talking to someone") because there is substantial evidence that telling a person *not* to think about something is apt to backfire and increase the "forbidden" thought (Wegner, Schneider, Carter, & White, 1987; Wenzlaff & Wegner, 2000). We therefore chose a less obtrusive approach, which was to use task instructions that focused on the speaker's individual performance and also to give them a microphone to speak into, in order to deflect their focus from the camera. During debriefing, we sought to ascertain whether they had been thinking of an addressee and dropped them from the sample if they did; see Participants, below. This was a conservative method that might not have entirely eliminated all implicit audiences and would therefore have worked against our hypothesis by making the monologue condition somewhat more dialogic. Before leaving this issue, it is worth noting that all of the above considerations bring to our attention what authors from Mead (1934) to Linell (2005) have proposed, namely, that speaking is an intrinsically social activity. If so, it may be that we can only ever approximate a pure monologue.

Visibility and gestures as a communicative resource

A second limitation of the previous literature is that the physical variable of visibility was often merely descriptive rather than conceptual or theoretical. The implied explanation for its effect is that because unseen gestures would not be useful to the addressee, speakers would withhold them, but this is a negative explanation. It is also limited to predicting effects on frequency or rate measures, and we agree with Krauss, Chen, and Gottesman (2000,, p. 261) that why different gestures take the particular physical form they do is one of the most important yet largely unaddressed questions in gesture research. As shown in Table 1, only Alibali et al. (2001) and Bavelas et al. (1992, Exp. 2) have tested theories that would predict effects of visibility on gesture type or function. We will invoke a more explicit, positive, and general communicative model of gestures, one that makes a number of precise tests possible and will examine the effects of visibility on several aspects of the form or function of the gestures as well as on their relationship to words.

Broadly stated, we propose that visibility is one aspect of the speaker's *communicative context* and that speakers adapt their communicative choices to the parameters of their particular communicative context. Even holding constant *what* they are going to convey, there may be different situational resources or constraints that determine *how* they can do so. Some of these situational parameters are social, for example, whether there is an addressee, as noted above, or whether the addressee shares common ground with the speaker (Gerwing & Bavelas, 2004). Other parameters are physical, such as whether the addressee can see the speaker. In the latter case, the speaker must adapt to the resources or constraints of the physical medium. Kendon (1987) proposed 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* as well as the impact a gestural formulation may have on a recipient as compared to a verbal formulation. (p. 90, emphasis added)

In this model, gestural and verbal formulations constitute a flexible system, which can shift roles as the transmission conditions change (see also Bavelas & Chovil, 2000, 2006; de Ruiter, 2006). Several experiments have demonstrated specific shifts between verbal and gestural representations as adaptations to the physical parameters of the communicative situation (e.g., Bangerter, 2004; Bavelas et al., 2002; Clark & Krych, 2004; Emmorey & Casey, 2001; Graham & Heywood, 1975; Özyürek, 2002).

Visibility is one physical parameter that makes a number of communicative resources, such as gestures, available to the speaker. When the addressee will see the speaker's gestures, they are available as a means of conveying information to the addressee, and there should be evidence in their form and in their relationship to the words that speakers are using them communicatively. As a corollary, when the addressee will not see the gestures, they should be less communicative in form, and speakers may rely more on words. For example, experiments by Bavelas et al. (2002) and Gullberg (2006) have shown that visibility can produce qualitative changes that make the gestures more communicative in the sense of being less redundant with words or clearer in form. We therefore predicted that, in face-to-face dialogue, when the gestures would be visible, they would not only occur at a higher rate (as shown in previous studies), but they would also be larger, less redundant with speech (i.e., more likely to convey independent information), more likely to be marked by verbal deixis, and more likely to be oriented to the addressee than to the object the speaker is describing. (It is important to point out that we were not gathering evidence about whether addressees actually use the information in the speaker's gestures; that would require a different design than the one reported here.)

Our focus on the precise communicative context reveals that not all conditions that lack visibility are the same and that some are confounded by other variables, which led us to make several changes from previous designs. Perhaps most notably, none of the previous research designs actually used a telephone; instead, they manipulated visibility by using an open intercom between participants in different rooms (Cohen, 1977; Cohen & Harrison, 1973; Krauss et al., 1995) or a partition between participants in the same room (Alibali et al., 2001; Bavelas et al., 1992, Exp. 2; Emmorey & Casey, 2001; Rimé, 1982). We chose to use a real telephone, for several reasons. First, previous researchers obviously chose these alternative formats in order to leave the speaker's hands free in both conditions, thereby avoiding an artifactual explanation for any decrease in the rate of gesturing. Their results have shown that holding a telephone is *not* responsible for the reduced rate of gesturing, because the reduction occurs even when both hands are free. Given that these designs have addressed (and eliminated) that potential internal validity problem, we chose to move in the direction of external validity.

A second reason for using a telephone was to avoid other potential internal validity problems such as differences in familiarity. Talking on the telephone is not identical to talking through an intercom or partition because a telephone is a much more familiar communicative context (as is face-to-face dialogue). Most individuals in the university samples that psychologists study may talk on the phone many times a day but rarely talk through an intercom or a partition. Therefore, unlike the telephone, both intercom and partition differ from face-to-face dialogue in familiarity as well as visibility. There is also evidence for a perceived social difference between telephone and partition conversations. Chovil (1989, 1991) found that a university sample ranked telephone and partition conversations differently on *sociality*, defined as "how close people would feel in the situation and how easily the people would find it to converse with each other" (p. 149). The average sociality ranking was highest for face-to-face dialogue (2.81), followed by telephone dialogue (1.81), then dialogue through a partition (1.28), and least for a tape-recorder monologue (.14). Moreover, these independent sociality ratings accurately predicted the rate of listeners' communicative facial displays, which was significantly lower through a partition (even with the speaker present in the same room) than in a telephone dialogue with participants in different rooms. If perceived sociality also affects the rate of conversational gestures, then the previous studies with a partition had a potential confound that could account, at least in part, for the lower rate of gesturing.

The main concern about using a real phone is that having only one hand free might reduce the rate or qualities of gesturing, independently of any visibility effect. Our pilot work suggested that holding a telephone did not reducing gesturing, and we took several steps in the present experiment to ensure that there would be no effect on the dependent variables: Our experience in analyzing gestures has shown that, while speakers often make symmetrical gestures that involve both hands, they rarely make simultaneous *different* gestures with different hands. Our stimulus was highly symmetrical (see Fig. 1), and, in all of our operational definitions, we ensured that gesturing the same feature with one hand versus two hands did not affect the measure; see Analysis, below. That is, having only one hand free could not, in itself, affect our quantitative (rate) or qualitative (e.g., size) measures of gestures.

In any case, the results of previous studies are available as a comparison that would reveal whether the telephone itself caused a further reduction in the rate of gesturing, over and above the lack of visibility. As shown in Table 1, the previous visibility studies varied widely in the absolute frequency or rate of gesturing that they found, probably due to task and other differences. However, for each study, the frequency or rate of gesturing in the not-visible (partition or intercom) condition was a surprisingly constant *proportion* of the face-to-face condition. In spite of a wide range of definitions and methods of analysis, these proportions clustered around .50 and did not go below .39. Therefore, if physically holding the telephone does indeed further suppress the rate of gesturing, we would find a proportion below .39; that is, our telephone condition would have an even

Fig. 1. The stimulus picture (from Blum, 1982, p. 14).

lower rate relative to the face-to-face condition than in the previous, hands-free studies. On the other hand, if we were to find a proportion in the same range as previous studies, it would then be hard to suggest that holding a telephone was a confound producing an artifactual reduction in the rate of gesturing.

A corollary design decision was what to do in the case of the tape-recorder condition. A hand-held microphone was a good choice for two reasons. As noted above, speaking into a microphone was an important part of the creation of a monologue. Also, because the most important comparison was between the telephone and tape-recorder conditions, which differed on the new variable of dialogue versus monologue, these should be otherwise as alike as possible. Therefore, the speakers in the tape-recorder condition used a hand-held microphone, so that they too would have one hand occupied.

Finally, the speakers' communicative conditions include not only their physical or social medium but what they are talking about, for example, whether the assigned topic was conducive to gesturing or not. As shown in Table 1, previous studies used stimuli that ranged widely in gestural encodability from giving spatial

directions or describing highly visual material (Alibali et al., 2001; Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss et al., 1995, Exp.1, design stimulus) to close-call stories, which often elicit gesturing (Bavelas et al., 1992, Exp.2) to abstract topics such as giving opinions on movies (Rimé, 1982) to completely nonvisual material (Krauss et al., 1995, Exps. 1 & 3, sound and taste stimuli). Three previous studies have shown a significant effect of stimulus characteristics on gesture rates (Bavelas et al., 2002; Cohen, 1977; Krauss et al., 1995). Therefore, to avoid curtailment of range that would obscure experimental effects, we used a visual stimulus that could elicit a high rate of gestures.

Design, measures, and hypotheses

The participants described a complex picture of an 18th century dress (Fig. 1) in one of three conditions: face to face with another participant who could not see the picture; on the telephone with another participant who could not see the picture; or to a tape recorder with no addressee. In the first two conditions, the task was explicitly dialogic: the speakers needed to describe the picture well enough that their addressee could later pick it out of a group of similar pictures; the addressee could speak freely and ask questions. In the third condition, the focus was as monologic as possible, with the emphasis on the quality of individual description the speaker could record.

We assessed the rate of *all gestures* and the rate of gestures specifically describing the picture (*topic gestures*); the latter excluded gestures that served other functions. If visibility were the only controlling variable, then the rate of gesturing in the telephone and tape-recorder conditions should be significantly different from the face-to-face dialogue condition *and the same as each other*. If, as we propose, dialogue also has an additional, independent effect on the speaker's production of gestures, then the rate of gesturing in the telephone condition should be higher than in tape recorder condition; that is, dialogue should produce more gestures than monologue even when visibility is the same.

Because we propose that adaptation to communicative context is the basis of a visibility effect, we also predicted differences in the communicative qualities of the gestures, namely, their form and relationship to words. Gestures in the telephone and tape-recorder conditions should be different from face-to-face gestures and similar to each other. That is, although speakers on the telephone would gesture at a higher rate because it is a dialogue, they would not use their gestures to communicate to an addressee. Therefore, the following aspects of gestures should show an effect of visibility but not of dialogue. Note that although most of these measures of a gesture's form and its relationship to accompanying words were initially nominal, we converted all of them to parametric variables by using each speaker's rate, average, or proportion (see Analysis section for full details).

The first measure of form was the average *size* of the speaker's gestures, ranging from the scale of the picture (which was on a laminated 8.5 in. × 11 in. sheet) up to life-size, that is, on the scale of the speaker's own body. Gestures that were life-size would take more effort but would also be an excellent communicative resource if the addressee could see them in relation to the speaker's body, whereas such gestures would not be useful in the telephone or tape-recorder conditions.

The other two aspects of form focused on how the speaker oriented his or her gestures. A gesture could be oriented to the picture that the speaker was describing, for example, pointing at the picture or even touching or tracing features directly on the picture. These *picture-oriented gestures* might be of assistance to the speaker but would be of little help to the addressee, even in the face-to-face condition, because the addressee could not see what the speaker was pointing at in the picture and often could not see the gesture itself because it was behind the raised picture. Indeed, the visible presence of an addressee who needed the information might socially inhibit such self-focused actions. Other researchers have noted differences between gestures that seem directed towards the addressee and ones that are either explicitly speaker-directed (e.g., hidden from the addressee; Melinger & Levelt, 2004) or implicitly speaker-directed (e.g., less available to the addressee; Furuyama, 2000).

A contrasting form could occur when the speaker oriented a gesture directly at the addressee (i.e., in the ventral space between them and either pointing or moving the hand toward the addressee). These *interactive gestures* (Bavelas, Chovil, Coates, & Roe, 1995; Bavelas et al., 1992) do not convey substantive content (such as information about the picture) but instead support the process of interacting in face-to-face dialogue (e.g., marking information as given or new, or requesting evidence of understanding). Three previous experiments have shown that interactive gestures were highly sensitive to communicative context: Their rate was significantly higher in face-to-face dialogue than when the speaker was alone (Bavelas et al., 1992, Exp. 1). The rate was also significantly higher when two participants were in face-to-face dialogue than when they spoke face-to-face but in alternating monologues (Bavelas et al., 1995, Exp.1). However, the addressee also had to be able to see these gestures, because they decreased significantly when a dialogue occurred through a partition rather than face to face (Bavelas et al., 1992, Exp. 2). Therefore, we would predict an overriding visibility effect, with the rate of gesturing being higher in face-to-face dialogue than in either the telephone or tape-recorder conditions, that is, when the addressee would not see the interactive gestures.

Finally, we examined two aspects of the relationship between a gesture and the concurrent words. The first was the use of verbal deictic expressions that referred to the gesture, such as *like this* or *down here*. Several researchers (Bangerter, 2004; Bavelas et al., 2002; Clark & Krych, 2004; Emmorey & Casey, 2001) have analyzed speakers' use of such deictics and found a visibility effect: verbal deixis was less frequent when the addressee could not see the gesture to which the deictic referred, which we would expect to replicate here. A second strong evidence that transmission conditions affect the use of gesture is whether the gesture is redundant with the words or, instead, conveys unique information about the referent. Slama-Cazacu (1976), Kendon (1987), Clark (1996), Bavelas and Chovil (2000, 2006) and de Ruiter (2006) have all proposed that gestures and words are an integrated system. Emmorey and Casey (2001) and Melinger and Levelt (2004) found that gestures often expressed information that was omitted in speech, that is, gestures were not simply redundant with speech. Bavelas et al. (2002) found that speakers used significantly more non-redundant gestures even when they were talking to an imaginary recipient who would see their videotape rather than only hear their audiotape. Therefore, speakers in the face-to-face condition should use more nonredundant gestures than in the other two conditions.

Table 3 summarizes our predictions for the effect of visibility and dialogue on the seven dependent variables.

Method

Participants

Sixty-one first-year psychology students signed up online and participated in return for one bonus mark (0.5%) towards their course grade. They knew they would be videotaped during the experiment (either alone or with another participant) and that they would control access to their video. We excluded and replaced a total of 11 participants (three dyads and five individuals) from analysis. In two of these dyads and one of the tape-recorder conditions, the speaker explicitly reported trying not to gesture. In one dyad, the experimenter made a procedural error. Finally, four individuals in the tape-recorder condition were excluded because, during debriefing, they answered affirmatively when the experimenter asked whether they had imagined talking to someone (e.g., the experimenters). We also examined all audiotapes in this condition for language that implied an addressee (e.g., saying "Sorry" for an error or "You know"). The five individuals who reported an implicit audience were the only ones to use such language. (These five included the one who had already been dropped for trying not to gesture; he had also reported treating the situation "like a job interview.") The final N was therefore the planned 50 participants: 20 participants in the face-to-face condition (forming 10 dyads), 20 in the telephone condition (10 dyads), and 10 individuals in the tape-recorder condition. These 50 participants gave us 30 speakers to analyze, 10 in each condition.

We randomly assigned the order of the three conditions at the outset. An exception would occur when one participant did not arrive for a pre-assigned dyadic condition, in which case the individual who did come was re-assigned to the tape-recorder condition. To avoid any possible temporal effects, we replaced the missing dyadic condition as soon as possible. Because our sign-up procedure prevented participants from knowing whether they were going to be in a dyad or alone, participants in all conditions were drawn from the same population (i.e., individuals who expected to be either alone or in a dyad and would actually show up at the experiment). In the dyadic conditions, we randomly assigned the roles of speaker and addressee.

Table 3
Predictions for different effects of visibility and dialogue on gesture rate, form, and relationship to words

Dependent variable	Condition		
	Face-to-face	Telephone	Tape recorder
Rate			
All gestures	Highest	High	Lowest
Topic gestures	Highest	High	Lowest
Form of gestures			
Size	Largest (life-sized)	Small (picture-sized)	Small (picture-sized)
Picture-oriented	Lowest	Higher	Higher
Interactive	Highest	Lower	Lower
Gestures' relation to words			
Proportion with deictic	Highest	Lower	Lower
Redundancy	Lowest	Higher	Higher

Equipment

Our Human Interaction Laboratory was equipped with four remotely controlled, tightly synchronized Panasonic WD-D5000 color cameras and two special effects generators (a Panasonic WJ-5500B overlaid on a customized Panasonic four-camera system). We used three cameras, configured in split-screen; see Fig. 2. For the telephone condition, the speaker used an ordinary hand-held telephone, and we tapped the telephone audio directly into the video system so that both participants were audible on the videotape. For the tape-recorder condition, we used a portable tape recorder and a hand-held microphone; however, the analysis used the synchronized audio from their videorecording. We digitized the analog video into AVI format using Broadway ProDVD (www.b-way.com) and analyzed the digitized data on an 18-in. ViewSonic GS790 color monitor, using Broadway.

Materials

The stimulus was a black-and-white picture (approximately 8 1/2 × 11 in.) of a very elaborate eighteenth-century dress (Blum, 1982, p. 14; see Fig. 1), which was laminated to a cardboard sheet and presented in a manila folder. The instructions for the task were printed on the outside of the folder. The speaker removed the picture and stood it in a clear plastic stand so that it would not be visible to the addressee in the face-to-face condition. There was an additional stimulus (a geometric maze), which we prepared in the same manner as the picture of the dress but which was not part of this experiment and was not analyzed. For the addressees' test, we created a large placard with four digitally edited versions of the same dress, one of which was identical to the original picture.

Procedure

In all conditions, there were two experimenters, one to conduct the experiment and another to handle the video equipment. Before recording began, the participants consented in writing to being videotaped. The experimenter then gave instructions at the outset for all of their main tasks, which were getting acquainted (or, in the tape-recorder condition, describing oneself) and then describing two pictures. (After these tasks,

Fig. 2. Three-way split-screen layout (face-to-face condition). The top of the screen is a face-on view of the speaker (on the left) and addressee (in the circle). The bottom of the screen is a side view of both participants. Note: In the telephone and tape-recorder conditions, the circle was empty.

the participants changed roles and received instructions for some unrelated pilot tasks.) The instructions for the main tasks were given only once at the beginning for all conditions, primarily in order to make the tape recorder condition as asocial as possible. That is, repeated interactions with the experimenter might increase the sociality of that condition. To prevent problems with remembering task instructions, there was a written copy of the instructions for the later tasks on the outside of the stimulus folder. At the conclusion of the experiment, the participants received an explanation of the study, asked any questions they had, and viewed the videotape of their participation. We then asked them to indicate, in writing, various levels of permission to view the data (e.g., permission to view for analysis only, to show to professional audiences, to reproduce as still photos in journal articles, etc.).

Face-to-face condition

After the participants had consented to being videotaped, they received instructions for their three tasks. First, they were to have a brief getting-acquainted conversation, spending approximately three minutes discussing their academic interests, hobbies, hometown, or whatever they chose to talk about. Then the speaker was to describe two different pictures, in counter-balanced order. The instructions emphasized that this was a dialogue. The experimenter asked the assigned speaker to take "the picture of an 18th century dress" out of the folder, place it in the stand so that the other person could not see it, and then describe it "to the other person, in the clearest and most detailed way that you can." The experimenter emphasized that they could "talk and ask questions whenever you need to" and that, when the addressee had a good idea of what the dress looked like, he or she would then have to choose this dress from pictures of four dresses. When the speaker and addressee felt they were finished talking about the dress, they announced that they were done. Then the experimenter re-entered, presented the addressee with the four dress options, and the addressee selected the dress that had been described. The choices were intended to be easy, and all addressees in this condition made the right choice. The speaker did not see the four options. The next task followed exactly the same procedure, but with a picture of a geometric maze. After these three tasks, the participants did two more tasks that were part of a pilot study.

Telephone condition

The procedure, tasks, and instructions were the same as the face-to-face condition except that, following the experimenter's instructions to both participants in person, the speaker stayed in the laboratory to be videotaped, while the experimenter took the addressee into a nearby office, connected them by telephone, and waited outside the office. When they were finished with getting acquainted and the speaker had described the dress, the addressee informed the experimenter, who presented the four options as above. One participant made the wrong choice, but we still told her that her choice was correct; again, the speaker did not see the four options. The participants then continued with the other (counterbalanced) picture. After this, the speaker and addressee changed places and did the pilot tasks on the telephone.

Tape-recorder condition

Participants in this condition arrived alone and spoke into an audio recorder with a hand-held microphone. To replace the getting-acquainted component of the dyadic conditions, we asked these participants to introduce themselves into the tape recorder (giving information about their academic interests, hobbies, etc., as in the other two conditions). They then took the picture of the dress (or maze) out of the folder, propped it up in the stand, and were to describe it "in the clearest and most detailed way" that they could. When they were finished, they repeated the procedure with the other picture, as above. As described earlier, their debriefing also included questioning about whether they had been talking to an implicit audience. Although this procedure screened out the clearest cases, all speakers knew they were being videotaped, so there was an implicit though muted audience in all three conditions, in the sense that the cameras and whoever would see the tape were overhearers (Schober & Clark, 1989).

Analysis

Data preparation and reliability procedures

Two individuals prepared and checked a transcript of the words of both participants in each description of the dress. At least two independent analysts were responsible for all of the analyses described below, using formal definitions and detailed guidelines (available from the first author). Reliability was assessed at two levels: First, during analysis, we aimed for analysts to agree on *each gesture* they located, so their initial reliability for these nominal decisions was percentage agreement (e.g., all agreements on individual gestures, divided by all agreements plus all disagreements, aggregated across all gestures of all speakers). Any disagreements were resolved either by the pair of analysts or by the research group. Second, because the ultimate dependent variables used in statistical analyses were the rates, averages, or proportions of all gestures made by each speaker, we also assessed reliability for these parametric data using r (e.g., the correlation between the two analysts for the

rate of gesturing they found for each speaker). The difference between these two levels of reliability is that rates, averages, or proportions yield one number per speaker and do not require that analysts agree on particular gestures, as long as they arrive at similar aggregate numbers for a speaker. Indeed, two analysts need only agree about the relative (and not absolute) figures per speaker in order to correlate highly. Our initial percentage agreement is therefore a more demanding and sensitive assessment of the reliability of analytic decisions. As will be seen below, both levels of reliability were high for all measures.

Rate of gesturing

The traditional dependent variable is the number of gestures divided by speaking time, or rate of gestures per minute. An alternative measure, the rate of gestures per 100 words (of the speaker), adjusts for differences in speaking rate, and we will report both here. In any case, both rates start with counting the number of gestures each speaker made. We identified gestures using McNeill's definition of a gesture stroke, that is, by identifying the meaningful part of the gesture that was synchronized with the accompanying speech (McNeill, 1992, p. 83). Each analyst located each gesture that the speaker made while describing the dress, which required two successive decisions. First, the analysts needed to differentiate between meaningful movements (i.e., gestures) and non-meaningful ones (i.e., adaptors), thereby eliminating from further analysis any noncommunicative movements such as rubbing the arms, pushing back hair, etc. We did include movements related to the picture, such as pointing or tracing. Second, because gesture strokes often occurred in virtually continuous succession (i.e., without retraction to a resting position), each analyst had to decide whether any sequence of contiguous gestures was a unitary whole or separate strokes. Sometimes, as in the following example, gesture strokes were separated by brief, post-stroke holds; in these cases, their timing would separate them. For example, Fig. 3 shows the actions of one speaker as he described the shape of the skirt; underlining indicates where his gestures occurred in relation to his words:

"It just juts completely out [Addressee: Okay] and drapes down."
gesture 1 gesture 2

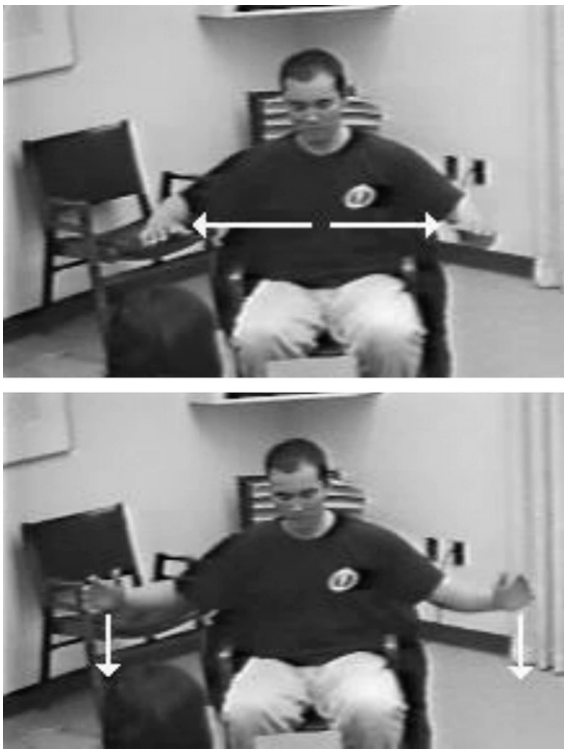

Fig. 3. Speaker describing the shape of the skirt (face-to-face condition). Arrows indicate the size and direction of the gestures. Top: Participant draws a line to indicate the top of the shape of the skirt. Bottom: Participant indicates the shape and location of the sides of the skirt.

As shown in the top picture, this speaker first used both hands to draw a line in front of himself, starting just below his chest and extending about 2 feet on either side (gesture 1). Then, after a pause of about 0.5 s, he drew lines going straight down from his previous ending point (gesture 2, bottom figure). The pause created two separate gestures. In other instances, when there was no pause, we relied on changes in meaning. For example, in Fig. 4, the same speaker described the trim material on the bodice of the dress as follows:

"[This material] that starts around the back of her neck, comes down,
gesture 1

and then right at her waist where the dress juts out ...."
gesture 2

As shown in the top and middle pictures, the speaker first used both hands to draw two lines from the back of his own neck down to his waist (gesture 1). Then, in the bottom picture and starting with the words "and then," he touched his waist several times, again with both hands (gesture 2). The first gesture depicted the outline of the material, whereas the second gesture indicated and emphasized the point where the dress began to jut out. Because these were two different meanings, we counted them as two gesture strokes even though there was no pause between them.

As noted above, we assessed reliability at the gesture level, that is, whether the two analysts agreed or disagreed on each gesture, expressed as a percentage of

Fig. 4. Speaker describing the material trimming the bodice of the dress (face-to-face condition). Arrows indicate the size and direction of the gestures. Top: Participant starts first gesture at the back of his neck. Middle: Participant traces the shape of the trim from his neck to his waist. Bottom: Participant touches his waist several times to indicate where the trim begins to "jut" out.

agreements. For the first decision, which was to distinguish gestures from other hand movements, they agreed 91.6% over all gestures by all speakers. For the second decision, which was to separate sequences of continuous gesturing (or not), they agreed 78.2%, over all gestures by all speakers. It was the total number of gestures per speaker that determined the rate measures (per minute or per word), and the reliability for the total number of gestures that each analyst found per speaker was $r = .98$, which is therefore the reliability of the actual dependent variable. All disagreements were resolved before proceeding on to the analysis of qualities of gestures.

Bias check

Because the analysts could not help seeing the experimental condition on the video (i.e., the presence of another person, a telephone, or a tape recorder), there was a potential for concern about bias in the direction of our hypothesis. Judgments about dividing gestures could be particularly vulnerable if the analysts tended to make more divisions (resulting in a higher rate of gesturing) in one condition than another. Note that the above high reliabilities at the level of each gesture argue against this possibility, because an overall tendency to divide (or not) would not produce high independent agreement on exactly which individual gestures to divide and how much. However, we also checked for potential bias by hiring three undergraduate honours students who had no knowledge of the project or hypotheses either before or during their period of work. Moreover, each student saw only one of the experimental conditions and had no knowledge that the other conditions existed. Other precautions to keep them blind to condition included relabeling all records in the research office and computers from "phone study" to "dress study"; not talking about the study when they were present; and instructing them not to talk to each other until all of them had finished. These students applied a much simpler analysis, which only required them to differentiate between intervals where there were *meaningful hand movements* (i.e., gestures) and intervals where there were *no meaningful hand movements* (i.e., an adaptor or no movement at all). Using frame-by-frame analysis and the timer on the software, they recorded the duration of each interval where there was meaningful hand movement and also the duration of any intervals of no meaningful hand movements. In order to preclude their having to make any other judgments, the preparatory phase of a gesture was included as meaningful hand movement. We converted the summed durations to the *proportion of meaningful hand movement per speaker*. The correlation between this measure and our measure of each speaker's rate of gesturing was .89, which was high enough to exclude the possibility of bias in our measure, especially given the differences in operational definitions.

Gesture function

Previous research (Bavelas et al., 1992, 1995) had demonstrated empirically a functional distinction between topic and interactive gestures, which we applied to these data. *Topic* gestures depict some aspect of the

current topic of conversation, in this case, the dress (e.g., Figs. 3–5). They include what many researchers call representational or iconic gestures. *Interactive* gestures refer directly (in form and meaning) to the addressee or to the conversational interaction. For example, the speaker might gesture toward the addressee to mark common ground or to seek evidence that the addressee understands. In these data, there was also a third function, *picture-oriented* gestures. The speakers sometimes pointed directly at or traced a feature of the picture they were describing, either touching the picture or gesturing very close to it. Note that because the addressee, even if present, could not see the picture, these movements were arguably for the benefit of the speaker.

Two analysts independently analyzed each gesture for which of the three functions it was serving. Their agreement for distinguishing among the three functions, gesture by gesture, was 93.4% over all speakers and gestures. We also calculated Cohen's Kappa for this qualitative distinction, and the value (.70) would be characterized as "substantial" agreement (Landis & Koch, 1977, p. 165). The actual dependent variables for the three functions were rates per speaker for each function, which was determined by the total number of each kind of gesture function per speaker. These reliabilities were $r = .996$ (number of topic gestures), $r = .87$ (number of interactive gestures), and $r = .93$ (number of picture-oriented movements). This analysis yielded our second pair of rate measures (rate of topic gestures per minute or per 100 words) as well as two pairs of form measures (rate of picture-oriented and of interactive gestures per minute or per 100 words).

Fig. 5. Speaker describing the top of the skirt (telephone condition). Arrows indicate the size and direction of the gestures.

Size of gestures

The gestures that speakers used to describe the dress ranged from the scale of the picture up to life-sized (e.g., when speakers described the dress as if it were on their own body). Two analysts rated the size of each gesture describing the outline of the skirt. We chose descriptions of the skirt both because it was the largest feature and therefore most sensitive to size differences and because almost all participants included the skirt in their description. The ratings were from 1 to 5, where 1 = the gesture was on the scale of the picture (e.g., Fig. 5) to 5 = the zgesture was on the scale of the speaker's body (e.g., Figs. 3 and 4). Note that this was a measure of the relationship between the gesture and either the picture or the participant's body and not a measure of the absolute size of the gesture. For example, a speaker could depict the skirt with a two-handed gesture showing the width on both sides of his or her body, while another speaker could use a one-handed gesture indicating the full width of only one side of the skirt; both would be treated as life-size (5) because both were scaled to match the size of the skirt as though it were on the participant's body. The reliability across all gestures and speakers was $r = .92$. The reliability for the average size of gesture per speaker, which determined the rate measures, was $r = .97$.

Gestures' relationship with words

Deictic expressions referring to a gesture

An additional measure of evidence regarding the speaker's use of gestures to communicate was the speaker's explicit verbal reference to his or her own gesture (Bangerter, 2004; Bavelas et al., 2002; Clark & Krych, 2004; Emmorey & Casey, 2001). *Deictic* (or indexical) expressions refer to and depend upon something in the immediate context, in this case, a gesture. Examples included "her waist is about *here*," "the bow is down *there*," "it goes like *this*," or "it's *that* long," accompanied by a gesture that depicted the location of "here" or "there" or the specific nature of "this" shape or "that" length. Two analysts examined the words accompanying each topic gesture for the use of a deictic expression referring to the gesture. The agreement across all gestures and speakers was 97.3%. The reliability for the aggregate dependent measure, which was the proportion of each speaker's topic gestures that were accompanied by a deictic expression, was $r = .98$.

Redundancy with words

We also examined the meaning of each gesture in relationship to the words it accompanied. A redundant gesture provided no additional information beyond the meaning that the words conveyed. A nonredundant gesture did convey information that was not in the words; the speaker was arguably relying on the gesture to carry

some of the information. To assess redundancy, we chose two features of the dress that the maximum number of participants included in their descriptions, the shape of the skirt and of the neckline.

The analysts examined each gesture that referred to these features and made a dichotomous decision whether or not there was additional meaningful information in the gesture that was not in the words and, if so, what that information was. For example, the gestures in Figs. 3 and 4 were both non-redundant. In Fig. 3, the gesture showed *how far* the skirt "juts completely out" and its *shape* as it "comes down." In Fig. 4, the first gesture showed *how far* the material "comes down." The second gesture indicated *location*, which was redundant with "right at her waist," but additionally maintained the *distance* between the two bands of material. That is, the words plus gesture in this phrase also meant "the two bands are still the same distance apart" rather than, for example, coming together in a V like the neckline trim. The gesture in Fig. 5 was redundant. Describing the top of the skirt, the speaker said it goes

"like a meter to the left"

Her gesture was a brief, vague movement to the left. Her words, rather than the accompanying gesture, gave the distance, and her splayed fingers contributed no additional information, such as shape. Note that although the vagueness or ambiguity of a gesture could lead to its being deemed redundant, this ambiguity was not confounded with the size of the gesture. Our large, high-resolution, color monitor made all gestures clearly visible, and small gestures were often well-formed. The per-gesture agreement on redundancy over all speakers was 84.4% for the 167 skirt gestures and 83.1% for the 123 neckline gestures. None of the speakers in the tape-recorder condition gestured when describing the neckline, so for the neckline redundancy measure there were only data in the face-to-face and telephone conditions. The reliability for the dependent variable, which was the proportion of each speaker's gestures that were redundant with words, was $r = .92$ for the skirt and $r = .71$ for the neckline. The lower correlation for the neckline analysis was greatly affected by a single speaker who made only one gesture about the neckline, on which the analysts disagreed, which led to the maximum negative relationship. If we exclude this speaker as well as one speaker at the other extreme (where the analysts agreed on the single gesture and therefore introduced the maximum positive relationship), the correlation becomes .95 for the remaining neckline gestures.

Figurative language

Finally, while conducting the above analyses of words, we became aware of speakers' frequent use of verbal imagery to describe features of the dress. Other researchers have examined various relationships between gesture and figurative language (Corts, 2006; Corts & Pollio, 1999; Hadar & Krauss, 1999; Rimé, Schiaratura, Hupet, & Ghysselinckx, 1984), but none have examined the effects of dialogue or visibility. In a study of figurative language (but not gestures), Boerger (2005) found that figurative expressions occurred at a *higher* rate when the participants could not see each other. That is, the rates were significantly lower in face-to-face dialogue than in three other conditions (which included an intercom, email, or being able to see only the other person's eyes). Here we were interested in whether verbal images might sometimes be an alternative to gesture and so developed a measure of figurative speech, including metaphor, simile, or analogy. To be considered figurative speech, the speaker's word or phrase had to describe a feature of the dress by referring explicitly to something that was not on the dress. For example, speakers described the design on the front of the skirt as "worms," "two arches," or "snakes." Our definition was conservative in the sense that it excluded names for abstract geometric shapes such as square, triangle, cube, or cylinder because these were arguably correct descriptions of a feature of the dress. We did include shape descriptions that invoked comparisons, such as "heart-shaped," "V-shaped," or "like a W."

The analysis of figurative speech used only a transcript of the speaker's words, so the analysts were blind to experimental condition. These two analysts had also never seen the videos. They worked together for eight speakers, then independently for the remaining 22. For the independent analysis, their agreement on which expressions should be considered figurative was 90%. Because some speakers used the same figurative expression more than once, we also checked the same analysts' agreement on frequency for 14 speakers, that is, how many figurative expressions each speaker used. They agreed on 87.5% of the figurative expressions they located, resulting in a correlation of $r = .98$ for the total number of figurative expressions per speaker. The latter number led to the dependent variable, which was the rate of figurative language per 100 words for each speaker.

Results

Qualitative description of data

Because the reader cannot view our video data here, we would like to preface our statistical findings with a qualitative description of the characteristics of gestures in the three conditions, which appeared to differ in some striking ways. In the face-to-face condition, both male and female speakers typically used their gestures to place

the dress around their own body. For instance, most speakers drew the unusual size and shape of the skirt by reaching outwards from their waist until their arms were fully extended, often with their hands shaping the two corners. They also drew the neckline of the dress on their own chest and the layers of ruffles on the sleeves down their own arms. If they described the closed fan she held in her hand, they would frequently position their own arm and hand like the figure in the picture and pretend to hold a fan. They often drew the hemline and its intricate designs just above the floor, leaning down and reaching well outside the typical gesture space. Some participants got out of their chair and stood up (like the woman in the stimulus picture) to draw the dress on and around their own body. In general, gestures in this condition were large, well-formed, and often continuous. They usually had a clear relationship to the words they accompanied, making analysis of their meaning relatively straightforward. In short, these gestures were maximally communicative to us as overhearers and presumably to the addressees.

The telephone gestures looked quite different. Often, these speakers described the dress while leaning forward in the chair and gazing intently at the picture (which rarely occurred in the face-to-face condition, although looking at the picture was equally necessary in all conditions). Their gestures were small movements of the hand, more often on the same scale as the picture itself rather than the speaker's body. These gestures were also brief; rather than occurring over long stretches of related words, they might occur with only one or two words, although they were still timed precisely with those few co-occurring words. From the analysts' point of view, these gestures were harder to describe, and their meanings were harder to derive. In this condition, the words seemed more descriptive and transparent while the gestures appeared to be contributing less.

Finally, the tape recorder gestures were tiny and strange. In this condition, participants also traced the picture with one hand. Or they sat back in their chair and made small motions with their free hand. Their motions were usually timed with the words, but sometimes finding a meaningful relationship to the words involved laborious decision-making for the analysts, because the gestures did not seem to be depicting clear referents. For instance, at one point, a speaker in this condition made two distinct, tiny movements with her little finger, both tightly synchronized with her words, but with no decodable meaning; we still called it a gesture.

Statistical findings

The test of our primary hypothesis on the independent effects of visibility and dialogue was by linear regression, using the additional sum of squares principle (e.g., Weisberg, 1985, pp. 37–41). Thus, we first tested the effect of visibility (face-to-face vs. telephone and tape-recorder) on a dependent measure such as the rate of topic gestures. Then we tested the effect of dialogue (face-to-face and telephone vs. tape-recorder) on the variability in the dependent variable that had *not* been explained by visibility. To paraphrase Weisberg using our own variables, "The main idea in adding [dialogue] is to explain the part of [rate of topic gestures] that has not already been explained by [visibility]" (p. 38).[1] We also used pairwise comparisons to elucidate specific effects.

We have grouped the results into four kinds of dependent variables: length of descriptions (Table 4), rates of gesturing (Table 5), form of gestures (Table 6), and relationships to words (Table 7). Each table includes descriptive statistics, simple main effects, pairwise comparisons of experimental conditions, and the regression analysis that tested for independent visibility and dialogue effects. Our pairwise comparisons between condition means were usually confidence intervals around the difference between means. When these confidence intervals were for variables with unequal numbers of participants, we used harmonic means. However, because confidence intervals around differences between means are particularly vulnerable to violations of the assumption of homogeneity of variance (Masson & Loftus, 2003), when the dependent variable had heterogeneous variance (as shown by Levene's test for equality of variances), the appropriate comparable test was Dunnett's T3.

Length of descriptions

It is of interest to note in Table 4 that the two dialogue conditions were virtually the same average length (face-to-face $M = 260$ s, $SD = 90.13$; telephone $M = 257.4$ s, $SD = 91.80$) and longer than the tape-recorder condition ($M = 111.6$ s, $SD = 39.86$). A one-way ANOVA indicated a significant main effect ($F_{(2,27)} = 11.93$; $MSE = 6046.77$; $p < .001$). Dunnett's T3 post hocs indicated that both of the dialogue conditions were significantly longer than the tape-recorder condition and that they were not significantly different from each other. The regression analysis revealed a marginal visibility effect and a strong dialogue effect. The dialogue effect is not simply due to the contributions of an addressee, because an analysis of the number of speaker's words shows a similar pattern; see Table 4. It seems that having an addressee stimulated the

[1] Note that we distinguish between independence and orthogonality. The design we required does not permit a test of whether visibility and dialogue are orthogonal. Our claim is that dialogue can account for variance over and above that which visibility can account for.

Table 4
Results of length measures

Dependent variable	Descriptive statistics: mean (SD)			One-way ANOVA		Regression	
	Face-to-face	Telephone	Tape recorder	Main effect	Pairwise comparisons[a]	Effect of visibility	Independent effect of dialogue
Description in seconds	260.00 (90.13) $n = 10$	257.40 (91.80) $n = 10$	111.60 (39.86) $n = 10$	$F_{(2,27)} = 11.93^{***}$ $MSE = 6046.77$	FF > TR;PH > TR	$F_{(1,28)} = 3.95$[b]	$F_{(1,27)} = 17.58^{***}$
Description in words	653.00 (242.49) $n = 10$	606.80 (228.79) $n = 10$	202.20 (95.66) $n = 10$	$F_{(2,27)}$=15.34*** $MSE = 40098.79$	FF > TR;PH > TR	$F_{(1,28)} = 6.06^{*}$	$F_{(1,27)} = 20.41^{***}$

[a] Pairwise comparisons tested with Dunnett's T3 because variance for these variables was non-homogenous.
[b] Marginally significant, $p = .057$.
* $p < .05$.
*** $p < .001$.

speakers to make fuller descriptions. Because all of the dependent variables were converted to rates, averages, or proportions, we eliminated any artifact of these differences in length. Also, as noted earlier, we calculated the rates of gesturing both per minute and per 100 words (of the speaker). When the results were the same for both measures, which was usually the case, we have reported the per-minute figures in the text.

Rate of gesturing

Our 30 speakers made a total of 1840 gestures. As we had found in previous experiments with descriptive tasks (Bavelas et al., 1992; Bavelas et al., 1995), the vast majority of these gestures were topic gestures, describing the dress itself. Therefore, the results for the rates of all gestures and the rates of topic gestures were virtually the same (see Table 5). Both of these dependent variables showed significant main effects; all gestures $F_{(2,27)} = 10.87$; $MSE = 67.96$; $p < .001$; topic gestures $F_{(2,27)} = 13.23$; $MSE = 58.55$; $p < .001$.

The regression analysis supported our primary hypothesis, with all four rate measures showing a significant dialogue effect in addition to and independently of the significant visibility effect. That is, while lack of visibility was reducing the rate of gesturing, dialogue was increasing the rate, even in the telephone condition. As a result, the means declined across the three conditions, with speakers in the face-to-face condition gesturing the most ($M = 21.80$ topic gestures per minute), followed by those interacting on the telephone ($M = 14.90$ topic gestures per minute) and then those talking into a tape recorder ($M = 4.32$ topic gestures per minute). The mean differences between the two dialogue conditions and the tape recorder condition both exceeded the 95% confidence interval of ±8.49. The rate of topic gestures per minute in the face-to-face condition was higher than the rate in the telephone condition, but the pairwise comparisons were not significantly different, which suggests that the visibility effect in the above regression analysis is primarily due to the tape-recorder condition. (All of the above tests showed the same results using the rate per 100 words; see Table 5.)

The lack of significant difference between face-to-face and telephone conditions may at first seem surprising, but it echoes a minority result in the previous literature. As shown in Table 1, the seven previous experiments all found more gesturing in their visible condition than their not-visible condition; this difference was significant for five of the experiments (Alibali et al., 2001; Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss et al., 1995) and not for two others (Bavelas et al., 1992, Exp. 2, result for topic gestures; Rimé, 1982). Although there are many reasons that a difference may not reach statistical significance, we are intrigued by a procedural difference between these two groups, namely, whether their conditions met all of the criteria

Table 5
Results of rate measures

Dependent variable	Descriptive statistics: mean (SD)			One-way ANOVA		Regression	
	Face-to-face	Telephone	Tape recorder	Main effect	Pairwise comparisons[a]	Effect of visibility	Independent effect of dialogue
All gestures per min	23.48(8.96) $n = 10$	17.60(9.80) $n = 10$	6.55(5.25) $n = 10$	$F_{(2,27)} = 10.87^{***}$ $MSE = 67.96$	FF > TR ± 9.15 PH > TR ± 9.15	$F_{(1,28)} = 9.93^{**}$	$F_{(1,27)} = 8.98^{**}$
All gestures per 100 words	15.02(4.00) $n = 10$	12.08 (5.80) $n = 10$	5.85(4.16) $n = 10$	$F_{(2,27)} = 9.82^{**}$ $MSE = 22.32$	FF > TR ± 5.24 PH > TR ± 5.24	$F_{(1,28)} = 8.59^{**}$	$F_{(1,27)} = 8.70^{**}$
Topic gestures per min	21.80(9.20) $n = 10$	14.90(8.38) $n = 10$	4.32(4.56) $n = 10$	$F_{(2,27)} = 13.23^{***}$ $MSE = 58.55$	FF > TR ± 8.49 PH > TR ± 8.49	$F_{(1,28)} = 12.96^{**}$	$F_{(1,27)} = 9.55^{**}$
Topic gestures per 100 words	13.92(4.40) $n = 10$	10.30(5.04) $n = 10$	3.77(3.64) $n = 10$	$F_{(2,27)} = 13.66^{***}$ MSE = 19.36	FF > TR ± 4.88 PH > TR ± 4.88	$F_{(1,28)} = 12.04^{**}$	$F_{(1,27)} = 10.98^{**}$

[a] Pairwise comparisons are reported as 95% confidence intervals (HSD) around the differences between means.
** $p < .01$.
*** $p < .001$.

Table 6
Results of form of gesture measures

Dependent variable	Descriptive statistics: mean (SD)			One-way ANOVA		Regression	
	Face-to-face	Telephone	Tape recorder	Main effect	Pairwise comparisons[a]	Effect of visibility	Independent effect of dialogue
Average size (skirt)	4.79(0.24) $n = 10$	1.67(0.60) $n = 8$	1.47(1.04) $n = 5$	$F_{(2,20)} = 78.35^{***}$ $MSE = 0.37$	FF > PH FF > TR	$F_{(1,21)} = 161.37^{***}$	$F_{(1,20)} = 0.35$
Picture-oriented gestures per min	0.48(0.93) $n = 10$	1.80(1.95) $n = 10$	1.84(3.01) $n = 10$	$F_{(2,27)} = 1.32$ $MSE = 4.57$	n.s.	$F_{(1,28)} = 2.74$	$F_{(1,27)} = 0.001$
Picture-oriented gestures per 100 words	0.32(0.63) $n = 10$	1.17(1.21) $n = 10$	1.71(3.15) $n = 10$	$F_{(2,27)} = 1.24$ $MSE = 3.94$	n.s.	$F_{(1,28)} = 2.16$	$F_{(1,27)} = 0.37$
Interactive gestures per min	1.21(0.74) $n = 10$	0.38(0.31) $n = 9$	0.39(0.66) $n = 10$	$F_{(2,26)} = 5.94^{**}$ $MSE = 0.37$	FF > PH FF > TR[b]	$F_{(1,27)} = 12.34^{**}$	$F_{(1,26)} = 0.001$
Interactive gestures per 100 words	0.77(0.42) $n = 10$	0.26(0.20) $n = 9$	0.36(0.62) $n = 10$	$F_{(2,26)} = 3.50^{*}$ $MSE = 0.21$	FF > PH	$F_{(1,27)} = 6.91^{*}$	$F_{(1,26)} = 0.27$

[a] Pairwise comparisons tested with Dunnett's T3 because variance for these variables was non-homogenous.
[b] Marginally significant: $p = .052$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

of a dialogue. In Rimé (1982), Bavelas et al. (1992, Exp. 2), and the present experiment, the speaker and addressee were both participants who could interact freely and spontaneously, when and as they wished. In contrast, the five experiments that found a significant effect of visibility were also the ones that constrained the addressee (who was usually the experimenter or a confederate) to a limited repertoire of responses. These highly constrained interactions were in fact closer to monologues.

The rate results provide several lines of statistical evidence that using one hand for the telephone was not an artifact that lowered the rate of gesturing here. First, speakers who were holding the telephone gestured at rates that were not significantly different from speakers in the face-to-face condition, who could use both hands. Second, we can compare the proportion of telephone to face-to-face rates in this study with the analogous proportions in earlier studies. The rate of topic gestures per minute in the telephone condition was .68 of the face-to-face condition, and the rate of all gestures per minute in the telephone condition was .75 of the face-to-face condition. These proportions are higher than in

Table 7
Results of gestures' relation to words measures

Dependent variable	Descriptive statistics: mean (SD)			One-way ANOVA		Regression	
	Face-to-face	Telephone	Tape recorder	Main effect	Pairwise comparisons[a]	Effect of visibility	Independent effect of dialogue
Proportion of topic gestures with a deictic expression	0.15 (0.08) $n = 10$	0.03 (0.03) $n = 10$	0.01 (0.02) $n = 8$	$F_{(2,25)} = 23.73^{***}$ $MSE = 0.002$	FF > PH FF > TR	$F_{(1,26)} = 47.43^{***}$	$F_{(1,25)} = 0.66$
Proportion of gestures redundant with words (skirt)	0.11 (0.10) $n = 10$	0.58 (0.28) $n = 8$	0.85 (0.34) $n = 5$	$F_{(2,20)} = 19.40^{***}$ $MSE = 0.05$	FF < PH ± 0.30 FF < TR ± 0.30	$F_{(1,21)} = 29.91^{***}$	$F_{(1,20)} = 4.26$[b]
Proportion of gestures redundant with words (neckline)	0.18 (0.12) $n = 8$	0.45 (0.37) $n = 9$	— $n = 0$	$t_{(9.72)} = 2.09^{*}$ (one-tailed)	n.a.	n.a.	n.a.
Figurative expressions per 100 words	0.57 (0.41) $n = 10$	1.11 (0.61) $n = 10$	0.37 (0.52) $n = 10$	$F_{(2,27)} = 5.36^{*}$ $MSE = 0.27$	PH > TR ± 0.58	$F_{(1,28)} = 0.53$	$F_{(1,27)} = 10.02^{**}$

[a] Pairwise comparisons are reported as 95% confidence intervals (HSD) around the differences between means when possible. When variance is non-homogenous, the differences between means tested with Dunnett's T3.
[b] $p = .052$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

the majority of the previous, hands-free studies. The following are the rank-ordered proportions of the not-visible (partition or intercom) condition to the face-to-face condition in previous studies (from Table 1), with our proportions inserted in boldface: .39, .40, .48, .50, .56, .57, **.68**, **.75**, .80, .82, .89. If holding the phone had decreased the rate of gesturing in addition to any visibility effect, then the proportions we obtained would have been lower than any of the previous, hands-free studies. Instead, we found proportions at the higher end of the previous range. A third line of evidence is the significant difference between the telephone and tape-recorder conditions. In both conditions, the speaker had only one hand free, yet these were the two conditions that differed significantly in rates of gesturing. Finally, the bias check confirmed that seeing the telephone did not affect the analysts' decisions. Thus, the data consistently support our thesis that holding the telephone did not, in itself, suppress the rate of gesturing as we measured it here.

Gestures' form

Size of gestures. Recall that analysts rated the size of gestures depicting the shape of the skirt on a continuum between picture-sized (1) and life-sized (5). As shown in Table 6, the one-way ANOVA indicated a significant main effect across conditions: $F_{(2,20)} = 78.35$; $MSE = 0.37$; $p < .001$. In the face-to-face condition, the gestures were largest and virtually entirely life-sized ($M = 4.79$), whereas in the other two conditions, the gestures were usually on the scale of the picture (telephone condition, $M = 1.67$; tape recorder condition, $M = 1.47$). The modal size in the face-to-face condition was 5, with no gesture rated lower than 3. In the other two conditions, the modal size was 1, and only one gesture (in the telephone condition) was rated as 5. Not surprisingly, pairwise comparisons (Dunnett's T3, because of heterogeneity of variance) showed that the gestures in the face-to-face condition were significantly larger than those in the telephone and tape recorder conditions, which did not differ from each other. As predicted, the regression analysis revealed only a visibility effect, with no additional effect of dialogue. The life-sized gestures may have required extra effort, but they would clearly be a useful resource for an addressee who could see them. These gestures often started at the speaker's waist and stretched out to the fullest, horizontal extension of the speaker's arms (as in Fig. 3). By depicting the shape of the skirt around the speaker's own body (which had the effect of making them life-sized), the body became an additional resource for the addressee's understanding of the gesture's meaning.

Picture-oriented gestures. As shown in Table 6, gestures directly oriented at the picture occurred at over three times the rate when the addressee would not see them. One interpretation is that these gestures might be helpful to the speaker but not to the addressee and were therefore inhibited (or replaced by other gestures) in the face-to-face condition. The ANOVA, however, revealed that this difference was not significant, just as the regression

analysis indicated no significant effect of either visibility or of dialogue, whether measured as rate per minute or per 100 words.

Interactive gestures. There is a small subset of conversational gestures that do not refer to the speaker's topic (in this case, the dress). Instead, these interactive gestures refer to and include the addressee in the conversation, for example, flicking the hand at the addressee to mark common ground (Bavelas et al., 1992, 1995). In these data, one speaker (in the telephone condition) was an extreme outlier on this variable. The mean rate for this condition without the outlier was 0.38 interactive gestures per minute ($SD = .31$), whereas this speaker made 5.53 interactive gestures per minute (16.61 *SD*s above his cohort's mean). We therefore conducted our analysis without this speaker, for this variable only, and these are the data in Table 6. There was a significant main effect across conditions; $F_{(2,26)} = 5.94$; $MSE = 0.37$; $p < .01$. As predicted, speakers in the face-to-face condition made interactive gestures at the highest rate ($M = 1.21$ interactive gestures per minute), and those in the other two conditions did them at virtually identical lower rates (telephone, $M = 0.38$; tape recorder, $M = 0.39$). The regression analysis confirmed our prediction that only visibility (and not dialogue) affects the use of interactive gestures. Dunnett's T3 (because of heterogeneity of variance) indicated that the rate per minute in the face-to-face condition was significantly higher than in the telephone condition, which replicates Bavelas et al. (1992, Exp. 2), where face-to-face dialogues produced a significantly higher rate of interactive gestures than dialogues through a partition. However, although the means of the telephone and tape recorder conditions were virtually identical, the face-to-face condition achieved only a marginally significant difference ($p = .052$) over the tape-recorder condition in rate per minute, presumably because of much higher variance in the tape-recorder ($SD = 0.66$) than in the telephone condition ($SD = 0.31$). The difference between the face-to-face and tape recorder conditions when calculated by rate per 100 words did not approach significance, although all other effects for this variable were similar to the rate per minute measure.

The different patterns of effects with topic versus interactive gestures illustrate the importance of functional distinctions among gestures. The rate of *topic* gestures on the telephone was close to that of the face-to-face condition and significantly higher than the tape-recorder condition, which was a dialogue effect: Interacting in a real dialogue appears to stimulate the rate of topic gestures, whether the dialogue is with a visible or not visible partner. In contrast, the rate of *interactive* gestures in the telephone condition was significantly lower than in the face-to-face condition and virtually identical to the rate in the tape-recorder condition, which was a visibility effect: Speaking with a visible addressee increased the rate of interactive gestures, whereas having no visible addressee (e.g., on the telephone) suppressed the rate of this particular kind of gesture.

Gestures' relationship with words

Deictic expressions referring to a gesture. When talking face to face, speakers sometimes referred explicitly to their gestures, using terms such as "like this" while gesturing the shape of a feature of the dress, or "over here" while gesturing the location of a feature of the dress. We calculated the proportion of topic gestures that were marked with a deictic expression; see Table 7. The main effect for condition on this variable was significant; $F_{(2,25)} = 23.73$; $MSE = 0.002$; $p < .001$. As predicted, the regression analysis showed a significant visibility effect for deictic expressions, with no additional dialogue effect. In the face-to-face condition, the mean proportion of topic gestures that included a deictic expression was 0.15. This mean was significantly different from the means in the other two conditions, where speakers rarely used deictic expressions with their gestures (telephone, $M = 0.03$; tape recorder, $M = 0.01$). Thus, when they were gesturing in person, the speakers more frequently drew the addressee's attention to their gestures, presumably because these gestures were useful for communicating, whereas such deictic references would only be confusing when the addressee could not see the accompanying gesture.

Redundancy with words. We assessed the average redundancy of gestures with their accompanying words for the two features of the dress that most speakers included in their descriptions, the skirt and the neckline. We called a gesture redundant with the words it accompanied when the only information it was contributing was also present in the words, which would suit conditions in which the addressee could not see the gesture. We also predicted that speakers in the face-to-face condition would take advantage of the visibility of their gestures, so the proportion of gestures that were redundant with words would be lowest in this condition. The results supported these predictions; see Table 7. There was a significant main effect for redundancy of skirt gestures; $F_{(2,20)} = 19.40$; $MSE = 0.05$; $p < .001$. In the face-to-face condition, the mean proportion of gestures that were redundant with the concomitant words was 0.11 for gestures depicting the skirt. In other words, almost 90% of the gestures used in this condition conveyed some information that was not in the immediately accompanying words. The gestures describing the skirt increased in redundancy over the three conditions (telephone $M = .58$; tape recorder $M = .85$). The mean differences between the face-to-face condition and the other two conditions both exceeded the 95% confidence interval

of ±0.30, which is consistent with the visibility effect shown by the regression analysis. However, the additional dialogue effect was almost significant as well ($p = .052$). Unfortunately, we could not clarify this unexpected effect by replication with the neckline descriptions because there were no neckline gestures in the tape-recorder condition. Speakers in the two other conditions did gesture about the neckline sufficiently to replicate the visibility effect: The speakers in the face-to-face condition produced fewer gestures that were redundant with their accompanying words ($M = .18$), and those in the telephone condition produced significantly more that were redundant ($M = .45$). For both features, speakers on the telephone were more likely to use gestures that did not contribute unique information when compared to their accompanying words, although they did often make nonredundant gestures. It is possible that our measure of redundancy may have been too crude, with no differentiation of the amount or quality of the nonredundant information, so it is difficult to interpret the middle status of the telephone gestures on this measure. (We are currently applying a more complex redundancy analysis to these and other data, assessing precisely which features are expressed in words versus gestures; Gerwing & Allison, manuscript in preparation.) However, to ascertain whether face-to-face and telephone gestures differ in the importance or salience of the nonredundant information, it would be necessary to use a stimulus that permitted precise gradations of redundancy and importance in its description. We can infer that, because gestures with deictic expressions were almost by definition not redundant with words, these two effects probably mean that speakers in the face-to-face condition were shifting some information entirely to a gesture and often verbally marking this shift for the addressee.

Figurative language. This was a purely verbal measure, assessed from the transcripts without reference to the gestures. Figurative images included likening the shape of the skirt to a "kettle cozy" or the decoration on the front as "worms butting heads." As shown in Table 7, there was a significant effect of experimental condition on the mean rate of figurative language per 100 words; $F_{(2,27)} = 5.36$; $MSE = 0.27$; $p < .05$. The regression analysis revealed only a dialogue effect, with no effect of visibility. Notably, the dialogue effect was due to the two not-visible conditions: When talking to another person on the telephone, speakers used figurative expressions at a significantly higher rate ($M = 1.11$ figurative expressions per 100 words) than when speaking to a tape recorder ($M = 0.37$). Even though figurative expressions did not rely on visibility and were therefore equally available in both conditions, speakers in monologue used them significantly less often. Lacking visual contact, one way to make descriptions clear would be to create a metaphoric verbal image rather than a visible gestural one, but the speakers were much less likely to create either verbal or gestural images when there was no addressee to imagine them. (Below, we will speculate that there is a close connection between a dialogic partner and iconic imagery). Finally, although the face-to-face condition was not significantly different from the telephone condition, its rate of figurative language ($M = .57$) was noticeably lower, which is consistent with Boerger's (2005) findings that the use of figurative language was lower when the interlocutors could see each other.

Summary and discussion

The results confirmed both of our hypotheses. First, dialogue had a significant effect on the speaker's rate of gesturing, an effect that was independent of and in addition to the effect of visibility. Speakers gestured at a significantly higher rate in a telephone dialogue than in a monologue to a tape recorder, confirming that visibility is not the only variable operating in telephone conversations. Second, the form of the gestures and their relationship to words showed a different pattern from their rate. These features changed only as a function of visibility (not dialogue), and the changes were consistent with speakers' use of gestures as a communicative resource. In the following, we will evaluate the findings for dialogue and for visibility in more detail, then engage in some wider theoretical speculation about a possible relationship between dialogue and demonstration as a mode of signaling.

Dialogue effects

The main purpose of the present design was to bring dialogue out of the background and to test its status as a separate independent variable. Telephone and face-to-face dialogues have in common the presence of and interaction with another person, which are completely lacking in monologues. The addition of the monologue (tape-recorder) condition made it possible to separate the two variables of visibility and dialogue, which had been intertwined in previous studies. The results showed that dialogue (face-to-face and on the telephone) significantly increased how much the speaker gestured, whether measured as the rate of all gestures or as the rate of topic gestures depicting the task stimulus. These effects confirmed our hypothesis that dialogue itself has different effects from monologue, even when other factors such as visibility are the same. It seems plausible to conclude that speakers gesture on the telephone largely because they are in a dialogue. The effect of dialogue on gesture rate replicated Cohen's (1977) suggestive finding and extended the effect that Chovil (1989, 1991) found for facial displays using a design sim-

ilar to ours. We can also add the intriguing effect of dialogue on figurative speech, to which we will return in the last section, below.

To clarify the role of dialogue, we addressed several potential problems in previous studies. Most previous visibility experiments (Alibali et al., 2001; Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss et al., 1995) had prevented full dialogues by constraining the actions of the addressee, whereas (like Bavelas et al., 1992, Exp. 2 & Rimé, 1982) we created unconstrained dialogues. (It is interesting that all of the five experiments that constrained dialogue, creating quasi-monologues, found that the lack of visibility significantly lowered the gesture rate, whereas the other two experiments, plus the present one, which used full dialogues, resulted in a rate that was lower but not significantly so.)

Another problem in the literature was that existing definitions of monologue were based on public speaking and therefore differed from conversational dialogues in more than one way. We created monologues that were more truly soliloquies and also took steps to eliminate an implicit audience, which was a potential confound in Cohen's (1977) tape-recorder condition.

The strong effect of dialogue on the rate of gesturing provides evidence for a social explanation for speakers' gesturing on the telephone, as an alternative to cognitive theories of these gestures (e.g., as encoding or lexical access; Cohen & Harrison, 1973; Krauss et al., 1995; Rimé, 1982). However, we were not seeking to undermine cognitive theories of gestures because, like many other gesture researchers, we do not see communicative and cognitive theories as incompatible (e.g., Bavelas, 1994; Özyürek, 2002; de Ruiter, 2000, 2006). Evidence for one function is not automatically evidence against another. The present experiment was not designed to address cognitive issues, so there are several features of our design and procedure that a cognitive theorist would not adopt: The addressee in the dialogue conditions could contribute spontaneously whenever he or she wished to do so, including helping the speaker find referents à la Clark and Wilkes-Gibbs (1986). Also, because all conditions used the same task, the cognitive requirements did not vary systematically across conditions, nor were they intended to. It might be tempting to interpret the gestures that did occur in the monologue condition (i.e., in the absence of both visibility and dialogue) as evidence for a primarily cognitive function of at least this small group of gestures. However, the presence of the cameras and the proximity of the experimenters were both residual social factors that worked against our hypothesis in the monologue condition and would also confound any cognitive interpretation. One design will not fit (or even inform) all theories. If, as lexical access theory implies, hand gestures are a kind of epistemic action (Kirsh & Maglio, 1994), then this prediction can best be tested by experimental tasks and conditions designed to reveal directly their contribution to word searches.

Returning to our focus on dialogue, we question the implicit reductionism that assumes that verbal monologue is the simplest and therefore basic form of language use, implying that face-to-face dialogue is a more complex variation to which an addressee and gesturing have been added. Our alternative is to propose that face-to-face dialogue with all of its natural features is the basic form of language use (e.g., Bavelas & Chovil, 2000, 2006; Bavelas et al., 1997; Clark, 1996; Fillmore, 1981; Garrod & Pickering, 2004; Goodwin, 1981; Levinson, 1983; Linell, 2005), which implies that monologues are a reduced or altered form in which some of the potential features are not available, not used, or suppressed. If so, the exceptional case is not when gestures are present but when they are absent. Indeed, asking why speakers gesture so much less in monologue may be as fruitful as asking why they gesture significantly more in dialogue. In any case, our goal was to bring dialogue into the forefront of the study of gesture and the study of language use more generally, rather than leaving it to be taken for granted and, even more problematically, studied through controlled interactions with confederates.

Visibility effects

The present experiment confirmed that visibility does have the expected effect on the rate of gesturing, but with several refinements on previous studies. First, we treated visibility as more than a descriptive, physical variable, proposing that it is one aspect of communicative context that would lead speakers to allocate their resources in different ways. That is, visibility affects not only whether speakers gesture but also *how* they do. Analyzing more features of the gestures than just their rate led to strong support for this hypothesis. The results for these variables, summarized below, were consistent with Kendon's (1987) theory that transmission conditions will affect gestural as well as verbal formulations and also with other similar positions (e.g., Bangerter, 2004; Bavelas & Chovil, 2000, 2006; Bavelas et al., 2002; Clark & Krych, 2004; Emmorey & Casey, 2001; Graham & Heywood, 1975; Özyürek, 2002; de Ruiter, 2006).

The difference in the *size* of gestures strongly implies that speakers were using them differently, depending on visibility. In face-to-face dialogue, the size was larger in a particular way: these gestures were life-sized, made in proportion to the speaker's own body. Indeed, the speakers often used their own body to depict the location or scale of features of the dress. Other researchers have seen similar gestures and noted that, for example, "the action ... is performed by the hand, but its meaning resides in a larger context that embraces salient features

of the [immediate] material environment, especially the speaker's corporal form'' (Koschmann & LeBaron, 2002, p. 257). In other words, the meaning of a speaker's gestures derived from the addressee seeing the gesturing hands in relationship to the speaker's body. In narrative terms, speakers in the face-to-face condition presented what McNeill called a *character viewpoint* to the addressee in which ''the narrator's hand plays the part of a character's hand, the narrator's body the part of the character's body, etc.'' (p. 190). As a result, the addressee in face-to-face dialogue could see directly the relationship of the dress to a real person's body.

The results also replicated the findings of Bavelas et al. (1992, 1995) that *interactive gestures* require both visibility and dialogue. As shown in Bavelas et al. (1992, Exp. 2), dialogues in which the participants could not see each other produced significantly lower rates of these gestures with social, interactive functions, presumably because they would not be useful to the addressee.

Equally important, other results demonstrated the *verbal effects* of permitting or restricting visibility of gestures. First, when their gestures would be visible to an addressee, speakers were significantly more likely to refer to them verbally with deictic expressions, replicating the findings of Bangerter (2004), Bavelas et al. (2002), Clark and Krych (2004), and Emmorey and Casey (2001). Second, as found by Emmorey and Casey (2001), Bavelas et al. (2002), and Melinger and Levelt (2004), when the speakers' gestures were visible, our speakers also shifted information to gestures that were not redundant with their words. This allocation of information is consistent with models that propose that language use in face-to-face dialogue comprises *integrated messages* (Bavelas & Chovil, 2000, 2006) or *composite signals* (Clark, 1996). These diverse and robust effects of visibility on both the form of gestures and their relationship to words should encourage other researchers to continue to examine the effects of communicative context on many other features of gestures in addition to their rate.

The present design also built on earlier visibility designs by addressing several of their limitations and potential confounds. In prior studies, dialogue had been a largely unrecognized variable, shared to some degree in both conditions, which particularly obscured the interpretation of gestures that occurred in the not-visible condition. Its explicit manipulation showed that dialogue is one reason speakers gesture even when there is no visible addressee. Although lack of visibility lowers the rate of gesturing, a true dialogue partially counteracts this effect.

Another potential confound in previous research was the manipulation of visibility by the use of a partition or intercom, which inadvertently confounded lack of visibility with unfamiliarity or low sociality (Chovil, 1989, 1991). The use of a real telephone eliminated this potential confound, and careful definition of the dependent variables ensured that having only one hand free would not affect the gesture measures. All of the statistical evidence confirmed that holding a phone did not artifactually lower the rate or form of gesturing as measured here. Indeed, contrary to what many would expect, the rates of gesturing when talking on a phone were closer to face-to-face dialogue than in most previous studies using a partition or intercom. It may well be that the greater familiarity and sociality of the telephone plus the presence of a spontaneous addressee removed artifacts that, in previous studies, had suppressed gesturing.

Thus, the results showed both visibility and dialogue effects. The proposals of Cohen and Harrison (1973, p. 279) and Clark (1996, pp. 179–180) that gestures are ''natural'' or ''integral'' to face-to-face dialogue could imply that the gestures on the telephone are accidental or unmonitored by-products of dialogue. However, it's not that simple. We cannot agree that these gestures were simply habitual or involuntary acts, because their form and relationship to words was systematically different than in the face-to-face dialogue. Speakers on the telephone gestured at a high rate but for the most part made gestures that the addressee would not need to see, and they used the highest rate of figurative language of the three conditions. Thus, rather than being accidental, their gestures were a precise adaptation to a dialogue without visibility. The visibility results suggest a subtle and skillful trade-off by speakers in the two dialogue conditions: Both gestured at high rate, but the speakers whose addressees would see them made gestures whose form and relationship to words would be informative and even essential to their addressees, while the speakers on the telephone did not.

Dialogue and demonstration

Our results, combined with earlier studies, begin to suggest a broader effect of dialogue on language use, which we will present here as two proposals: First, what the three conversational phenomena that have shown dialogue effects in the present or previous experiments (gestures, facial displays, and figurative language) have in common is that they are demonstrations and, second, that demonstrations are intimately tied to dialogue because they rely on creating an image for an addressee. In monologue, there is no addressee, which may therefore suppress this method of signaling.

We start by reviewing Clark and Gerrig (1990) and Clark's (1996) reformulation of Peirce's (1960, pp.359–360) three forms of signals. In Clark and Gerrig's view, these three forms (symbol, icon, and index) are not only abstract categories but three different modalities that interlocutors use, in combination, as methods of signaling in a social interaction. For *describing* they use conventional symbols, primarily words. For *indicating*

they use words such as demonstrative pronouns and also nonverbal acts such as finger-pointing or eye gaze. For *demonstrating* they use selective depictions of the referent, which Clark and Gerrig initially illustrated with verbal quotations. We will propose that demonstration is the modality used by the three features empirically linked to dialogue, namely, gestures, facial displays, and figurative language.

The first evidence for an independent dialogue effect was Cohen's (1977) experiment on gestures, which we have replicated here with improved controls and statistical assessment. By definition, gestures are demonstrations (Clark & Gerrig, 1990, p. 764–769; Clark, 1996, pp.176–180). What the present data show is that their use as demonstrations depends on dialogue as well as visibility.

The next dialogue effect was Chovil's (1989, 1991) experiment showing an independent effect of dialogue on listener's facial displays. When listening to a tape-recorded monologue instead of a "live" speaker, participants' faces were almost impassive, significantly lower in expressiveness than listeners in dialogues on the telephone or through a partition. Clark (1996, pp. 180–182) and Bavelas and Chovil (1997) have argued that conversational facial displays also fit the criteria for demonstrations. Like gestures, facial displays present a selective image for the addressee.

Finally, the present results also showed an independent effect of dialogue on the use of figurative language, which was significantly higher in the telephone dialogues than in the monologues. Both Peirce (1940, p. 105) and Clark (1996, p. 157) identified metaphors as icons or demonstrations. Metaphors such as figurative language offer a selective depiction of the perceptual experience, in that the addressee must understand, for example, that the decoration on the front of the dress has some but not all features of "snakes" (i.e., their shape but not their tactile or behavioral characteristics). Although there is a tendency to equate demonstrations (or icons) with nonverbal acts, both Peirce and Clark explicitly noted that these could use verbal symbols. For example, Peirce (1960,, p. 360) emphasized that

> a symbol may have an icon or an index incorporated into it [and] may require its interpretation to involve the calling up of an image, or a composite photograph of many images of past experiences, as ordinary common nouns and verbs do.

Clark and Gerrig (1990) argued that purely verbal quotations are demonstrations, in which reported speech recreates an image of the original speech. Similarly, figurative language relies on "the calling up of an image" that is evoked by the words. The speakers to a tape recorder rarely called up such an image, presumably because there was no one to visualize it. It is this reliance on Peirce's "calling up of an image" that, we propose, unites conversational gestures, facial displays, and figurative language, all of which have shown dialogue effects. Hand and facial gestures present a selective image visually and directly, while figurative language does so through words.

In other words, demonstration is particularly tied to the presence of an addressee. Clark (1996, p.156) emphasized that speakers describe, demonstrate, or indicate *to someone*. For all forms of demonstrations (quotations, metaphors, gestures, communicative facial displays, etc.), "the point of demonstrating a thing is to enable addressees to experience selective parts of what it would be like to perceive the thing directly" (Clark, 1996, p. 174). That is, these iconic signals are selective transformations of the referent that co-opt the addressee's perception for their meaning. However, the addressee does not simply passively experience the perception but has to sort it into figure and ground; for example,

> We assume that demonstrators intend recipients to recognize that their demonstrations divide into four parts: (i) depictive aspects ... (ii) supportive aspects... (iii) annotative aspects... (iv) incidental aspects..." (Clark & Gerrig, 1990, p. 768).

That is, the speaker relies on the addressee to do some perceptual and/or cognitive work in order to apprehend the intended referent of a demonstration. More specifically, conversational hand gestures, facial displays, and figurative language are all spontaneously improvised for a certain moment in the dialogue, and they may well be even more polysemic than words as symbols. Addressees need to construct the intended meaning, and speakers need to be able to assume or see evidence that they have done so. The fact that monologue cannot satisfy these requirements may be another reason that it does not favor demonstrations.

Returning, finally, to the effects of visibility, it is obvious that a demonstration cannot draw on the addressee's extraction of the intended image unless the addressee has access to the signal that would evoke it. In our face-to-face dialogue condition, the speakers made qualitatively different gestures that took advantage of visibility by creating a vivid and immediate perceptual experience for the addressee. They often "drew" the dress on their own bodies; they sometimes explicitly called their addressee's attention to a gesture by marking it with a deictic verbal reference; and their gestures contributed significantly more nonredundant information than when they would not be visible. They were demonstrating *to* someone who would see the gestures. In contrast, the speakers in the tape-recorder condition had no addressee; they made significantly fewer (and less communicative) gestures and seldom used figurative language. One might suggest that the speakers in monologue suppressed demonstration in favor of

description partly because they were less motivated to be vivid and immediate. However, it is equally plausible that they did so because no addressee would "see" the dress as it might look on the speaker's body, no one would put the nonredundant gesture together with the words, and no one would create a mental image from their figurative language. These speakers limited themselves almost entirely to strict verbal description and used very little gestural or verbal demonstration, even though the latter was freely possible and would have been recorded as part of their monologue: It may be that description versus demonstration is a better distinction here than verbal versus nonverbal.

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.

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

Bavelas, J. B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction, 27*, 201–221.

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

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology, 19*(2), 163–194.

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

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin, 21*(4), 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*(6), 941–952.

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

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

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

Blum, S. (1982). *Eighteenth-century French fashion plates in full color: 64 engravings from the "Galeries des Modes, 1778–1787*. New York: Dover.

Boerger, M. A. (2005). Variations in figurative language use as a function of mode of communication. *Journal of Psycholinguistic Research, 34*, 31–49.

Chovil, N. (1989). Communicative functions of facial displays in conversation. Unpublished Ph.D. dissertation, University of Victoria, Victoria, B.C.

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

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*, 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.

Corts, D. P. (2006). Factors characterizing bursts of figurative language and gesture in college lectures. *Discourse Studies, 8*, 211–233.

Corts, D. P., & Pollio, H. R. (1999). Spontaneous production of figurative language and gesture in college lectures. *Metaphor and Symbol, 14*, 81–100.

Crystal, D. (2001). *A dictionary of language* (2nd ed.). Chicago: University of Chicago Press.

Emmorey, K., & Casey, S. (2001). Gesture, thought and spatial language? *Gesture, 1*, 35–50.

Fillmore, C. (1981). Pragmatics and the description of discourse. In P. Cole (Ed.), *Radical pragmatics* (pp. 143–166). New York: Academic Press.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiality by an implicit audience. *Journal of Personality and Social Psychology, 60*, 229–240.

Fridlund, A. J., Sabini, J. P., Hedlund, L. E., Schaut, J. A., Shenker, J. I., & Knauer, M. J. (1990). Audience effects on solitary faces during imagery: Displaying to the people in your head. *Journal of Nonverbal Behavior, 14*, 113–137.

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.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Science, 8*, 8–11.

Gerwing, J., & Allison, M. (manuscript in preparation). The coordinated semantic integration of words and gestures.

Gerwing, J., & Bavelas, J. B. (2004). Linguistic influences on gesture's form. *Gesture, 4*(2), 157–195.

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

Graham, J. A., & Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech

performance. *European Journal of Social Psychology, 5*, 189–195.

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56*, 155–196.

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

Kendon, A. (1987). On gesture: Its complementary relationship with speech. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 65–97). Hillsdale, NJ: Lawrence Erlbaum.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*, 513–549.

Koschmann, T., & LeBaron, C. (2002). Learner articulation as interactional achievement: Studying the conversation gesture. *Cognition and Instruction, 20*(2), 249–282.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). 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.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

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

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

Mahl, G. F. (1961). Sensory factors in the control of expressive behavior: An experimental study of the function of auditory self-stimulation and visual feedback in the dynamics of vocal and gestural behavior in the interview situation. *Acta Psychologica, 19*, 497–498.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*, 203–220.

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

Mead, G. H. (1934). *Mind, self, and society: From the standpoint of a social behaviorist*. Chicago: University of Chicago Press.

Melinger, A., & Levelt, W. J. M. (2004). Gesture and the communicative intention of the speaker. *Gesture, 4*, 119–141.

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

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

Peirce, C. S. (1960). In C. Hartshorne & P. Weiss (Eds.), *Collected papers of Charles Sanders Peirce, Volume IV: The simplest mathematics*. Cambridge, MA: The Belknap Press of Harvard University Press.

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.

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

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

de Ruiter, J. P. A. (2006). Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech-Language Pathology, 8*, 124–127.

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

Slama-Cazacu, T. (1976). Nonverbal components in message sequence: "Mixed syntax". In W. C. McCormack & S. A. Wurm (Eds.), *Language and man: Anthropological issues* (pp. 217–227). The Hague: Mouton.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*, 5–13.

Weisberg, S. (1985). *Applied linear regression*. New York: Wiley.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology, 51*, 59–91.