Multiple Audience Problem: A Strategic Communication Perspective on Social Perception

John H. Fleming
University of Minnesota

James L. Hilton
University of Michigan

John M. Darley
Princeton University

Brian A. Kojetin
University of Minnesota

Examined how communicators send mixed messages containing an explicit surface content and a covert hidden content. In Study 1, Ss wrote constrained essays presenting either an introverted or extraverted personality. Although authors reported manipulating essay credibility and readers reported relying on credibility to make their judgments, readers succumbed to correspondence bias. In Studies 2 and 3, Ss again prepared either constrained essays (Study 2) or constrained videotapes (Study 3) and included in them a hidden message that would be understood by only their friends but not by strangers. Observers then read these essays or watched these videotapes. Friends detected and decoded the hidden messages, whereas strangers did not. We discuss these findings in terms of social perception and strategic communication.

A large proportion of everyday social behavior is strategic: People tell lies, feign emotions, ingratiate powerful others, and take stands with which they do not agree. That is, people frequently attempt to manage the impressions and the information they present to others in order to further their interpersonal goals. Although people's strategic behavior may often accurately reflect their underlying personalities, beliefs, or values, people sometimes behave in ways that do not accurately reflect who they are or what they believe.

This behavior may result from a commitment on the part of the person to mislead the audience; the person intends to deceive. The medicine-show man, for example, seeks to convince his audience that the elixir he hawks is of great value. Similarly, the greedy gold digger seeks to convince a wealthy widower that her love is genuine. In contrast, at other times this behavior flows from constraints inherent in the situation in which people find themselves. A defense attorney for an accused pornographer, for example, must play the role of advocate for the client regardless of the attorney's personal feelings toward pornography; debaters who are constrained by a flip of a coin to support a particular side of an issue do not necessarily hold the views they endorse during the debate. But how good are people at detecting misrepresentations that arise either from active attempts to deceive the audience or from the constraints inherent in the situation? At least two literatures in social psychology address these issues. Both reach similar conclusions.

In the past 20 years, over 700 articles have been devoted to understanding the antecedents and consequences of intentionally misleading deceptions (Hyman, 1989; see DePaulo, Stone, & Lasater, 1985, DePaulo, Zuckerman, & Rosenthal, 1980a, 1980b, Ekman, 1985, Knapp & Comadena, 1979, Kraut, 1980, G. R. Miller & Burgoon, 1981, and Zuckerman, DePaulo, & Rosenthal, 1981, 1986 for reviews). These lie detection studies encompass two broad categories: (a) those that address aspects of actors' behaviors that differentiate lying from truthful behavior; and (b) those that examine the processes by which observers determine, first, whether a person is lying or telling the truth and, second, what the concealed truth is in a deceptive performance. The typical finding is that despite the fact that there are documented differences between truthful and deceptive behaviors (for reviews, see Ekman, 1985; Krauss, 1981), people are not very good at perceiving those behaviors and are not very good at detecting lies. By and large, observers take a person's deceptive behavior at face value and detect deception at levels only slightly better than chance. This tendency to take behavior at face value is known in the lie detection literature as the truthfulness bias (Zuckerman, Koestner, Colella, & Alton, 1984).

A similar tendency has emerged in the literature on the dispositional attributions that are made when situational constraint is high. One of the major discoveries to emerge from the attribution literature has been variously labeled the fundamental attribution error (Ross, 1977), observer bias (Jones & Nisbett, 1972), overattribution (Quattrone, 1982), or correspondence bias (Gilbert & Jones, 1986). All of these terms refer to a persistent tendency among observers to underestimate the effects of
situational constraints on behavior, and to attribute causes to a person's internal dispositions instead. Observers who have been asked to read pro- and anti-Castro essays, for example, have assumed that the authors endorse the positions advocated in the essays even when they have been told that the direction of the essays was assigned by the experimenter (Jones & Harris, 1967). Although originally unexpected, this finding has proven both robust (e.g., Jones, Worchel, Goethals, & Grumet, 1971) and generalizable across variations (A. G. Miller, Jones, & Hinkle, 1981; Schneider & Miller, 1975). It has also survived suggestions that it was produced by experimental artifact (Kelley, 1972; Snyder & Jones, 1974), and by now numerous studies have found evidence documenting the existence of correspondence bias (for reviews, see Fiske & Taylor, 1984, Jones, 1979, 1986, Nisbett & Ross, 1980, and Ross, 1977).

The similarities in the findings from the lie detection literature and the attribution literature are not accidental. Both demonstrate the same fundamental tendency for observers to assume that people's behavior accurately reflects their attitudes or dispositions. It is possible, however, that these studies may underestimate communicators' abilities to send complex messages and observers' abilities to interpret those messages, because these studies have relied on a limited class of situations in which a person generates complex communications. What happens, for example, when one sees a captive American decry American imperialism and terrorism on a videotape made by the captors (ABC News, 1986, 1987)? Or, more mundanely, what happens when one sees teenagers who, in the presence of their peers, are forced by their parents to promise to walk the straight and narrow? In other words, what happens when actors must (realistically) concerned that their behavior, constrained by the presence of one audience, will result in unfortunate or undesirable judgments being made about them by another audience (e.g., the American public who sees the hostage videotape, or the teenager's friends)? We refer to these situations as multiple audience problems, because actors face the task of simultaneously sending two different messages to two different audiences receiving the same communication (also see Fleming & Darley, 1989).

Our goal in this research was to begin examining the strategies that communicators and observers adopt when they are confronted with these kinds of situations. We have attempted to address questions such as: What (if anything) do constrained actors attempt to communicate about their true personalities or opinions, and how do they attempt to communicate it when they are constrained to behave in personally discrepant ways? What attributes of the target audience (or audiences) determine the types of communications they will use? What strategies or cues do observers use to interpret communicators' messages and infer actors' dispositions, and how do these strategies parallel communicators' strategies for conveying the underlying truth in their constrained behavior? The answers to these questions may not only reveal more about the nature of lie detection and correspondence bias but may also provide a foundation for a more general understanding of strategic interpersonal communication.

We designed three different studies to examine from a communication perspective how constrained actors attempt to send cues in messages prepared under conditions of high constraint to signal the actual nature of the message to observers. Likewise, we examined the responses and reactions of the observers to see whether they detected those cues or focused on other aspects of the communications instead. Study 1 followed the typical personality attribution/assigned essay paradigm used in many correspondence bias studies. In those studies, authors are assigned a position to present that may or may not contradict their own opinions or personalities. Observers are then given the essays to read with full knowledge that the direction of the essays was determined by the experimenter. The observers' task is then to determine the true attitudes or personalities of the authors. (For a more complete review of studies using this paradigm, see Jones, 1979.) With a few additions to be described later, we followed this general procedure. We also asked the essay authors to do something not normally done in a social perception study: We asked them to write a second essay, under the same constraints as the first, with additional instructions to attempt to signal their own true personalities while keeping the surface content of the essay the same.

STUDY 1

Phase 1: Generating the Essays

Method

The 32 subjects contacted by telephone who served as essay authors were paid volunteers enrolled at Princeton University. After consenting to participate in a communication experiment, they heard descriptions of an introvert and an extravert containing adjectives that have been found to have high prototypicality ratings for introversion and extraversion (Cantor & Mischel, 1977). Authors then rated themselves on a 9-point introversion-extraversion scale. Equal numbers of subjects who rated themselves on either side of the midpoint of the scale agreed to come into the laboratory for an experiment lasting 45–60 min.

Standard Essays

We asked the authors to write an essay designed to convince another person that they were either an extreme introvert or an extreme extravert. Half of the authors were randomly assigned to present a trait corresponding to their self-ratings, and half were randomly assigned to present a trait opposite of their self-ratings. The instructions used in this experiment were previously used by A. G. Miller et al. (1981, p. 590). The experimenter pointed out to the authors that the set of instructions they each received (e.g., "Write about being introverted") was randomly determined and that the only restriction in writing the essays was that the words extravert and introvert could not be used. Finally, they were informed that the essays should be approximately one page long and were given 15 min to complete the task. At this point, if there were no questions, the experimenter left.

Communicate Essays

The experimenter returned 15 min later, collected the first essays, and then presented the written instructions for the second essay. These instructions initially made salient the fact that the instructional set for the first essay had been determined by chance and continued as follows:

What we want you to do now is to rewrite your first essay, only this time we want you to try to get across what you really are. Notice, however, that if you first wrote presenting yourself as an introvert you must still write presenting yourself as an introvert. Similarly, if you first wrote presenting yourself as an extravert you must still write as an extravert. At an overt level the essay topic must remain
The authors were told that they could use as much or as little of their first essay as they could remember. They were then told that there were three constraints on the essay. First, the essay had to be similar in form to the first essay. Second, the words introvert and extravert again could not be used. Third, phrases such as "I'm writing this essay because I have to" were not to be included. Again, they were given 15 min to complete the essay. After the experimenter collected the second essays, authors were asked to complete a two-page questionnaire that contained the dependent measures.1

**Dependent Measures**

We first asked the authors whether they would describe themselves as introverts or extraverts. (There were no shifts from previous self-ratings.) We then asked them which of these traits they had been asked to portray in the two essays; all accurately recalled their assigned traits. Next, we asked them to rate both of their essays on two 9-point scales, rating the extremity of their presentations of the trait first and the credibility of their presentations second. On a final 9-point scale, authors rated how concerned they were with communicating their true personality traits in each essay.2 When they completed the dependent measures, authors were debriefed, thanked for their participation, and excused.

**Results**

**Manipulation Check: Ratings of Concern**

To examine the authors' motivations to transmit their true personalities, a 2 (essay assignment: portray introvert vs. portray extravert) X 2 (actual trait: introvert vs. extravert) X 2 (essay: first vs. second) mixed-model analysis of variance (ANOVA) with repeated measures on the last factor was performed on the authors' reports of concern for presenting their true personalities. Consistent with the instructions they received, a significant essay main effect emerged; authors expressed greater concern for presenting their true personalities in the second essay (M = 7.54) than in the first essay (M = 3.67), F(1, 20) = 71.04, p < .001. A significant Essay Assignment X Actual Trait interaction also emerged; authors who were assigned positions corresponding to their true personalities were more concerned about presenting their true personalities than were authors assigned to present a noncorresponding trait, F(1, 20) = 27.28, p < .001. The interpretations of both of these effects are qualified, however, by the presence of a significant Essay Assignment X Actual Trait X Essay interaction indicating that these differences are confined primarily to the first essay, F(1, 20) = 13.15, p < .01. As the top pair of graphs in Figure 1 indicates, the instructions for the second essay greatly increased the reported motivation of all subjects to convey their true personalities (and especially those who presented a false image of their personalities).

**Ratings of Credibility**

The credibility ratings provided some insight into how the authors attempted to convey their true personalities. A significant essay main effect emerged; authors rated their second essay as less credible (M = 6.12) than their first essay (M = 6.89), F(1, 28) = 5.73, p < .03. A significant Essay Assignment X Actual Trait interaction also emerged; authors who wrote noncorresponding essays rated them as less credible than authors who presented a corresponding trait, F(1, 28) = 23.75, p < .001. Once again, however, the interpretations of these effects are qualified by the presence of a significant Essay Assignment X Actual Trait X Essay interaction indicating that the credibility differences that emerged among authors who wrote noncorresponding essays were limited primarily to the second essay, F(1, 28) = 16.45, p < .001 (see the middle pair of graphs in Figure 1).

**Ratings of Extremity**

Analysis of the authors' perceptions of the extremity of their presentations of the assigned trait revealed that they rated their first essay as more extreme (M = 7.17) than their second essay (M = 6.42), F(1, 28) = 6.59, p < .02. A marginally significant effect for essay assignment emerged; authors who presented themselves as extraverts tended to rate their essays as more extreme (M = 7.19) than did authors who presented themselves as introverts (M = 6.40), F(1, 28) = 3.54, p < .08. A marginally significant Essay Assignment X Actual Trait interaction also emerged; authors who presented a corresponding personality tended to rate their essays as more extreme (Ms = 6.78 and 7.62 for introverts and extraverts, respectively) than did authors who presented a noncorresponding personality (Ms = 6.03 and 6.75 for introverts and extraverts, respectively), F(1, 28) = 3.83, p < .07. Importantly, the three-way interaction was not significant in this analysis; the patterns of extremity ratings did not differ between the first and second essays (see the bottom pair of graphs in Figure 1).

Together, these results suggest that authors in the typical assigned essay paradigm are not particularly motivated to communicate their true personalities when assigned to describe their personalities in a noncorresponding way. When they were motivated to communicate their true personalities, however, they attempted to do so and reported greater concern with communicating an accurate self-image. Apparently, as their debriefing comments suggest, authors attempted to manipulate the credibility of their essays to signal their true personalities. Our next question naturally follows: If authors try to signal their true personalities by manipulating the credibility of their essays, do the readers of the essays notice those cues, and if so, how do those cues affect readers' judgments of the authors' personalities? We addressed this question in Phase 2.

**Phase 2: Readers' Judgments**

**Method**

The 16 subjects who served as essay readers were paid volunteers enrolled at Princeton University and were contacted by telephone to par-

---

1 We gave the questionnaire measures about both essays after the second essay, because we did not want to run the risk of alerting subjects to dimensions useful in writing their second essays. This means, however, that their ratings of their first essays were made with some possibility of comparison with their second essays, and vice versa.

2 Because of experimenter oversight, 8 of the 32 subjects did not receive this question, because we initially failed to recognize its importance. The 24 subjects who did receive the question were distributed evenly across the four conditions.
participate in a communication experiment. All subjects were randomly assigned to conditions. On arriving readers read instructions that stated that in a previous study the author-subjects had written essays designed to convince another person that they were either an extreme introvert or an extreme extravert. Readers then read the same descriptions of an introvert and an extravert that the authors had seen and learned that the authors had been instructed not to use the words introvert and extravert in their essays. We emphasized that the instructional set each author received had been determined by chance. We told the readers that they would read 16 essays that had all been written under the assigned position instructions. They then read and rated the 16 essays; half of the subjects read essays under the assigned position instructions, and half read essays written under the instructions directing authors to convey their true personalities.

After reading each essay, the readers rated the author's true personality trait on a 19-point scale with endpoints labeled extremely introverted (1) and extremely extraverted (19). They then indicated to which of the four experimental conditions the author of the essay had been assigned and, additionally, which personality trait the author had presented.

Figure 1. Authors' perceptions of essays written under the two sets of instructions in Study 1.

3 Recall that we had 32 essays written under the standard set of instructions and 32 essays written under the communicate set of instructions. In order to keep the observers' task within a manageable time framework, the essays written under each set of instructions were divided into two sets of 16 essays. Thus, each of the four sets of four observers read one of the sets of 16 essays. Two sets of observers read essays written under the standard set of instructions, and two sets of observers read essays written under the communicate set of instructions.

4 Authors were first asked to indicate whether they were an introvert or an extravert and then were asked to indicate the extremity of their possession of the trait on a 9-point scale. Similarly, they were asked whether the essay they wrote was in the direction of introversion or extraversion and then were asked to rate the extremity of their presentation of the trait. For observers and paired observers the 19-point scale incorporated both direction and extremity by ranging from extremely introverted (1) to extremely extraverted (19). Reported correlations involving extremity ratings were calculated separately for introverted and extraverted essays using recoded extremity scores that eliminated the direction of the rating.
Results

We used readers' estimations of the true personality of each author (ratings of inferred personality) and readers' ratings of the extremity and credibility of each essay to examine three related questions. First, what information did readers use to infer the authors' personalities? More specifically, what was the relation between judged extremity and credibility of the essays and the resulting judgments of the authors' true personalities? Second, did our readers succumb to correspondence bias, or in the case of the second essays, did they instead detect the changes the authors reported making? Third, what were the relations between the authors' and readers' judgments of the authors' essays?

Informational Bases of Readers' Judgments

As might be expected, readers' inferences about the essay authors' true personalities were strongly related to readers' judgments of the credibility of the essays and less strongly to the readers' judgments of the extremity of the essays; the average correlations between these variables were both significant—\( r(62) = .63, p < .001 \), and \( r(62) = .24, p = .05 \), for credibility and extremity, respectively.\(^3\) Credibility and extremity, however, were not correlated \( r(62) = -.07 \). Because the correlation pattern did not differ significantly for readers of the first and second essays, these results were pooled across the two essays.

To determine the relative contribution of these variables to the readers' judgments of the authors' true personalities, we performed multiple regression analyses on the data from each reader. The averaged multiple regression equation predicting inferred personality from the credibility and extremity of the essays resulted in a correlation of .70 and accounted for 53% of the variance. The average beta weight for credibility in the predictor equations was generally high (.61) and was statistically reliable for all but 5 of the 16 readers. The average beta weight for extremity was significantly lower (.25); for 4 readers, however, it was a statistically reliable predictor of judgments of the author's true personality. These results suggest that most readers' personality judgments were related to the authors' estimates of the essay's credibility, although a few readers' judgments were also related to authors' estimates of extremity over and above credibility (also see Fleming & Kraut, 1988).

Accuracy of Readers' Judgments and Correspondence Bias

As we mentioned earlier, during debriefing, authors reported attempting to manipulate essay credibility to convey their true personalities in the second essay, but not in the first. Furthermore, readers reported basing their judgments of the authors' personalities on their judgments of essay credibility. But it remains to assess the actual success of their joint efforts. This can be determined by examining the relation between the readers' ratings and the authors' self-reported personality traits for the two essays. The average correlations between authors' and readers' ratings of the essays are presented in columns 2 and 3 in the upper panel of Table 1, and as the table shows, these correlations for essays written under both sets of instructions were quite low; readers generally misjudged the authors' true personalities, even when authors tried to convey their true personalities.

If readers' judgments were not related to the authors' true personalities, then to what were they related? An examination of the average correlations between essay personality assignment (i.e., write presenting yourself as an introvert vs. extravert) and the readers' ratings of inferred personality under both sets of instructions provides an answer (see column 4 in the upper panel of Table 1). As Table 1 shows, readers tended to ascribe personality in line with the direction of essay assignment for both essays. In other words, readers tended to fall prey to correspondence bias even when they were reading essays in which authors attempted to convey their true personalities.

Relations Between Authors' and Readers' Judgments

To understand why readers apparently missed the cues authors reported placing in their essays (despite the fact that readers reported using them), we analyzed the correlational relations between readers' and authors' ratings of essay extremity, essay credibility, and author personality type. These correlations are presented in the left-hand portion of Table 2. As Table 2 shows, for essays written under both sets of instructions, the noise in the earlier analyses is evident here in the low average correlations; authors and readers did not seem to agree on what constitutes credibility and, to a lesser extent, extremity.

These results suggest that although authors and readers agreed that essay credibility and extremity were valid cues to signal authors' true personalities, the two groups seemed to disagree on exactly what constitutes those cues. As a result, readers did not detect the authors' credibility and extremity cues, and made correspondent judgments of the authors' personalities. Our next question naturally follows: If readers in the standard assigned essay paradigm (first essays) did not accurately detect the authors' cues signaling their true personalities, under what conditions could readers be made to do so? We addressed this question in Phase 3.

Phase 3: Paired Readers' Judgments

Method

Four additional subjects or paired readers read sets of the first and second essays written by the same author.\(^6\) These paired readers were paid undergraduate volunteers enrolled at Princeton University and received the same information as the readers in Phase 2. In addition,

\(^3\) The average correlations reported were computed by performing a Fisher's \( Z \) transformation on the \( r \) for each subject, averaging the \( Z \) across subjects, and then transforming the resultant \( Z' \) to \( r \). Following Cohen and Cohen (1983), the probability associated with each averaged correlation was determined by first performing Fisher's \( r \) to \( Z \) transformation for each subject and then transforming each \( Z' \) to \( Z \) using a hypothetical population value of \( r = 0 \). We then treated each \( Z \) score as an independent test of the hypothesis and combined them as suggested by Rosenthal (1978). Specifically, we summed the \( Z \)s and divided by the square root of \( N \). The associated probability of the combined \( Z \)s is reported.

\(^6\) In line with the logic of Footnote 3, the 32 pairs of essays were divided into two sets of 16 pairs. Thus, two paired observers read one set of 16 pairs and, the remaining two paired observers read the second set of 16 pairs.
Table 1

Average Correlations Between Readers’ Ratings of Authors’ Personalities and Selected Variables From Study 1

<table>
<thead>
<tr>
<th>Reader ratings of authors’ true personalities</th>
<th>Authors’ self-reported personality trait</th>
<th>Authors’ ratings of actual personality</th>
<th>Authors’ assigned personality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First essays</td>
<td>.07</td>
<td>.15</td>
<td>.55**</td>
</tr>
<tr>
<td>Second essays</td>
<td>.08</td>
<td>.14</td>
<td>.30*</td>
</tr>
<tr>
<td>Paired readers</td>
<td>.20</td>
<td>.22</td>
<td>.53**</td>
</tr>
</tbody>
</table>

Note. \( N = 32 \) ratings for readers, 64 ratings for paired readers.

* Correlations reported are point-biserial correlations calculated by grouping authors as introverts or extroverts on the basis of their self-reports of the trait they possessed.  
  ** Correlations were calculated using authors’ self-reports of introversion–extraversion on 9-point scales.  
  *** Correlations reported are point-biserial correlations calculated using the trait that authors were instructed to present.

paired readers were told of the existence of the essays written under the second instructional set and were told these instructions. We emphasized that each pair of essays was written by the same author but that the order of presentation within pairs had been randomized. After reading each pair of essays, paired readers completed the same dependent measures as those previously completed by the original readers with appropriate changes for the task of rating pairs of essays.

Results

For paired readers we addressed the same three questions we asked of the original readers. First, what information did paired readers use to infer authors’ personalities? Second, were paired readers able to detect the changes the authors reported making in their second essays, or did they again succumb to the correspondence bias? Third, what were the relations between authors’ and paired readers’ judgments of the essays?

Informational Bases of Paired Readers’ Judgments

The average correlations between paired readers’ ratings of inferred personality and essay credibility for essays written under both sets of instructions were significant; average—\( r_s(30) = .46 \) and .63 for the first and second essays, respectively; both \( p < .01 \). The correlations between essay extremity and essay credibility failed to reach significance, as did correlations between inferred personality and essay extremity. Paired readers’ inferences of the authors’ true personalities, like those of the original readers, were strongly related to paired readers’ judgments of essay credibility. Once again, this pattern does not appear to differ for essays written under the two instructional sets, and there is no clear relation between judgments of essay credibility and extremity for either set of essays.

Accuracy of Paired Readers’ Judgments and Correspondence Bias

Apparently, the strategy that paired readers adopted to decode the essays was quite similar to the strategy adopted by the original readers. Paired readers, however, had access to more information than the original readers, in that they knew of the existence of both instructional and could compare each author’s attempts at the two types of essays. We can now ask whether this additional information increased the accuracy of the paired readers’ judgments and whether paired readers were less likely to succumb to correspondence bias. Again, we correlated the authors’ true personalities with paired readers’ ratings of inferred personality (see columns 2 and 3 in the lower panel of Table 1). The resulting correlations were quite low in comparison; paired readers were also relatively unsuccessful in their attempts to infer the authors’ true personalities. We again corre-

Table 2

Average Intercorrelations Among Selected Variables From Study 1

<table>
<thead>
<tr>
<th>Authors’ ratings</th>
<th>Readers’ ratings</th>
<th>Paired readers’ ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay credibility</td>
<td>.14</td>
<td>-</td>
</tr>
<tr>
<td>Essay extremity</td>
<td></td>
<td>- .02</td>
</tr>
<tr>
<td>Essay credibility</td>
<td></td>
<td>- .27*</td>
</tr>
<tr>
<td>Essay extremity</td>
<td></td>
<td>- .28**</td>
</tr>
<tr>
<td>Essay credibility</td>
<td></td>
<td>- .39***</td>
</tr>
<tr>
<td>Essay extremity</td>
<td></td>
<td>- .27*</td>
</tr>
</tbody>
</table>

Note. \( N = 32 \) ratings. Dash means no data are available.

* \( p < .05 \).  
  ** \( p < .01 \).  
  *** \( p < .001 \).
lated the essay personality assignment with the paired readers' ratings of ascribed personality. The resulting correlation again was quite high and positive (see column 4 in the lower panel of Table 1); paired readers tended to ascribe personality in line with the direction of essay assignment.

**Relations Between Authors' and Paired Readers' Judgments**

To understand why paired readers also apparently missed the cues authors reported placing in their essays (despite the fact that readers reported using them), we again analyzed the correlational relations between paired readers' and authors' ratings of essay extremity, essay credibility, and author personality type. These correlations are presented in the right-hand portion of Table 2. As Table 2 shows, the relations between paired readers' ratings and authors' ratings of essay credibility and essay extremity are quite small; apparently, authors and paired readers, like the original readers, did not agree on what constitutes credibility but did seem to agree somewhat on what constitutes extremity.

At this point, because even paired readers who were allowed to see both essays could not infer the authors' true personalities, we became concerned that authors, despite their claims to the contrary, had failed to imbed clues about their true personalities in the second essays. To investigate this possibility, the third author read all 32 pairs of essays with full knowledge of the equal distribution of introverts and extraverts across the four conditions, the order in which the essays were written, and with no time constraints. In other words, the only thing of which the third author was kept unaware was the actual assignment to conditions of each pair. To our relief, 26 of the 32 authors were correctly identified using this procedure (sign test, \( p < .01 \)). Thus, authors apparently did indeed imbed clues in their second essays, but normally instructed readers failed to detect those clues.

**Discussion**

The main findings that emerged from Study 1 are clear: First, authors given the usual rationale for writing a constrained essay presenting a particular personality reported that they were not motivated to convey their true personalities. Not surprisingly, readers did not successfully detect the true personalities of those authors. More surprisingly, authors motivated to convey their true personalities did not seem to be able to communicate that information successfully. Despite the fact that authors and readers claimed to use the same strategies (i.e., manipulating and assessing credibility) and that authors reported being motivated to convey their true personalities in the second essay, they were not successful. Apparently, credibility is not easily manipulated in ways that are clearly understood by the readers.

The failure on the part of our readers to detect the clues that authors embedded in the essays stands in rather stark contrast to the success experienced by others in similar situations. Zuckerman et al. (1979) found that observers were able to detect posed deception (e.g., "Tell a lie, but communicate that you are lying") from the facial expressions and voices of videotaped targets at levels greater than chance. Along different lines, American prisoners of war during the Korean war, who were forced to state false beliefs during "self-criticism" sessions, were able to convince their captors that they held those beliefs while simultaneously conveying to other prisoners that they actually held attitudes in opposition to the ones they were publicly proclaiming (Schein, 1956; Schein, Schneier, & Barker, 1961). Similarly, in a recently released videotape, a Beirut hostage communicated the fact that his confession was made under duress and should not be believed (ABC News, 1987). But of course, from a communication perspective, those prisoners had certain advantages over the subjects in our study: They shared a culture and a language that were alien to their captors. They could rely on shared rules of communication among their friends and acquaintances (or among fellow prisoners) to signal their true beliefs. Friends and acquaintances of the Beirut hostage were the first to detect the falsity of his videotaped communication. In contrast, the authors in our experiment knew nothing specific about the audience to whom they were to convey their true personalities, nothing that might have given them clues about successful communication techniques. They had no chance to spend time with the audience to develop a shared language or set of experiences that could provide the bases for reference that would cue the audience to the fact that the message was other than what it seemed.8

This discussion suggests that knowledge shared with the intended target audience enables a communicator to devise a communication strategy; such shared knowledge may be a critical component of successful covert communication. For example, suppose that the essay readers had been the authors’ friends. If the author wanted to make sure his or her friends disregarded the overt content of the essay, the author could use some common point of reference or experience shared with his or her friends to signal the constrained nature of the essay, perhaps contradicting or lying about some thing or event that the friends would know to be true. In their work on overheard communications, Clark and his colleagues (Clark, 1985; Clark & Schaefer, 1984) have suggested that knowledge shared with the intended target audience enables a communicator to devise a communication strategy; such shared knowledge may be a critical component of successful covert communication.
1987) have referred to these techniques collectively as private keys. We would expect that the larger the body of shared knowledge that exists between the communicator and the audience, the easier it will be to convey a subtle message to them while simultaneously conveying the demanded contradicting message to a different audience. In other words, the successful communication of mixed messages may be less a function of the communicator's motivation to convey the truth and more a function of the nature of the audiences to which the messages are directed. In Study 2, we examined the possibility that audience differentiation (which creates the possibility of differentiated shared knowledge) may serve as one mechanism through which constrained communicators can effectively communicate one message to one audience while simultaneously conveying a second, contradictory message to a different, second audience.

STUDY 2

Overview of Hypotheses and Design

Study 2 examined the multiple audience problem, which was described earlier, in two phases. In the first phase, 8 authors came to the laboratory to write four essays each. In each essay they attempted to persuade readers to select one object from a set of four objects. The authors were told that the essays would be read by two groups of readers: (a) a group of their friends (the friends audience), and (b) a group of strangers (the strangers audience). Operationally, the use of these audiences represents a manipulation of the amount of shared knowledge between the authors and the readers. The four essays comprised two truthful messages and two false messages; one of the false messages contained a hidden message directed at the friends. Authors were told that the communication would be successful if the two audiences correctly received their intended messages; it was unsuccessful if either audience failed to detect their respective messages or if they detected the message intended for the other audience. In the second phase, readers in each of the audiences (friends and strangers) individually read and rated each of the essays.

Initially, we planned to replicate Study 1 by having authors communicate a potentially false message about their introversion or extraversion. This approach, however, would have created the artificial possibility that friends would know or would be able to guess the author's true personality with considerable accuracy. Thus, we shifted the topic of the essay from a communication about personality type to one conveying which object from a set of four the reader should choose. We made one more procedural alteration that was intended to broaden the generality of the results if the predicted findings were obtained. In Study 1, authors described themselves in terms of a dichotomous trait dimension: introversion or extraversion. Consequently, if a reader detected that the overt communication (e.g., "I'm an introvert") was untrue, he or she also knew that the author must be an extravert. In Study 2, we removed this possibility by having the author attempt to signal which of four objects was to be chosen.

The hypotheses tested were first, that authors who were constrained to make a deliberately false statement would attempt to send hidden messages within the constrained communication and would do so by reference to a domain of shared knowledge between themselves and their friends. Second, we predicted that the authors would be successful: That is, essay readers who shared knowledge with the author (his or her friends) would be able to detect and translate a hidden message embedded in a false communication. They would use their shared knowledge to discount (or disregard) the surface content of the message and would correctly determine the concealed content of the message. In contrast, essay readers who did not possess shared knowledge (the strangers) would be significantly less able (or unable) to detect and decode the hidden message embedded in a false communication. Instead, they would receive (and believe) the manifest or surface content of the communication.

Method

Subjects

Eight undergraduate students from the introductory psychology subject pool at the University of Minnesota volunteered for an experiment on social communication and served as the essay authors. Each author-subject received course credit and $5 for participating in the experiment. We recruited the subjects for the stranger audience from the introductory psychology subject pool as well as other introductory psychology classes at the University of Minnesota; all subjects in the stranger audience received course credit for their participation. Thirty-two students served as readers in the stranger audience, 4 readers for each of the 8 authors. Each author also nominated four friends, the majority of whom also tended to be university students. The 31 friends who finally agreed to participate received either course credit or $5 for the half-hour judging session.

Procedure

Essay Authors

Essay authors reported to the laboratory and prepared four short essays ostensibly to be used in a study on lie detection. After completing a set of demographic and filler questionnaires, and reading a pamphlet describing the essay writing task, authors were told that the essays would be read by two different audiences: a group of friends whom they were
asked to nominate and a group of strangers from the introductory psychology subject pool.

The written instructions presented a role-playing scenario in which authors were to imagine themselves as members of an underground resistance movement who had been captured by the enemy. Authors were also instructed to imagine that they were being held hostage and that they were being forced by the enemy to write a set of essays that would later be read by their fellow resistance members (their friends). The essays were to consist of a short persuasive argument that should convince readers in the two different audiences (friends and strangers) to choose one randomly predetermined box (the target) from a set of four boxes ostensibly containing a set of secret maps that would be of assistance to the resistance. We randomly preselected one of the remaining boxes as the decoy and informed the authors which box had been so designated. Each of the boxes (labeled A through D) was unique in color, size, and shape.

To mimic conditions in the typical assigned essay paradigm, the authors wrote four essays: two truthful messages (the truth trials) and two deceptive messages, one consisting of an outright falsehood (the lie trial) and one containing a hidden message designed to get the friends to select the target box (the multiple message trial). The actual content of the arguments for and against each of the four boxes could consist of anything the communicator desired, but in order to minimize the variability among the essays, we instructed authors to follow a prescribed format for the presentation of their arguments. Specifically, they were to request the readers to select the target and then to give a reason (or reasons) for them to do so. Then, communicators were to identify the remaining three alternatives and provide reasons why they should not be selected. Authors were instructed to conclude their arguments with a reiteration of their requests to select the target. The order of trials was randomized for each author.

**Truth trials.** For the two truthful essays, authors were told that they should try to convince both audiences to select the preselected target box and not to select any of the other three boxes. In other words, authors were to lead both audiences to the location of the secret maps (ostensibly) hidden inside the target box. For example, if the maps were hidden inside Box A, then on the truth trials, authors were to get both audiences to select Box A.

**Lie trial.** For the lie trial, authors were told that they should try to convince both audiences to select the preselected decoy box and not to select any of the other three boxes. In other words, authors were instructed to mislead both audiences as to the true location of the secret maps (the target). For example, if the maps were hidden inside Box A (the target) and Box D was identified as the decoy, then on the lie trial, authors were to get both audiences to select Box D. Thus, the difference between the truth trial and the lie trial was that on the lie trial, authors knew which box was the target and they were not directing any of the readers to it; this corresponds closely to other experimental conditions in which introverts write as extraverts with the author and the experimenter being the only ones who know the truth.12

**Multiple message trial.** Finally, in the multiple message trial, which was of primary interest in Study 2, authors again were provided with both a decoy and a target randomly preselected from among the four boxes. On this trial, authors were told that they should try to convince only the strangers to select the decoy box and not to select any of the other three boxes; however, they were also asked to simultaneously communicate the identity of the target box only to their friends so that their friends would select that box. In other words, authors were to mislead the strangers as to the location of the secret maps while simultaneously communicating the true location to their friends. For example, if the maps were hidden inside Box A (the target) and Box D was identified as the decoy, then on the multiple message trial, authors were to get the strangers to select Box D and their friends to select Box A. Thus, in the multiple message trial, authors were aware not only that they were not telling the truth, but also that they were actually attempting to actively deceive the strangers and to send a hidden message to their friends.

The experimenter placed the boxes on the table in front of the authors and gave them as much time as they wanted to write their essays. The experimenter answered any questions but would not give any direct aid in deciding on a particular strategy to use to communicate the various messages. When they were finished, authors reviewed their essays and were allowed to rewrite any essay with which they were not satisfied. Finally, we arranged the four essays for each author in rating booklets, each in a different counterbalanced order.

**Essay Readers**

Readers arrived for a study on lie detection, completed a set of demographic questionnaires, and were told that they would be reading a series of four essays consisting of a short persuasive argument to get them to choose one box from a set of four boxes. Each then received a booklet describing the essay reading task. For the friends, the written instructions presented a role-playing scenario in which they were to imagine themselves as a member of an underground resistance movement designated to receive secret communications from other resistance members. We also instructed readers to imagine that the essay author might have been captured by the enemy and there was reason to suspect the truthfulness of the essays. Specifically, we told the readers that of the four essays they would read, two truthfully directed them to the secret maps and two directed them to incorrect locations. Strangers received identical instructions, except that their affiliation was with the enemy rather than with the resistance. None of the readers was explicitly told about the existence of the hidden communication.

The experimenter placed the four boxes on the table in front of the readers. After reading each essay, all readers indicated which box they believed had been designated as the target and rated their confidence in that choice and the other alternatives on an 11-point scale with endpoints labeled very confident the box was the correct choice (11) and very confident the box was not the correct choice (1) with an unmarked, neutral midpoint. We slightly restricted readers' responses on this scale such that their ratings of the chosen box had to be placed above the midpoint toward being confident that the box was correct. The other nonchosen alternatives had to be rated on the other side of the midpoint toward being confident that the box was not correct. In other words, readers rated all four boxes on the same scale, with the stipulation that their choice be placed on the way side of the midpoint and the other three alternatives be placed on the way side. If they had simply guessed, we told them that they could put all of their choices on the scale midpoint. Finally, on an 11-point scale with endpoints labeled not at all (1) and very well (11), all participants indicated how well, prior to the experiment, they knew the essay author. Also, the participants wrote a brief explanation of what information or cues they had used to decide on the chosen box.

After completing their task, all subjects were told about the existence of the hidden message, debriefed about the actual purpose of the study, thanked for their participation, and excused.

**Results**

**Ratings of How Well Readers Knew Author**

In Study 2, we operationalized shared knowledge by using two groups of readers who differed in the extent to which they

---

12 Of course, we need to acknowledge that the box paradigm differs from the assigned essay/personality attribution paradigm in several ways. Specifically, in the personality attribution paradigm, authors presumably have some investment in being accurately perceived; in the box paradigm, this investment is lower. We attempted to reduce this difference by making the communication task as involving as possible for the authors.
knew the author. As a check on this manipulation, readers’ ratings of how well they knew the author were submitted to a 2 (reader audience: friends vs. strangers) × 8 (author) ANOVA. This analysis revealed a significant main effect for the reader audience, F(1, 47) = 295.25, p < .0001. Friends reported knowing the author significantly better (M = 11.48) than did the strangers (M = 2.44). Neither the author main effect nor the Reader Audience × Author interaction was significant on this measure, both Fs(7, 47) < 1.9.

**Accuracy at Selecting Target Box**

Because we were interested in differences between our two reader audiences in their ability to correctly select the target, our primary dependent measure was the number of hits (correct selections of the target from each set of boxes on each message trial) registered by members of the two reader audiences. In other words, our main dependent measure was the number of correct guesses on each message trial.

With categorical data such as these, it is generally appropriate to use one of the more recently developed nonparametric analytic techniques such as loglinear models, logit or probit regressions, or categorical modeling programs. Because the large number of structural zeros in the present design would render such procedures too unreliable, however, we decided to use a different strategy. Specifically, within each audience, we calculated the proportion of hits among the readers for each author on each message trial (thereby collapsing the data across each author’s readers). Then, following procedures recommended by Bishop, Feinberg, and Holland (1975, pp. 367–372), Cohen and Cohen (1983, pp. 265–270), and Kenny (1985, pp. 502–503), we transformed those proportions using an arc sine (angular) transformation; this transformation strategy yields data that as closely as possible approximate parametric data that can then be analyzed using conventional ANOVA procedures. Thus, we submitted the transformed percentages (as well as the raw proportions) to separate 2 (reader audience: friends vs. strangers) × 3 (message trial: truth, lie, and multiple message) fixed effects ANOVAs, with repeated measures on the second factor. Because the results of the analyses of the transformed and raw proportions were virtually identical, for ease of explication we will report only the raw proportion results.

We expected both reader audiences to achieve approximately equal numbers of hits on both the truth and the lie trials. On the multiple message trial, however, we expected the friends to achieve a significantly larger proportion of hits than the strangers. In other words, we predicted a Reader Audience × Message Trial interaction. As can be seen in Table 3, our predictions were supported. Specifically, a significant Reader Audience × Message Trial interaction emerged with all means in the anticipated directions, F(2, 28) = 6.64, p < .01. A simple effects analysis revealed no significant differences between the reader audiences in the proportion of hits on the truth trials or the lie trial, Fs(1, 41) = 1.33 and 1.45, respectively. The predicted simple effect difference did emerge, however, on the multiple message trial. Friends registered a significantly higher proportion of hits (correctly guessing that the actual target location was the correct choice on that trial) than did the strangers, F(1, 41) = 13.08, p < .001.

In addition, a significant repeated-measures effect emerged for the message trial variable, F(2, 28) = 13.58, p < .001. Across the audiences, readers were significantly more accurate on the truth trials than they were on the lie trial or multiple message trial. The reader audience main effect was marginal in this analysis, F(1, 14) = 3.40, p < .09. Across all message trials, friends registered a marginally higher proportion of hits than did the strangers, reflecting that the friends had significantly greater accuracy on the multiple message trial.

**Readers’ Confidence in Their Choices**

Readers’ ratings of their confidence that the target box was or was not the correct choice were submitted to a 2 (reader audience) × 3 (truth, lie, multiple message) × 5 (number of confidence responses) mixed ANOVA. Corresponding z scores for lie detection analyses

<table>
<thead>
<tr>
<th>Reader audience</th>
<th>Truth</th>
<th>Lie</th>
<th>Multiple message</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Truth detection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends</td>
<td>53.12</td>
<td>6.25</td>
<td>58.25</td>
<td>38.54</td>
</tr>
<tr>
<td>Strangers</td>
<td>41.15</td>
<td>18.75</td>
<td>18.75</td>
<td>26.22</td>
</tr>
<tr>
<td>Marginal</td>
<td>47.14</td>
<td>12.54</td>
<td>37.50</td>
<td>—</td>
</tr>
<tr>
<td><strong>Confidence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends</td>
<td>8.70</td>
<td>3.69</td>
<td>9.07</td>
<td>7.15</td>
</tr>
<tr>
<td>Strangers</td>
<td>7.20</td>
<td>5.86</td>
<td>5.89</td>
<td>6.32</td>
</tr>
<tr>
<td>Marginal</td>
<td>7.95</td>
<td>4.78</td>
<td>7.48</td>
<td>—</td>
</tr>
<tr>
<td><strong>Lie detection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends</td>
<td>46.87</td>
<td>40.62</td>
<td>75.00</td>
<td>54.17</td>
</tr>
<tr>
<td>Strangers</td>
<td>58.85</td>
<td>46.87</td>
<td>50.00</td>
<td>51.91</td>
</tr>
<tr>
<td>Marginal</td>
<td>52.86</td>
<td>43.73</td>
<td>62.50</td>
<td>—</td>
</tr>
<tr>
<td>Corresponding z scores for lie detection analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friends</td>
<td>−0.35</td>
<td>−1.06</td>
<td>2.83**</td>
<td>0.82</td>
</tr>
<tr>
<td>Strangers</td>
<td>1.00</td>
<td>−0.35</td>
<td>0.00</td>
<td>0.37</td>
</tr>
<tr>
<td>Marginal</td>
<td>0.46</td>
<td>−1.00</td>
<td>2.00*</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. Cell n = 8. Except for the lie detection means, cell means within a column or row that share a lowercase subscript, and marginal means that share an uppercase subscript, do not differ at p < .05 by Newman-Keuls test. Dash means no data are available. *p < .05. **p < .01.

13 To balance the number of truthful and deceptive trials each observer would watch, we originally used two truth trials. Because these two truth trials were identical in every respect, we pooled observers’ responses on these two trials into a single combined truth trial in which the mean number of hits across the truth trials was used.
of significance in the analyses of the transformed proportions. Specifically, a significant Reader Audience × Message Trial interaction emerged, $F(2, 94) = 8.74, p < .001$. Simple effects analysis revealed no significant differences between the two reader audiences in their confidence that the actual target box was the correct choice on the truth trials, $F(1, 135) = 2.23$. Strangers reported significantly more confidence, however, that the actual target box was the correct choice on the lie trial, $F(1, 135) = 4.71, p < .05$. Finally, the predicted simple effect difference emerged on the multiple message trial, $F(1, 135) = 10.12, p < .002$. Friends expressed significantly greater confidence that the actual target box was the correct choice on that trial than did the strangers.

In addition, a significant repeated-measures effect emerged across the message trial variable, $F(2, 94) = 13.71, p < .001$. Readers were significantly more confident that the target was the correct choice on the truth trials and the multiple message trial than they were on the lie trial. Confidence ratings for the truth trials and the multiple message trial did not differ reliably. Finally, the reader audience main effect was not significant, $F(1, 47) = 1.62$.

### Lie Detection Versus Truth Detection

The first major question addressed by Study 2 was whether readers who shared common knowledge with an author were more accurate than those who did not in determining the underlying truth in a potentially false communication. The results suggest that they were more accurate. A second and equally important question, however, is whether this increased accuracy was obtained at the cost of being discovered sending hidden messages by the second audience. In other words, did authors sacrifice secrecy in order to convey their hidden messages, or were they able to communicate them without being discovered? Multiple audience problems differ from situations in which one can be overheard (e.g., Clark & Schaefer, 1987) in that communicators caught in multiple audience problems must not only communicate a different message to one of the present audiences but must also do so without being detected by the other audience.

The data from this experiment also allowed us to address this issue. Specifically, within each reader audience, we recoded the original truth detection data and calculated the proportion of responses in which the readers selected a box other than the one the author told them to select (i.e., selected a box other than the target box on the truth trials or the decoy box on the lie trial and multiple message trial) for each author on each message trial. In other words, in this analysis we used the proportion of responses in which readers were not convinced by the surface content of the essay and suspected the author of lying. As with the truth detection analyses, we transformed these proportions using an arcsine transformation and then submitted these data sets to separate 2 (reader audience) × 3 (message trial) ANOVAs, with repeated measures on the second factor. Again, because the patterns of significance in the analyses of the transformed and raw proportions were identical, we report only the raw proportion results.

We reasoned that if authors were communicating their hidden messages only to their friends without betraying to the strangers the fact that they were sending hidden messages, then no differences should emerge across message trials in the extent to which the strangers suspected a lie (i.e., they selected a box other than the one the author requested them to select). Moreover, we expected that the friends would not differ reliably on this measure either, with the exception that their demonstrated enhanced accuracy on the multiple message trial would also translate into better lie detection on that trial. Consistent with this interpretation, results of this analysis revealed no significant main or interaction effects (all $Fs < 2.14$, all $ps > .13$; see the third panel of Table 3). Across all message trials, neither audience selected a box other than the one the communicator requested them to select more frequently, $F(1, 14) < 1$; nor were the combined audiences better able to detect a lie from among the three message trials, $F(1, 14) = 1.90, p < .17$. We also compared each of these marginal means with the chance guessing rate of 50%. As the bottom panel in Table 3 shows, with the exception of the multiple message trial mean, none of the marginal means differed significantly from chance levels. The guessing rate for the multiple message trial did exceed the chance level as a result of the friends' significantly better hit rate on that trial. Thus, in general, it appears that readers in both audiences adhered to our original baseline parameters for lying and truth telling: Two of the trials were true, and two were false. More important, these data suggest that authors successfully communicated their hidden messages without arousing the suspicions of the strangers. Consequently, this demonstrates that the authors successfully executed their strategic communication task and effectively eliminates a number of guessing strategy explanations for the previously demonstrated effects.

### Authors' Hidden Message Strategies

While completing the essay writing task, all 8 authors reported that the task seemed extremely difficult. Nonetheless, all senders were able to complete the task. Without help or guidance from the experimenter, 7 of the 8 authors arrived at essentially similar strategies for communicating their hidden messages: They all used the falsification of some self-relevant information that only their friends would know about them. For example, one author who had a love for Wales and chocolate, but not for a local department store, communicated the hidden message by talking about the possibility of traveling to Wales while describing the box she wanted her friends to select, and by mentioning the department store while praising the box she wanted the strangers to choose. Another author mentioned the names of several potential dating partners while presenting reasons why the target box should not be chosen. The eighth author used a different (and substantially less effective) strategy; the author capitalized every word in the essay that began with the letter B to communicate that Box B should be chosen. Compar-

---

14 Again, observers' responses on the two truth trials were pooled. Specifically, we used the mean confidence rating across both truth trials for each observer.
isons of the readers’ explanations for their strategies, coupled with measures of communication success, suggested that although the hidden communications were, in general, quite successful, some (such as the capitalization ruse) were less successful than others. Moreover, communicators who were unsuccessful usually failed not because their friends could not understand their messages, but instead because the strangers were able to detect and decode them. That is, failures were due to using a strategy that had too large a scope, strategies that did not effectively differentiate among the audiences.

STUDY 3

In Study 2, the communicators prepared written essays to communicate the various messages to the readers in the strategic communication task we gave them. Obviously, readers had access only to the verbal content of the messages in making their judgments; written essays do not incorporate the numerous nonverbal and paralinguistic cues that not only may facilitate the communication of hidden messages but also may lead to their increased detection by unintended audiences (see Zucker-...)

Thus, in Study 3, individuals made a set of four videotapes, in one of which they were asked to communicate a false message to two different audiences while simultaneously conveying a true and contradictory message to a third audience. The (student) senders knew that the first two audiences would be fellow students who did not know them, whereas the third audience would be students whom they had identified as their friends.

Method

Subjects

We recruited 10 senders to participate in Study 3: 8 graduate students (4 women, 4 men) from six different specialty areas within the psychology department at the University of Minnesota, 1 male undergraduate research assistant, and 1 male undergraduate office worker. The mean age of the senders was 24 years. Each received $10 for the 2-hr videotaping session. Each sender nominated four friends to serve as observers, the majority of whom tended to be University of Minnesota social science graduate students. The 40 friends (20 women, 20 men, mean age = 27 years) were paid $5 for the half-hour judging session.

We also recruited a group of strangers who were matched to the friends on several variables, including age and academic status. The 40 matched strangers (24 women, 16 men, mean age = 27 years), 4 for each of the 10 senders, were also University of Minnesota social science graduate student volunteers and advanced psychology undergraduate students; each student received $5 for their participation. Finally, the unmatched strangers in Study 3 were recruited from the introductory psychology subject pool as well as other introductory psychology classes at the University of Minnesota. These subjects either received course credit or $5 for their participation. A total of 40 students (22 women, 18 men, mean age = 21 years) served as unmatched strangers, 4 for each of the 10 senders.

Procedure

Senders

With the exceptions described later, the procedure for Study 3 was identical to that used in Study 2. In the first phase, the 10 senders came to the laboratory to make four videotaped messages following the procedures previously described. In each of the videotapes, they attempted to persuade observers to choose one randomly predetermined song from a set of four popular songs. To minimize variability in the content and style of the songs, we used four popular songs from the same artist or group for each videotaped message. The senders were told that the videotapes would be viewed by two groups of observers: (a) a group of their close friends (the friends audience), and (b) two groups of persons whom they did not know (the matched and unmatched strangers audiences). Again, the use of these groups represented a manipulation of the amount of shared knowledge between the senders and the observers. As in Study 2, the four videotaped trials comprised two truthful messages and two false messages; one of the false trials contained a hidden message directed at the friends.

Before making the videotapes, senders were allowed to listen to a 1-min segment of each of the songs and were given as much time as they wanted to think, make notes, or write an argument verbatim. The experimenter answered questions but would not give senders any direct aid in deciding on a particular strategy to use. When senders were ready, they were videotaped individually presenting their arguments. The senders reviewed their videotapes and could remake any tape with which they were not satisfied; only two senders decided to remake one trial each. Finally, we transferred the four videotapes for each sender to a master tape, each arranged in a different counterbalanced order.

Observers

Observers arrived for a study on lie detection and completed a brief demographic information sheet. We then told them that they would be watching a series of videotapes of a person delivering a short persuasive argument to get them to choose one song from a set of four popular songs. Unlike Study 2, however, the role-playing task that the senders had completed was not described to the observers. Instead, observers were told only that in two of the four tapes they would see, the person was telling the truth and that in the other two, the person was lying. Observers were asked to carefully watch the videotapes and try to decide whether the person was lying or telling the truth, and to guess which of the four songs had been preselected as the target song. The observers were also allowed to listen to 1-min segments from each of the songs before watching the videotapes. After viewing each tape, all observers indicated which song they believed was the correct choice (had been designated as the target) and completed the same dependent measures used in Study 2. After completing their task, all subjects were told about the existence of the hidden
Results

8.59) than the un-

Accuracy at Selecting Target Song

Message trial

Table 4

Observer audience | Truth | Lie | Multiple message | Marginal
---|---|---|---|---
Friends | 57.50<sub>A</sub> | 10.00<sub>b</sub> | 42.50<sub>A</sub> | 36.67<sub>A</sub>
Matched strangers | 51.25<sub>A</sub> | 20.00<sub>b</sub> | 17.50<sub>b</sub> | 29.58<sub>A</sub>
Unmatched strangers | 48.75<sub>A</sub> | 17.50<sub>b</sub> | 20.00<sub>b</sub> | 28.75<sub>B</sub>
Marginal | 52.50<sub>A</sub> | 15.83<sub>C</sub> | 26.67<sub>b</sub> | —

Confidence<sup>b</sup>

Friends | 6.66<sub>a</sub> | 4.29<sub>b</sub> | 6.01<sub>b</sub> | 5.65<sub>A</sub>
Matched strangers | 6.02<sub>a</sub> | 4.56<sub>b</sub> | 4.55<sub>b</sub> | 5.05<sub>A</sub>
Unmatched strangers | 6.24<sub>a</sub> | 4.14<sub>b</sub> | 4.26<sub>b</sub> | 4.88<sub>A</sub>
Marginal | 6.30<sub>a</sub> | 4.33<sub>c</sub> | 4.54<sub>b</sub> | —

Lie detection<sup>c</sup>

Friends | 42.50<sub>a</sub> | 47.50<sub>a</sub> | 67.50<sub>a</sub> | 52.50<sub>A</sub>
Matched strangers | 48.75<sub>a</sub> | 52.50<sub>a</sub> | 52.50<sub>a</sub> | 54.58<sub>A</sub>
Unmatched strangers | 51.25<sub>a</sub> | 62.50<sub>a</sub> | 52.50<sub>a</sub> | 52.08<sub>A</sub>
Marginal | 47.50<sub>a</sub> | 54.17<sub>C</sub> | 57.50<sub>C</sub> | —

Corresponding z scores for lie detection analyses

Friends | -0.95 | -0.32 | 2.21<sup>**</sup> | 0.55
Matched strangers | 0.16 | 0.32 | 0.32 | 0.46
Unmatched strangers | -0.16 | 1.58 | 0.32 | 1.00
Marginal | -0.55 | 0.91 | 1.64<sup>*</sup> | —

Note. Cell n = 10. Cell means within a column or row that share a lowercase subscript, and marginal means that share an uppercase subscript, do not differ at p < .05 by Newman-Keuls test. Dash means no data are available.

* Entries are proportions of correct selections of the target song.

<sup>a</sup> Entries are ratings made on 15-point scales of observers' confidence in the song that was the target on each trial. <sup>b</sup> Entries are proportions of responses in which observers selected a song other than the one senders requested them to select (i.e., observers suspected a lie and did not use the surface content of the message to make their choices). <sup>c</sup> p = .10. <sup>**</sup> p < .01.

the three audiences, observers were most accurate on the truth trials, significantly less accurate on the multiple message trial, and even less accurate on the lie trial (see Table 4 for specific mean comparisons).

Observers' Confidence in Their Choices

Because no significant main or interaction effects emerged from preliminary analyses that included the senders variable, this variable was not included in subsequent analyses. Consequently, we submitted observers' confidence ratings to a 3 (ob-
server audience) × 3 (message trial) ANOVA, with repeated measures on the last factor.

We expected the results of this analysis to be identical to those from our main dependent measure, and as the middle panel of Table 4 shows, they were. Specifically, a significant Observer Audience × Message Trial interaction emerged, $F(4, 234) = 2.54, p < .05$. No significant differences emerged among the observer audiences in their confidence that the actual target song was the correct choice on the truth trials or the lie trial, both $F_s(2, 329) < 1$. Again, however, the predicted simple effect difference emerged on the multiple message trial, $F(2, 329) = 6.77, p < .001$. Friends expressed significantly greater confidence that the actual target song was the correct choice on that trial than did either the matched or unmatched strangers.

Finally, a significant repeated-measures effect emerged across the message trial variable, $F(2, 234) = 28.79, p < .001$. Observers were most confident that the target was the correct choice on the truth trials, significantly less confident that the target was the correct choice on the multiple message trial, and least confident that the target was the correct choice on the lie trial. A marginal main effect also emerged among the three observer audiences, $F(2, 117) = 2.79, p < .07$. Friends were marginally more confident in their ratings of the actual target than were the matched or unmatched strangers, which were not reliably different.

### Lie Detection Versus Truth Detection

Again, one of the major questions addressed by Study 3 was whether observers who shared common knowledge with a sender were more accurate than those who did not in determining the underlying truth in a potentially false communication. The results demonstrated that even when the communication channel is broad (i.e., using videotaped rather than written essays) and even when observers are not told about the senders’ task, observers are more accurate. As in Study 2, a second and equally important question was whether this increased accuracy was obtained at the cost of being discovered sending hidden messages by other observers. It is possible that increasing the information available to all observers by using a broad communication channel could actually lead to higher lie detection rates by all three audiences in Study 3. The same lie detection analyses that we performed in Study 2 can also be applied here; the cell and marginal means from these analyses are presented in the lower panel of Table 4.

Results of the lie detection analysis revealed no significant main or interaction effects (all $F_s < 1.06, ps > .35$). As Table 4 shows, across all message trials, none of the audiences selected a song other than the one the sender requested them to select more frequently; nor were the combined audiences better able to detect lies from among the three message trials. Again, we compared each of the marginal and cell means with the chance guessing rate of 50%. As in Study 2, with the exception of the multiple message trial mean, none of the marginal means differed significantly from chance levels. The guessing rate for the multiple message trial did marginally exceed the chance level (significant at $p = .10$) as a result of the friends’ significantly better hit rate on that trial. Thus it appears that observers in the three audiences in Study 3 also adhered to our original baseline parameters for lying and truth telling: Two of the trials were true, and two were false. More important, these data suggest that senders successfully communicated their hidden messages without arousing the suspicions of the strangers, even though the communication channel had been broadened. Consequently, this supports the major findings of this research: Not only were the senders able to communicate different messages to the different witnessing audiences, but they managed to communicate a message to their friends that was undetected by the strangers.

### Senders’ Hidden Message Strategies

While completing the videotaping task, all 10 senders reported that the task seemed extremely difficult. (Nonetheless, all senders were again able to complete the task.) Without help or guidance from the experimenter, all 10 senders arrived at essentially similar strategies for communicating their hidden messages: Eight senders used the falsification of some self-relevant information that only their friends would know about them, and 2 senders used a set of uncharacteristic (or characteristic but inappropriate) facial expressions while making the videotapes. For example, 1 sender who had recently lost a great deal of weight communicated his hidden message while describing the target song by talking about how he was presently gaining weight. Another sender mentioned the name of her fiancé while presenting reasons why the target song should not be chosen. A different sender smirked and raised her eyebrows when telling the observers not to select the target song—a characteristic facial expression that to her friends meant she was being sarcastic.

### Discussion

The present results provide strong evidence for a multiple audience effect. Communicators in both Studies 2 and 3 spontaneously generated techniques to conceal cues to their friends that the overt content of the messages they were communicating to strangers was sometimes false. These techniques involved the use of private keys—references to knowledge that is shared between the communicator and the friends (but not the communicator and the strangers)—to alert the friends both to the fact that this was one of the trials on which deception was taking place and to what the correct selection should be. In Studies 2 and 3, the most frequent variant of this strategy involved referring to some personal information accessible to friends but not to strangers. In addition, senders in Study 3 used another variant of the private keying strategy. This second strategy involved using facial expressions that because of the history surrounding these expressions in interactions with friends, apparently had meanings that were known and highly likely to be recognized within the friendship group but were not likely to be so idiosyncratic or out of place as to give away their meanings to strangers.

These strategies were generally successful, in that friends, who were alerted to the existence of a certain frequency of deceptive messages within the set of four essays or videotapes, were often able to infer the correct answer in the one message in which the communicators attempted to encode the truth in a way that would be detectable by the friends. Although equally alerting to the existence of a certain frequency of deceptive messages within the set of four messages, strangers were not gener-
ally able to detect or decode the true content of the hidden messages.

In addition, it is worth noting that none of the observers in Study 3 were aware of the possibility that hidden messages were being sent when the task began. Instead, we simply told them that the sender had lied on half of the trials. Interestingly, this created a predicament for the observers. Although many of the friends detected and correctly interpreted the meaning of the hidden signals, some who detected them did not interpret them correctly. Using an unusually Byzantine logic, they reasoned that their friends (the senders) used the clues they detected to throw them further off the correct track. That is, because they knew that the sender was instructed to lie on some trials, they reasoned that the senders were using the hidden messages to deceive them more effectively rather than to provide correct information. What is clear is that the task we gave the senders and the observers was by no means a simple one. These observations underscore at least one way in which covert communication attempts can go wrong.

Because we used multiple, rather than two, choice alternatives within each communication and because the communications were conveyed both in written form and on videotape, we are able to conclude that the major path to the preceding result was not solely through the detection that the communicator was lying. Nor, more specifically, was it through the increased likelihood that the communicator's friends were better able to recognize the emission of the traditional lying cues of deception cues and leakage (Ekman & Friesen, 1969, 1974; cf. Zuckerman et al., 1979). Rather, friends were able to recognize the meanings of the clues that the senders planted in the communication directed solely at them and were able to decode the meaning of those clues. For instance, one sender in Study 3 pointed out that the presence of religious symbolism in the target song greatly annoyed him; he claimed to have a tremendous disdain for religion and stressed his belief that religion had no place in popular music. In fact, he held very deep and strong religious convictions. His friends reported that this passage alerted them to the falseness of the communication and signaled the actual location of the target song. Strangers, who did not have this shared piece of common knowledge, were not clued in by it.

Finally, because we matched one stranger audience with the friends audience on several important variables (specifically age, method of payment for participation, and academic status), we are also able to conclude that the major path to the differential detection abilities on the multiple message trial was not exclusively a result of these characteristics of the observer audiences. The two groups of strangers did not differ from one another in terms of their performance, despite the fact that one audience was matched to the friends audience. If any of these audience variables were responsible for the results we obtained, we would have expected the friends and the matched strangers to perform more similarly, and they did not.

GENERAL DISCUSSION

Taken together, the results from the current studies demonstrate that behavior is not always taken at face value. Under certain circumstances, people are capable of conveying accurate information (or possibly an accurate self-image) even when they are constrained. Whether they are successful depends on several factors. First, the communicator must be motivated to convey an accurate image of his or her true personality, opinion, or information. As we have demonstrated, essay authors in the assigned position paradigm are not particularly motivated to present their true personalities in the essays they write, and readers of the essays fall prey to the correspondence bias. Similarly, numerous studies in the lie detection literature demonstrate that when people seek to conceal their true personalities or opinions, they can be quite successful (see Ekman, 1985).

Second, the audiences witness to the behavior must have some reason to doubt the veracity of the presentation (Ekman & Friesen, 1969; Zuckerman et al., 1984). Most of the time, audiences operate in a conspiracy mode (Kraut, 1978) in which they not only honor a person's claim to a specific role but also actively help support that claim (see also Goffman, 1959). As a result, and assuming that their criteria for assessing the validity of a performance are met (e.g., our readers' assessments of essay credibility and extremity in Study 1), observers will take behavior at face value and infer a corresponding disposition. Sometimes, however, audiences become suspicious and begin questioning the communicator's intent. In some cases, expectations about the person or situation can arouse general suspicions on the part of an observer. It is the wealth of the widower, for example, that leads us to question the intentions of his pursuer (Fein, Hilton, & Miller, in press). Similarly, when observers are told by an experimenter that a sender may be lying, they will be suspicious (e.g., Toris & DePaulo, 1984). In other cases, suspicion may be aroused by some aspect of the display that seems unexpected, untoward, or unnatural (see Ekman & Friesen, 1969, 1974; Krauss, 1981; Zuckerman, Amidon, Bishop, & Pomeronitz, 1982). If observers suspect that a communicator is being untruthful, they may move into an agonistic or sparring mode (Kraut, 1978) and begin searching for diagnostic information in the presentation (Fein et al., in press; Hilton, Miller, Fein, & Darley, 1989; Pyszczynski & Greenberg, 1981). As we have shown, however, unless the cues observers seek out are unequivocal or particularly informative, they are not likely to arrive at an accurate impression of the communicator.

We have pointed to the existence of a class of situations, multiple audience problems, in which individuals attempt to convey different messages to different, but simultaneously present, audiences. In one large subclass of these situations, communicators want the larger audience to accept the surface content of their messages, while desiring that some targeted subset of that audience receives a more complex message that is hidden from the larger audience. For this complex task to be carried out successfully, we suggest that the communicator and the targeted subaudience must have some shared knowledge that creates the possibility of a communication strategy. By referencing this shared knowledge, constrained communicators can both signal the presence of constraint, and reveal to that subaudience the true message, while masking its content and even its existence from the larger audience. Friends in Studies 2 and 3 detected and correctly interpreted the clues that communicators put into their messages and arrived at an accurate appraisal of the underlying message. Strangers who did not possess the requisite shared knowledge, on the other hand, took the communicators' messages at face value.

At a more general level, this research is consistent with a growing research tradition that attempts to take into account...
the strategic nature of our social lives in our experimentation. Impression management theory (Schlenker, 1985; Tedeschi, 1981), for example, emphasizes people’s motives to convey positive impressions of themselves to others. Likewise, researchers interested in strategic self-presentation stress the instrumental component of presenting a desired self-image (Jones & Pittman, 1982). Similarly, recent conceptions of intelligence have taken multidimensional approaches to understanding intelligence that frequently include concerns of a strategic nature (e.g., Cantor & Kihlstrom, 1987; Sternberg, 1986). Still others have argued that human communication is itself fundamentally instrumental and that the goals of the participants in the communication game must be taken into account in order to reach a comprehensive understanding of social behavior (Higgins, 1981; Kraut & Higgins, 1984).

Finally, the results presented here suggest that a communication perspective may help us to understand certain aspects of the social perception process. Our typical experimental methods for studying social perception tend to eliminate the possibilities of developing shared understandings of their task between observers and targets. In the everyday world, however, those interactions are often possible, and as Goffman (1959, 1961, 1969) has noted, both actors and observers are capable of strategic interaction behaviors that facilitate decoding quite complex social messages. We have demonstrated that the existence of shared knowledge with a targeted subaudience, social actors spontaneously generate and use complicated communication strategies to convey mixed messages, and the targeted observers can detect and interpret those strategies and arrive at accurate perceptions of the communicators’ messages and dispositions. If conditions that allow for these possibilities are included in future social perception research, a more complex empirical picture may emerge of the conditions under which a person’s true opinions or attitudes do or do not become apparent to the perceiver.

References


MULTIPLE AUDIENCE PROBLEM


Received June 27, 1984
Revision received November 15, 1989
Accepted December 5, 1989 •