

HOW PEOPLE DISQUALIFY: EXPERIMENTAL STUDIES OF SPONTANEOUS WRITTEN DISQUALIFICATION

JANET BEAVIN BAVELAS AND NICOLE CHOVIL*

Disqualification is nonstraightforward communication—messages that say something without really saying it. The four experiments described here examined whether naive, normal individuals might generate disqualified messages. The participants were presented with hypothetical communicative conflicts to which they wrote their own replies. These messages were significantly higher in quantitative measures of disqualification than were the messages written in control conditions. Our conclusion is that disqualified communication is a systematic product of the sender's situation and that anyone trying to avoid saying something will generate such messages.

WHILE looking for a used car in good condition, you see these classified ads:

VOLKSWAGEN FOR SALE. BODY AND ENGINE IN GOOD CONDITION. ONE OWNER, MUST SELL BEFORE DEC. 15. PHONE 721-7550. ASK FOR MARK.

FOR SALE. 1966 VOLKSWAGEN. VERY CHEAP. PERSON WHO LIKES WORKING ON CARS WOULD BE WISE TO BUY THIS CAR.

These two ads, which were written by two participants in the experiments to be described in this report, are similar yet different. The first is direct and clear considering the constraints of classified ad format, but the second seems to become more ambiguous with each reading. Its slippery, equivocal quality is what Watzlawick (1964), among others, called *disqualification*, "a technique that enables one to say something without really saying it" (p. 18). We have been engaged in a program of experimental research on disqualified communication (Bavelas, 1983, 1985; Bavelas & Smith, 1982), and this report will describe the findings of four further experiments in the series.

The prerequisite to studying disqualified communication is to be able to identify it, which requires both a conceptual and an empirical definition. Bavelas and Smith (1982) began with Haley's (1959) insight into such communication, that it avoids what might be seen as essential elements of communication: sender, content, receiver, and context. A straightforward message would convey, "I am saying *this* to *you* in *this situation*," but a disqualified message would be ambiguous on one or more of these four elements.

For example, the second ad above is disqualified on all four dimensions. The *sender* seems distant from the message (in contrast to the personal tone of the first ad). Its *content* is not only cryptic but equivocal, bordering on double entendre. The message is re-addressed from the obvious *receiver* (anyone reading car ads) to a "person who likes working on cars." Finally, nowhere does the message answer

*Janet Beavin Bavelas is Associate Professor of Psychology, University of Victoria, where Nicole Chovil is a doctoral student in Psychology. We wish to thank Deborah Bolt, Mona McAllister, Susan Cathro, Jennifer Mullett, and Jennifer Evans for their assistance in conducting the experiments or scaling sessions. This research was supported by grants from the Social Sciences and Humanities Research Council of Canada and the University of Victoria to the first author. Order of authors on project publications is alphabetical.

clearly the question that is implicitly the *context* of a car ad, "What is the condition of the car?"

We translated Haley's four essential elements of communication into four continuous dimensions on which messages could be scaled, each measuring a different way in which messages could be ordered from straightforward to disqualified. That is, the *sender's own opinion* may be given clearly or, at the other extreme, be unknown. The *content* (what is being said) can range from very clear to totally unclear. The *receiver* may be addressed directly or, at the other extreme, not at all. *Context* is defined as the general situation in which the message occurs, with emphasis on the particular preceding statement—which may be answered precisely or, on the other hand, not at all.

We then developed a procedure whereby lay judges could learn to scale written messages on the above four dimensions (Bavelas & Smith, 1982). Although this method is costly in time (eight to 10 judges are individually trained for up to seven sessions before scaling experimental messages), it has two major advantages: It emphasizes the pragmatic impact of communication—i.e., what do naive decoders make of what they read—and it is extremely reliable in practice (Bavelas & Smith, 1982, pp. 222–224). (The reader is referred to the original article for further details; a step-by-step procedure is available from the first author.)

The next step was to seek the antecedents of disqualification. Although unclear communication is widely attributed to characteristics of the sender, there is a persistent minority opinion that it is "the only possible reaction to an absurd or untenable communication context" (Watzlawick, Beavin & Jackson, 1967, p. 78). Our research has been aimed at developing and testing such a situational theory (Bavelas, 1983, 1985). Using an experimental strategy and forced choice among messages, Bavelas (1983) established that, even in hypothetical situations that might be said to favor "good" communication, disqualified messages were chosen far too often to be due to individual differences in communicative ability or pathology. Rather, they were preferred (or not) as a function of the situation, not the individual. Moreover, the necessary and sufficient characteristic of a situation leading to the choice of disqualification was shown to be *avoidance-avoidance conflict*, in which all of the direct communicative alternatives are negative yet communication is required. (A common example is trying to avoid lying and at the same time to avoid telling a hurtful truth.) In our theory, disqualification is seen as the only possible solution to this conflict—a solution because it manages to avoid the negative alternatives that would surely be encountered by a direct, clear message.

In summary, what appears at first to be fuzzy, inchoate communication has been revealed as systematic in two senses: It can be quantitatively measured with considerable precision, and it can be shown to vary directly and almost completely with specifiable parameters of the situation. Encouraging as these results were, it would be quite unsatisfactory to stop there, principally because these previous experiments used only messages written by the researchers and offered as choices to the respondents. We do not often communicate by choosing among messages written by another, and it could be argued that the results of these experiments are not generalizable. The present experiments address several questions that remain: Do people opt for disqualified messages only when presented to them in forced-choice format? Or would they spontaneously generate their own disqualifications? If so, could messages unique to each writer be scaled reliably by our method? Would these

scale values vary according to the writer's situation, as predicted by our conflict theory? Finally, what would such messages be like—how do people disqualify? In the research to be reported here, there were real senders writing real messages that were measured by their impact on real receivers who scaled the messages.

THE EXPERIMENTS¹

In four independent experiments, volunteer university students wrote their own responses to hypothetical situations. Two of the four situations (the poor class presentation and the unwelcome gift) had been used in earlier studies establishing them as avoidance-avoidance conflicts (Bavelas, 1983), and two new situations (the politician's dilemma and a used car ad) were constructed on the same principles.² In each of the four situations, there were two conditions to which individuals were randomly assigned in equal numbers. In the experimental condition, the participant was faced with an avoidance-avoidance conflict, while in the control condition there was no such conflict. In each experiment, each participant saw only one version of a particular situation, to which she or he wrote a reply. These unedited messages were typed onto cards with no identifying information (for example, as to condition).

All messages for a given situation were arranged in random order and scaled as a set on each of the four dimensions (sender, content, receiver, context) by nine to 11 lay judges according to the procedure described in Bavelas and Smith (1982). The next step of this procedure is to standardize the raw scores for each message on each dimension for each judge and then to average across all judges in the group. This results in very stable scores for each message on each dimension, with low (negative) values indicating relative clarity and high (positive) values indicating disqualification. The judges were in fact the groups identified as "D" and "E" in the above article. As reported there, these two groups had a median intraclass correlation (Ebel, 1951; Winer, 1971, pp. 124–128) of .96 on test trials and a median bivariate correlation of .94 with all previous groups of judges.

Note that once we had set this procedure in motion by designing and presenting the experimental situations, our intrusion as experimenters became peripheral. The participants wrote their own messages; the judges scaled these messages; and we calculated scale scores at the end of the sequence.

The situations, conflicts, and results were as follows: First, 20 people wrote a note to a fellow student in answer to "How did I do?" following a class presentation which (in the experimental condition) this person had done quite badly. This is an avoidance-avoidance conflict between lying or hurting the other's feelings. In the control condition, the asker had done very well. The 20 unique messages were scaled by judges, with intraclass reliabilities of .90 to .94. On all four dimensions and the sum across dimensions, the means of the control messages were negative and of experimental messages positive, as predicted. Sample messages from the control and experimental conditions, respectively, were "I think your presentation was well thought-up and well delivered," and "It was O.K. but there were things that could be improved." The mean scale values of messages written in the two conditions differed significantly on the context dimension, that is, the experimental messages were less responsive to the question asked ($t = 2.29$, $df = 18$, $p < .025$).³

In the second situation, 14 persons⁴ wrote a brief "telegram" as a Member of

Parliament who must reply to a reporter's question about an issue on which (in the experimental condition) his or her constituency is badly divided. The conflict is to avoid offending either group. In the control condition, the constituency was united. Again, each message was different from the others, yet they were scaled with intraclass reliabilities of .89 to .98, and the mean scale values were positive in the experimental condition and negative in the control condition on all four dimensions and their sum. A typical control message was, "Dear Sir, I would prefer Route A due to its obvious favorable aspects. Signed xxxxxx." In contrast, two experimental messages read, "I will send the facts to a committee for further consideration," and "Undecides (sic) as of yet." The mean differences were significant for content, context, and the sum across all four dimensions (smallest $t = 1.93$, $df = 12$, $p < .05$). In other words, experimental messages were significantly less clear in what they said, less responsive to the question asked, and more disqualified overall.

In the third situation, 18 people wrote a thank-you note to a friend who had (in the experimental condition) sent a gift so bizarre that it was unclear whether or not it was intended as a joke. In the control condition, the gift was welcome and suitable. The conflict is between two tacks that might be taken (the gift is serious vs. the gift is a joke), either of which might be wrong and therefore offensive. Because of curtailment of range in the message set, the intraclass reliabilities were from .65 to .93, but all mean scale values split into negative and positive as predicted. A typical control message read, "I can't tell you how thrilled I was to receive the present you sent me. Obviously, you spent some time selecting it, and I appreciate it." Some experimental messages were themselves bizarre, for example, "Your gift, although much appreciated (sic) was definately (sic) uncalled for. Can I see you to talk about it, as I think my feelings were hurt by it. Thanks." Note that both of these sample messages state their appreciation, but the second is almost incoherent as it seems to dodge the twin dangers of ingratitude and insult. The mean scale values differed significantly on three dimensions (sender, content, context) and their sum (smallest $t = 2.67$, $df = 16$, $p < .01$). Thus, the experimental messages were not only less clear in content and less responsive to the question asked (as in the first two situations), they were also judged to be less clearly the writer's own opinion than were the control messages.

Finally, 18 people wrote an ad for a car that must be sold but which (in the experimental condition) was in bad condition. In the control condition, the same car was in good condition. The conflict was between dishonesty with potential buyers and betraying one's self-interest. The intraclass reliabilities were .75 to .98. A sample message from each condition was quoted at the beginning of this article. Other ads written in the experimental condition were cryptic: "AS IS, QUICK SALE FOR CASH," or wordy but equivocal: "MECHANIC'S DREAM—1966 VOLKSWAGEN—BEST YEAR OF THE BUG—HOWEVER NEEDS SOME BUGS REMOVED BY CARING MECHANIC..." As predicted, all experimental means were positive and control means negative, and these differences were significant on all four dimensions and their sum (smallest $t = 1.99$, $df = 16$, $p < .05$). That is, the experimental messages, compared to the control messages, avoided giving the sender's opinion, were unclear in content, did not address the receiver, and did not reply to the implicit question.

The reader may have noticed that there was a significant effect in all four

experiments on the *context* dimension, in three on *content*, in two on *sender*, and in one on *receiver*. This pattern has continued in subsequent experiments and seems to reveal more specifically how people disqualify.

CONCLUSION

These results support our situational theory of disqualification as well as the construct validity of the scaling method. People will generate their own disqualified messages in avoidance-avoidance conflicts, and our scaling procedure will detect this property even when every message is unique. Our next reports will describe the results of a field experiment where the conflict was real for the respondents (Black, Mullett, Bryson & Bavelas, 1985), then further laboratory experiments with spoken and face-to-face messages where both verbal and nonverbal disqualification is possible (Bavelas, Black, Chovil & Mullett, 1985).

NOTES

¹Full details are available from the senior author.

²In addition, subsequent *spoken* versions of these situations have produced differences in latency of response as predicted, for example, by Barker (1942) for avoidance-avoidance conflicts.

³One-tailed tests were used for all *t* comparisons.

⁴Two other participants were excluded because they did not understand the situation.

REFERENCES

- Barker, R.G. (1942). An experimental study of the resolution of conflict by children: Time elapsing and amount of vicarious trial-and-error behavior occurring. In Q. McNemar & M.A. Merrill (Eds.), *Studies in personality* (pp. 13-34). New York: McGraw-Hill.
- Bavelas, J.B. (1983). Situations that lead to disqualification. *Human Communication Research*, 9, 130-145.
- Bavelas, J.B. (1985). A situational theory of disqualification. Using language to "leave the field." In J. Forgas (Ed.), *Language and social situations* (pp. 189-211). New York: Springer.
- Bavelas, J.B., Black, A., Chovil, N., & Mullett, J. (1985). *Truth, lies and alternatives: Disqualification in spoken communication*. Unpublished manuscript, Department of Psychology, University of Victoria.
- Bavelas, J.B., & Smith, B.J. (1982). A method for scaling verbal disqualification. *Human Communication Research*, 8, 214-227.
- Black, A., Mullett, J., Bryson, L., & Bavelas, J.B. (1985). Political disqualification: A situational theory. Paper presented at the annual meeting of the International Communication Association, Honolulu.
- Ebel, R.L. (1951). Estimation of reliability ratings. *Psychometrika*, 16, 407-424.
- Haley, J. (1959). An interactional description of schizophrenia. *Psychiatry*, 22, 321-332.
- Watzlawick, P. (1964). *An anthology of human communication: Text and tape*. Palo Alto, CA: Science and Behavior Books.
- Watzlawick, P., Beavin, J., & Jackson, D.D. (1967). *Pragmatics of human communication. A study of interactional patterns, pathologies, and paradoxes*. New York: Norton.
- Winer, B.J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.