

The Impact of Anonymity on Responses to Sensitive Questions¹

ANTHONY D. ONG

University of Southern California

DAVID J. WEISS²

California State University, Los Angeles

To estimate frequencies of behaviors not carried out in public view, researchers generally must rely on self-report data. We explored 2 factors expected to influence the decision to reveal: (a) privacy (anonymity vs. confidentiality) and (b) normalization (providing information so that a behavior is reputedly commonplace or rare). We administered a questionnaire to 155 undergraduates. For 79 respondents, we had corroborative information regarding a negative behavior: cheating. The privacy variable had an enormous impact; of those who had cheated, 25% acknowledged having done so under confidentiality, but 74% admitted the behavior under anonymity. Normalization had no effect. There were also dramatic differences between anonymity and confidentiality on some of our other questions, for which we did not have validation.

Perhaps the most fundamental issue in survey research is that posed in the title of Hyman's (1944) paper, "Do they tell the truth?" The quality of questionnaire data is especially worrisome when the behavior in question is likely to be viewed as embarrassing or sensitive (Armacost, Hosseini, Morris, & Rehbein, 1991; Becker & Bakal, 1970; Bradburn, Sudman, Blair, & Stocking, 1978). Nevertheless, the questionnaire is the primary tool for estimating the frequencies of behaviors not carried out in public view (Sudman & Bradburn, 1974).

Although the validity of the self-report is often called into question, there is little alternative when the researcher seeks information about events in a respondent's past, because only the respondent is likely to have ready access to such information. For questions that seem innocuous, such as those concerning employment status or area of residence, expectations of truthful responses are usually high. For others, such as inquiries about sexual behavior, incidents of dishonesty, or dramatic experiences such as abuse, the nature of the question

¹This report is based on a thesis submitted by the first author, and supervised by the second, to the Department of Psychology, California State University, Los Angeles, in partial fulfillment of the requirements for the M.A. degree. We wish to thank James H. Hogge for comments on a draft of this article.

²Correspondence concerning this article should be addressed to David J. Weiss, Department of Psychology, California State University, Los Angeles, 5151 State University Drive, Los Angeles, CA 90032. e-mail: dweiss@calstatela.edu.

suggests that some respondents will not respond accurately (Bradburn & Sudman, 1979).

For the researcher seeking to investigate methods designed to elicit truthful responses, the principal experimental problem is that a respondent's history is unknown. It is difficult, therefore, to determine which method is most effective and practical at allowing respondents to reveal potentially sensitive behaviors, because no verification is available.

Given that validation is not practical, researchers in applied settings have relied on two plausible assumptions as a basis for methodological comparisons. The first is the key assumption that lies occur in a predictable direction. Greater reported frequencies of stigmatized behaviors are taken to indicate greater response accuracy (Bradburn et al., 1978). This assumption allows the researcher to conclude that one response method is superior to another. The assumption, although compelling, is not guaranteed to be correct. Consider a subculture in which antisocial acts are considered a rite of passage (Catania, Chitwood, Gibson, & Coates, 1990). Members might falsely claim to have committed acts that others would strain to conceal.

The second assumption is that the researcher knows which behaviors are stigmatized, thus supporting the use of question sensitivity as a substantive factor. A standard design seeks to show that a methodological improvement yields more revelations for sensitive items, but has no effect on innocuous ones (Linden & Weiss, 1994). There are two bases to this assumption: that the researcher knows which answer the respondent will find embarrassing (Tourangeau & Smith, 1996), and that respondents will agree on what is sensitive (Catania et al., 1990; Lee & Renzetti, 1990).

These assumptions may be correct in some cases and not in others. Their applicability may explain why studies have not always shown the methodological differences anticipated by researchers. For example, a recent meta-analysis (Singer, Von Thurn, & Miller, 1995) summarizes the inconsistent effects of confidentiality assurances in the literature. Some studies have found that individuals disclose more sensitive information under anonymous conditions (Hill, Dill, & Davenport, 1988; Klein & Chevront, 1990; Werch, 1990); but others have found the opposite, with respondents disclosing more under confidential conditions (Boruch & Cecil, 1979; Campbell & Waters, 1990; Esposito, Agard, & Rosnow, 1984). Still other studies have found no difference between participants' responses under anonymity or confidentiality (Brink, 1995; Fuller, 1974; King, 1970).

Null findings may occur if the behavior explored is either not stigmatized or not common within the sample. If the behavior is not stigmatized, there is no motivation for respondents to lie, so no difference between methods emerges. Fidler and Kleinknecht (1977) reported that anonymity generated more acknowledgment of masturbation than did confidentiality; while Begin, Boivin, and

Bellerose (1979) reported no such difference for the same behavior. The percentages acknowledging masturbation were much higher for the latter study. If we assume that almost all students masturbate (Person, Terestman, Myers, Goldberg, & Salvador, 1989), a possible interpretation of this disparity is that masturbation was stigmatized for Fidler and Kleinknecht's respondents, but not for those of Begin et al.

For rare behaviors, any comparison of methods for eliciting sensitive information is likely to be weak. The problem is lack of statistical power. For example, when Tourangeau and Smith (1996) compared three methods for eliciting acknowledgment of marijuana use, they found differences for lifetime use but not for past year or past month use. Of course, the average proportions reported were much higher (66% in the most favorable condition) for lifetime use. Cocaine is used much less frequently, with only 20% reported lifetime use in the most favorable condition. The method variation yielded no significant differences at all for cocaine use. Low prevalence rates make it difficult to discriminate among conditions, as differences between response proportions will inevitably be small, even if one condition encourages more revelations than does another. Uncontrolled variation in prevalence rates may produce a literature characterized by inconsistency.

The survey methodologist can hope to address the power issue by judiciously choosing to explore common (but stigmatized) behaviors. In a laboratory study, the prevalence rate problem may be addressed more directly. Validation information affords the possibility of bypassing the prevalence rate problem through conditionalizing. By examining responses for the subset of people for whom the accurate response to a question about a stigmatized behavior is known to be "Yes," the researcher obtains an effective base rate of 100%. Conditionalization thereby increases the statistical power to distinguish among methods.

Methodological comparisons in the self-report domain often contrast modes of data collection. For example, surveys may be administered face to face or via telephone. These modes differ in several properties (Catania et al., 1990). The properties are likely to affect the respondent's decision to answer truthfully. Examples include the credibility of the interviewer's presentation, the amount of time allowed for the respondent to reflect on an answer, and the degree to which the process is interactive.

When potentially sensitive matters are explored, the perceived privacy inherent in the data-collection mode may be the most important property. People convey a particular impression of themselves to others so as to minimize their own discomfort. This process has been labeled *self-presentation* (Goffman, 1959; Jones & Pittman, 1982). Responding to a survey can similarly be viewed as a process of interpersonal communication, so one might expect self-presentation concerns to bias the responses (Catania et al., 1990; Sudman & Bradburn, 1974).

Variables that have been shown to affect self-presentation may be expected to play a role in the respondent's decision to reveal personal information on a

survey. The experimental paradigm commonly used to study self-presentation in a social context involves manipulating whether opinions are expressed publicly or privately (Baumeister, 1982; Baumeister, Hutton, & Tice, 1989). If the likelihood of being identified causes individuals to respond differently, it is assumed that the change reflects concern about what these reports communicate. Conversely, anonymity reduces concern with self-presentation because one's actions are no longer monitored by others (DePaulo, 1992; Patterson, 1991; Schlenker & Weigold, 1990).

In the survey literature, privacy is manipulated by contrasting the two most frequently used guarantees: confidentiality and anonymity. Researchers who administer questionnaires inquiring about past behaviors routinely guarantee participants confidentiality to enhance response validity (American Psychological Association, 1996). *Confidentiality* refers to an implicit or explicit agreement that no traceable record of the participant's data will be disclosed (Nation, 1997); only the researcher knows the response. *Anonymity*, on the other hand, refers to a condition in which the researcher does not know the identity of the respondent.

The specific way in which a sensitive question is phrased may be another important property of the data-collection mode. Phrasing can normalize behaviors by influencing the extent to which a behavior is seen as abnormal and thereby inappropriate to disclose (Catania et al., 1996). Normalization may be implemented by providing information about how others have responded (Churchill, 1987; Clark & Tiffi, 1966; Fowler, 1985). The assumption is that a behavior of dubious propriety may be acknowledged to have taken place if it is thought to be a commonly occurring one. This approach has the potential advantage that it may be applied routinely and systematically without regard to specific items.

In the present study, we explore how expectations of privacy and normalization of the behavior affect the disclosure rate among people whom we know to have cheated. Cheating was chosen as the stigmatized behavior for two reasons. First, cheating is known to be common among college students; the prevalence rate has been reported to be approximately 75% across several large-scale surveys (McCabe & Bowers, 1994). The high base rate made it likely that we would be able to obtain a considerable number of participants for whom the accurate answer to a question about cheating would be positive. Second, we were able to assess its occurrence in our laboratory, thus validating the responses. Prior to our survey, we had surreptitiously arranged a situation in which cheating was possible. We knew which of our participants had, in fact, cheated, but they were unaware of our knowledge. Proportions of "Yes" responses for the subset may be contrasted with those for the entire sample; the latter analysis mimics the usual methodological study. We expected the effect of our independent variables to be more pronounced within the subset of cheaters, since there is no dilution attributable to people for whom a "No" response is a truthful one.

Our methods were designed to mimic those used in a typical questionnaire survey. We asked some questions we thought were highly personal and others that seemed innocuous. In the anonymous condition, we kept track of participants with private identification numbers so that we could truthfully assure them that no one, including the experimenter, could learn how they responded. In contrast, the confidential condition was designed to resemble studies in which the experimenter knows the participant's face and name. Explicit assurance was given that no individual responses would ever be disclosed by the research team. We hypothesize that for sensitive questions, there will be more disclosures when respondents are promised anonymity than when they are promised confidentiality.

To explore the effect of normative information, we provided bogus reports on how previous respondents had answered our survey questions. We included either a high or low percentage for each item. We hypothesize that there will be more disclosures when the sensitive behavior in question is one that the respondent believes most people would report.

Along with the question about cheating, we asked questions about other behaviors for which we had no validation. We thought some of the questions would be sensitive, whereas others were anticipated to be innocuous. In everyday survey work, of course, validation evidence is not available, so we hope to extrapolate the effects of privacy and normalization to those questions.

Method

Participants

The survey participants were 155 students enrolled in five Introductory Psychology courses at a large, urban university in California. Participants were surveyed in classes assigned randomly to one of four experimental conditions by random permutation, with one class shifted afterward to achieve near equality of group sizes. There were 102 (66%) females and 53 (34%) males in the sample. All participants received credit toward fulfillment of a course requirement.

Apparatus

The questionnaire was entitled "Student Behavior Questionnaire" and consisted of 15 items. Three of the questions were adapted from Linden and Weiss (1994), and the remainder were constructed for this study. The response options on the questionnaire were "Yes" and "No" (see Appendix for the instructions and Table 1 for the questionnaire).

Design

A 2×2 factorial design was employed, with privacy and normalization as the two independent variables. There were two privacy levels: anonymous (*i.e.*,

Table 1

Proportions of "Yes" Responses for Anonymity and Confidentiality Groups

	Anonym- ity	Confiden- tiality	<i>F</i>
1. Are you between 18 and 25 years old?	.83	.76	1.21
2. Have you ever, even once, been charged by a policeman for driving under the influence of alcohol?	.25	.15	1.38
3. Do you or anyone in your house own a gun?	.21	.28	<1
4. Have you ever tried marijuana?	.22	.31	<1
5. Do you speak more than one language reasonably well?	.83	.76	1.55
6. Were you born in California?	.31	.52	9.76*
7. In the past year, have you ever, even once, used unapproved material on an exam, quiz, or any other form of test?	.47	.13	24.35*
(7. Peekers only)	.74	.25	13.73*
8. Have you observed anyone engaging in physical abuse of campus property within the last year?	.11	.17	<1
9. Are you an only child?	.09	.09	<1
10. Have you ever, even once, taken something from a store without paying for it?	.74	.40	20.65*
11. Are you carrying a study load of more than 12 units this quarter?	.69	.60	2.22
12. As an undergraduate, did you ever collaborate on an assignment when individual work was required?	.57	.27	17.12*
13. In the past month, have you ever engaged in masturbation?	.61	.24	23.29*
14. Have you ever, even once, turned in work done by someone else?	.53	.27	10.86*

Note. Significance assessed using Bonferroni's adjustment.

* $p < .003$.

respondents were explicitly told to avoid indicating their names) and confidential (i.e., respondents were explicitly asked for name and Social Security number). The normalization variable was implemented by including in the questionnaire a bogus distribution of Yes and No responses from "previous students" at the participants' university. Two forms were used. In the commonplace form, most of the items had high percentages (78% to 96%) of Yes responses furnished. In the rare form, the complements of these percentages were given, so that most items had low percentages (4% to 22%) attached to them. For Items 1, 6, and 9, moderate percentages (51% to 58%) were presented for both forms in order to make the pattern of percentages appear less extreme. Participants were run in class sections, with each section randomly assigned to one of four conditions as follows: (a) anonymous, commonplace ($n = 33$); (b) confidential, commonplace ($n = 53$); (c) anonymous, rare ($n = 39$); and (d) confidential, rare ($n = 30$).

Procedure

Students participated in a two-phase experimental task, separated in time and carried out by different investigators. In the first phase, 79 students from five sections of Introductory Psychology volunteered to compete individually in a study of vocabulary aptitude. In the second phase, the same students, along with their classmates (total $N = 155$), were asked to complete a questionnaire in the classroom. From the participant's perspective, the first phase had no relation to the second. From the experimenter's perspective, however, the two phases were connected. Specifically, response validation in the questionnaire phase was supported by a manipulation introduced in Phase 1.

Phase 1

In the first phase, 79 students were tested individually in a small laboratory room. Each participant was told that he or she would take a vocabulary test that consisted of 20 multiple-choice words that appeared on a computer screen. The keyboard and monitor were placed on a large table. Before starting the test, the student was required to enter a code consisting of the last four digits of his or her Social Security number. Participants were informed by the experimenter that they were free to preview, review, or skip items, as well as to change their responses by moving the arrow keys on the keyboard. In addition, the experimenter advised participants that the test was timed by the computer for 10 min and that they were free to leave if they finished early. Each participant completed a sample problem before beginning the actual test. The researcher made sure that the participant was comfortable using the computer before leaving the room.

The first phase was designed to make it easy for the student to engage in a specific kind of cheating. Cheating was operationalized as having consulted a small dictionary unobtrusively placed among a row of six books on the table, 3 feet (1.08 m) to the left of the monitor. A bookend was inserted between two designated pages in the dictionary. It was therefore possible to tell if a participant had cheated by observing if the dictionary had either been moved or put back with the bookend between the wrong pages. Participants were told to ignore the books the researcher had carelessly left in the room. They were also told not to use any outside material for the test. There was no explicit mention of the dictionary.

The 20-item multiple-choice vocabulary test consisted of 5 relatively easy words and 15 extremely difficult words. Each test incorporated a random selection of words drawn from a larger pool of 60 vocabulary items. In addition to receiving extra credit, students were advised that they would win \$25 if they achieved a score of 17/20 (85%) or better on the test. In order to ensure that the words would be difficult enough to inspire cheating, we made up the last three words. No one obtained a score greater than 14 of the 79 students who took the test, 55 (69%) cheated by consulting the dictionary.

Phase 2

At least 3 weeks later, the students who had been in Phase 1 participated in the second phase of the study along with their classmates. A second experimenter went to each class and invited all students to participate in a survey. Instructions to students in this phase were similar to those provided by Linden and Weiss (1994). Participants were asked to complete a questionnaire and were told that some of the questions would be "of a sensitive nature." The students were instructed that they were free to leave without providing a reason (and without forfeiting extra credit) if at any time they became uncomfortable. Participants were advised that the questionnaire was part of a follow-up survey to one conducted at the same university a few quarters earlier.

Participants were told that their responses would be kept either completely anonymous or confidential, depending on the condition. The experimenter instructed students in the anonymity groups to place the completed questionnaire in a cardboard box placed at the front of the room. Participants in these groups were asked to provide the last five digits of their Social Security numbers on the questionnaire. The questionnaire was not to be handed directly to the experimenter who stayed in the classroom. In contrast, participants in the confidentiality groups were asked to put their full names and Social Security numbers on the questionnaire and to hand the questionnaire to the experimenter when it was completed. The Social Security numbers, whether whole or fragments, were sufficient for the researchers to unambiguously identify those who had participated in the first phase.

Response Validation

Information about who had cheated in the first phase was used to validate individual responses to survey Question 7 (“In the past year, have you ever, even once, used unapproved material on an exam, quiz, or any other form of test?”). Of interest in the questionnaire phase were the effects of privacy and normalization on the responses of those individuals who cheated. Additionally, evidence was sought for the overall effects of anonymity and normalization on the 13 other survey items for which there was no validation information.

Postexperimental Session

At the conclusion of the second phase, participants were debriefed in the classroom setting. During the debriefing, participants were informed of the purpose of the study, and the deception was explained. The word “cheating” was not used to describe consultation of the dictionary. Because students served as participants to earn class credit and to learn about psychological research, the experimenter also discussed issues related to the use of human participants in psychological research.

Results

The question of central interest concerns the accuracy of the responses obtained with the privacy and normalization manipulations. Privacy had an enormous impact, while normalization had virtually none. To assess accuracy, we confine our attention to Question 7, the question for which we had validation information. The proportion of “Yes” responses by “peekers” is shown in Figure 1, with a factorial plot displaying both variables. It may be seen that 74% of those who cheated admitted the behavior under anonymity, but only 25% of respondents did so with the guarantee of confidentiality. ANOVA provides statistical confirmation of a privacy difference, $F(1, 51) = 13.73, p < .003$.

Also apparent in Figure 1 is the lack of effect of the normalization variable (commonplace, as opposed to rare). The proportions of Yes responses were similar (.50 vs. .48), and this difference was not significant, $F(1, 151) < 1$. There was also no interaction between privacy and normalization, $F(1, 151) < 1$. The responses also furnished evidence on the issue of whether the direction of inaccurate answers was predictable. None of the individuals who did not peek answered “Yes” to the cheating question.

The effect of the conditionalization allowed by the validation is apparent when we look at the overall proportions of Yes responses to the cheating question. Here, our 55 peekers and 24 nonpeekers were mingled with 76 other individuals for whom we had no validation information. There were thus 155

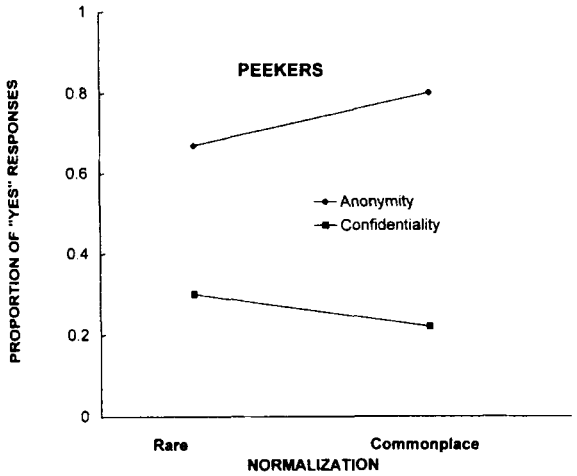


Figure 1. Proportion of "Yes" responses to Question 7 for peekers under anonymity versus confidentiality and for commonplace versus rare.

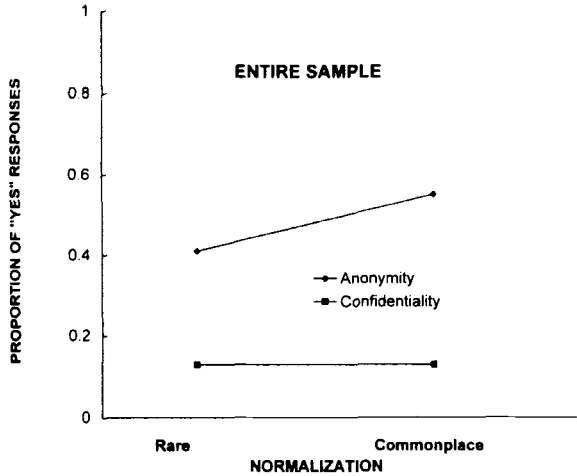


Figure 2. Proportion of "Yes" responses to Question 7 for entire sample under anonymity versus confidentiality and for commonplace versus rare.

respondents in the sample, whose response proportions we plot factorially in Figure 2 as the usual survey report would. Anonymity induced 47% of the respondents to disclose cheating, while confidentiality brought out only 13% "Yes" responses. The advantage of anonymity, while still pronounced, is muted. With conditionalization, anonymity elicits 49% more disclosures; without, anonymity

elicits 34% more disclosures. In either case, though, anonymity is dramatically more effective.

The proportion of "Yes" responses to our other questions for all participants, partitioned by the privacy variable, is given in Table 1. Because the analysis comprised multiple independent tests, a Bonferroni adjustment procedure with accordingly reduced significance level was employed for all questions. With 15 tests, we used a significance level of $.05/15$ (.003) for each comparison. There was a significant main effect of privacy for six items (Items 6, 7, 10, 12, and 13), as well as for the test of peekers on Item 7. Compared to respondents in the confidentiality group, more anonymous respondents acknowledged taking something from a store at least once without paying for it, $F(1, 151) = 20.65, p < .0001$; collaborating on an assignment when individual work was required, $F(1, 151) = 17.12, p < .0001$; and engaging in masturbation in the past month, $F(1, 151) = 23.39, p < .0001$. Conversely, more confidential respondents than anonymous respondents stated that they were born in California, $F(1, 151) = 9.76, p < .003$.

There was no significant main effect of normalization for any of the items (all questions except Items 1, 6, and 9 were tested). Similarly, there were no interactions between privacy and normalization.

Our preconceptions of question sensitivity were not supported by the responses. The predictions made by the two authors were not wholly in accord. The disagreement emphasizes the subjective nature of this designation. At least one of us anticipated that Questions 2, 3, 4, 7, 10, 12, 13, and 14 would be sensitive. Our operational definition of sensitivity is less subjective; a sensitive question yields a significantly different proportion of "Yes" responses for anonymity as opposed to confidentiality. According to this criterion, six of these items (Items 6, 7, 10, 12, 13, and 14) were sensitive.

"Were you born in California?" (Question 6) emerged as a sensitive question, with "Yes" being the socially desirable response. Only 31% of the anonymous respondents reported being born in California. Under confidentiality, 52% said California was their birthplace, a significant difference, $F(1, 151) = 9.76, p < .003$. Neither author anticipated sensitivity here. The outcome may reflect the strong anti-immigrant sentiments expressed in recent California elections, or it may reflect the fact that tuition is higher for non-residents. We have no evidence to support either speculation.

Discussion

The results show that anonymity and confidentiality should not be seen as interchangeable. Anonymity induced many more revelations. When the contrast was experimentally magnified with the conditionalization allowed by validation, the difference in the proportions of "Yes" responses to an inquiry about a socially

disapproved behavior was 49%. Even when the privacy guarantees were compared without conditionalization, as they would be in ordinary survey research, anonymity yielded 34% more revelations of cheating.

We are assuming that those who did not peek in our study, all of whom denied having cheated, in fact never cheated during the previous year. Our validation is incomplete because this assumption is not verifiable. The fact that 31% did not acknowledge cheating is consistent with the national norms reported by McCabe and Bowers (1994). There is a minority of students who simply do not cheat.

The failure of the normalization variable to affect the responses may be the result of a weak implementation. Our "peekers" were no more likely to disclose cheating when the question addressed a behavior that "96% of students in a prior survey" reported than when it addressed a behavior that "4% of students in a prior survey" reported. In retrospect, it appears that assigning most survey questions extreme values may have been a design error, in that the percentages were not credible.

The other possibility is that the participants believed the cited statistics, but that normalization implemented in this way simply does not affect revelation. This interpretation suggests that students maintain their view of their own behavior independent of how they may think others behave. It might have been a good idea to assess formally the effectiveness of the manipulation during the debriefing. We could have asked respondents to estimate the proportion of their peers who would acknowledge cheating. The idea that perception of prevalence may be altered without affecting personal disclosure is consistent with an observation by Levin, Schnittjer, and Thee (1988). In a study focused on how information was presented, they found that the framing of cheating statements affected students' ratings of the incidence of cheating, but did not affect their expressed likelihood of personally cheating.

Other means of implementing a normalizing variable might have been more effective. For example, wording items to make stigmatized behaviors appear more normative without citing specific numbers has been found to increase disclosure (Catania et al., 1996). Recent research by Teigen (1998) has suggested that numerical probabilities are not well understood, and that verbal probabilistic terms can have more consistent effects.

The present findings raise several issues of concern for survey researchers. First, the term "sensitive question" is often used in the literature as if it were self-explanatory (Lee & Renzetti, 1990; Tourangeau & Smith, 1996). Investigators have attempted to operationalize sensitivity in terms of question threat or to identify questions that concern contranormative behaviors (Lee & Renzetti, 1990; Locander, Sudman, & Bradburn, 1976). Asking respondents which questions are sensitive (Catania et al., 1990) substitutes the collective insight of the participants for that of the researchers, but it is not obvious why one respondent should know

what another will choose to conceal. Even after seeing our results, we are not able to find a basis for deciding whether a novel question will be sensitive. We propose to treat sensitivity purely as an outcome.

Our proposed definition of a sensitive question is that it is one for which a privacy manipulation yields a difference in response proportions. This requires that the investigation have sufficient power. If a question concerns a rare, albeit disapproved behavior, such as murder or sexual abuse of a child, the low prevalence rate makes it unlikely that we can observe an impact of privacy on revelation. Of course, sensitivity must be defined for a specific population of respondents; prevalence rates for many behaviors may be expected to vary across populations. If the prevalence rate for the behavior is low, large samples (e.g., Turner et al., 1998) may be needed before sensitivity can be determined.

The unexpected result showing that California birth was exaggerated under confidentiality highlights the difficulty of designating question sensitivity in advance. The result also supports the insight of Bradburn et al. (1978), that socially desirable distortions may take the form of either underreporting or overreporting. We did not anticipate the direction, of course, but if we had, we might have asked the question in a different format. "Were you born outside of California?" elicits the same information, and presumably would generate higher proportions of "Yes" responses for the anonymity condition. This example illustrates the principle that one cannot simply assume that more "Yes" responses connote more true responses. A difference in "Yes" proportions across conditions implies that a question is sensitive, and that some respondents are lying. Our evidence suggests that the anonymous responses are more likely to be truthful, no matter which response predominates.

Another challenge arises when the question is one that may elicit lies in either direction. Such questions as "Are you a virgin?" or "Did you have intercourse within the last week?" may generate confounded response proportions. We do not know how to state a general rule for identifying such questions. The key element in the problem is group heterogeneity. Some respondents may feel enhanced self-presentation with an incorrect "Yes," others with an incorrect "No." Without validation, there is no way to resolve this matter. We have assumed that our questions generate lies in a predictable direction, but we have verified this assumption only for Question 7.

Anonymity is so predominant that it may obscure other effects sought by survey methodologists. For example, Linden and Weiss (1994) explored the hypothesis that the random-response method of data collection (Warner, 1965) yields more revelations of sensitive information than does a direct questioning method. The theoretical basis of the method is that feelings of privacy are enhanced by the inability of the researcher to connect the individual respondent with his or her answer. Linden and Weiss's (1994) respondents were guaranteed anonymity in both conditions. They found that, contrary to expectation, the random-responding

method offered no advantage over the simpler direct-questioning approach. Our perspective is that, with anonymity guaranteed, there was no room for the random-response method to show an advantage because maximal revelation had already been achieved.

Anonymity is clearly very powerful. Even so, 26% of those who "peeked" in the present study responded "No" with the strong guarantee of privacy. Perhaps they simply did not want to admit it, even if no one else would know. It is also possible that they did not view their behavior as fitting the definition provided in the cheating question or that they forgot the behavior. The existence of this logical possibility implies that our use of the term "lie" when someone responds inaccurately may be overly judgmental.

The present results may not generalize beyond self-administered questionnaires. In some field settings, anonymity may not be attainable. For telephone surveys, the respondent knows that the researcher has his or her telephone number. In such cases, confidentiality is as much privacy as is available. Here, characteristics of the interviewer (human or computer), gender of the interviewer, and personal style may play a major role (Catania et al., 1996; Turner et al., 1998). Moreover, there may be situations in which privacy guarantees backfire. Singer, Hippler, and Schwarz (1992) suggested that, under some circumstances, elaborate assurances of confidentiality can frighten participants into refusing to respond.

All too often, applied researchers have glossed over the distinction between confidentiality and anonymity. For understanding sensitive behaviors that have important consequences, such as risky sexual practices, this distinction may be critical. Lies induced by weak privacy guarantees are likely to occur across experimental conditions, thereby making it difficult to see the value of a useful intervention.

References

- American Psychological Association. (1996). Rules and procedures: June 1, 1996. *American Psychologist*, *51*, 529-548.
- Armocost, R. L., Hosseini, J. C., Morris, S. A., & Rehbein, K. A. (1991). An empirical comparison of direct questioning, scenario, and random response methods for obtaining sensitive business information. *Decision Sciences*, *22*, 1073-1090.
- Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3-26.
- Baumeister, R. F., Hutton, D. G., & Tice, D. M. (1989). Cognitive processes during deliberate self-presentation: How self-presenters alter and misinterpret the behavior of their interaction partners. *Journal of Experimental Social Psychology*, *25*, 59-78.

- Becker, G., & Bakal, D. (1970). Subject anonymity and motivational distortion in self-report data. *Journal of Clinical Psychology*, **26**, 207-209.
- Begin, G., Boivin, M., & Bellerose, J. (1979). Sensitive data collection through the random response technique: Some improvements. *Journal of Psychology*, **101**, 53-65.
- Boruch, R. F., & Cecil, J. S. (1979). *Assuring the confidentiality of social research data*. Philadelphia, PA: University of Pennsylvania Press.
- Bradburn, N. M., & Sudman, S. (1979). *Improving interview method and questionnaire design*. San Francisco, CA: Jossey-Bass.
- Bradburn, N., Sudman, S., Blair, E., & Stocking, C. (1978). Question threat and response bias. *Public Opinion Quarterly*, **42**, 221-234.
- Brink, T. L. (1995). Sexual behavior and telling the truth on questionnaires. *Psychological Reports*, **76**, 218.
- Campbell, M. J., & Waters, W. E. (1990). Does anonymity increase response rate in postal questionnaire surveys about sensitive subjects? A randomized trial. *Journal of Epidemiology and Community Health*, **44**, 75-76.
- Catania, J. A., Binson, D., Canchola, J., Pollack, L. M., Hauck, W., & Coates, T. J. (1996). Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behavior. *Public Opinion Quarterly*, **60**, 345-375.
- Catania, J., Chitwood, D. D., Gibson, D. R., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin*, **108**, 339-362.
- Churchill, G. F. (1987). *Marketing research* (4th ed.). Chicago, IL: Dryden.
- Clark, J. P., & Tiffit, L. L. (1966). Polygraph and interview validation of self-reported deviant behavior. *American Sociological Review*, **31**, 516-523.
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, **111**, 203-243.
- Esposito, J. L., Agard, E., & Rosnow, R. L. (1984). Can confidentiality of data pay off? *Personality and Individual Differences*, **5**, 477-480.
- Fidler, D. S., & Kleinknecht, R. E. (1977). Random responding versus direct questioning: Two data-collection methods for sensitive information. *Psychological Bulletin*, **84**, 1045-1049.
- Fowler, F. J., Jr. (1985). *Survey research methods*. Beverly Hills, CA: Sage.
- Fuller, C. (1974). Effect of anonymity on return rate and response bias in a mail survey. *Journal of Applied Psychology*, **59**, 292-296.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday/Anchor Books.
- Hill, P. C., Dill, C. A., & Davenport, E. C. (1988). A reexamination of the bogus pipeline. *Educational and Psychological Measurement*, **48**, 587-601.
- Hyman, H. (1944). Do they tell the truth? *Public Opinion Quarterly*, **8**, 557-559.

- Jones, E. E., & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1. pp. 231-262). Hillsdale, NJ: Lawrence Erlbaum.
- King, F. W. (1970). Anonymous versus identifiable questionnaires in drug usage surveys. *American Psychologist*, *25*, 982-985.
- Klein, K., & Chevront, B. (1990). The subject-experimenter contract: A reexamination of subject pool contamination. *Teaching of Psychology*, *17*, 166-169.
- Lee, R. M., & Renzetti, C. M. (1990). The problem of researching sensitive topics. *American Behavioral Scientist*, *33*, 510-528.
- Levin, I. P., Schnitjjer, S. K., & Thee, S. L. (1988). Information framing effects in social and personal decision. *Journal of Experimental Social Psychology*, *24*, 520-529.
- Linden, L. E., & Weiss, D. J. (1994). An empirical assessment of the random response method of sensitive data collection. *Journal of Social Behavior and Personality*, *9*, 823-836.
- Locander, W., Sudman, S., & Bradburn, N. (1976). An investigation of interview method, threat and response distortion. *Journal of the American Statistical Association*, *71*, 269-275.
- McCabe, D. L., & Bowers, W. J. (1994). Academic dishonesty among males in college: A 30-year perspective. *Journal of College Student Development*, *35*, 5-10.
- Nation, J. R. (1997). *Research methods*. Upper Saddle River, NJ: Prentice-Hall.
- Patterson, M. L. (1991). Functions of nonverbal behavior in interpersonal interaction. In R. S. Feldman & B. Rime (Eds.), *Fundamentals of nonverbal behavior* (pp. 458-495). Cambridge, UK: Cambridge University Press.
- Person, E. S., Terestman, N., Myers, W. A., Goldberg, E. L., & Salvador, C. (1989). Gender differences in sexual behaviors and fantasies in a college population. *Journal of Sex and Marital Therapy*, *15*, 187-198.
- Schlenker, B. R., & Weigold, M. F. (1990). Self-consciousness and self-presentation: Being autonomous versus appearing autonomous. *Journal of Personality and Social Psychology*, *59*, 820-828.
- Singer, E., Hippler, H., & Schwarz, N. (1992). Confidentiality assurances in surveys: Reassurance or threat? *International Journal of Public Opinion Research*, *4*, 256-268.
- Singer, E., Von Thurn, D. R., & Miller, E. R. (1995). Confidentiality assurances and response: A quantitative review of the experimental literature. *Public Opinion Quarterly*, *59*, 66-77.
- Sudman, S., & Bradburn, N. (1974). *Response effects in surveys: A review and synthesis*. Chicago, IL: Aldine.
- Teigen, K. H. (1998, November). *A good chance is not a number: Determinants and implications of verbal probabilistic terms*. Paper presented at the meeting of the Judgment/Decision Making Society, Dallas, TX.

- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, **60**, 275-304.
- Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science*, **280**, 867-873.
- Warner, S. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, **60**, 63-69.
- Werch, C. E. (1990). Two procedures to reduce response bias in reports of alcohol consumption. *Journal of Studies on Alcohol*, **51**, 327-330.

Appendix

Instructions for Anonymity Groups

This is a follow-up survey to one conducted here 2 years ago. We are interested in knowing the reported frequency of certain behaviors on college campuses. We are not interested in any individual response, but rather what proportion of individuals engage in certain behaviors. Some of the questions will be of a sensitive nature. The percentage of students responding “Yes” and “No” in the previous survey is given next to each item. Please answer each question honestly. All answers will remain anonymous. To receive credit for participating, please print the last five digits of your Social Security number in the space provided on the questionnaire. Do not put your name on the paper. When you have completed the questionnaire, place it in the cardboard box at the front of the classroom. This way I cannot tell which participant goes with which questionnaire. Are there any questions? [Pause for questions.]

Instructions for Confidentiality Groups

This is a follow-up survey to one conducted here 2 years ago. We are interested in knowing the reported frequency of certain behaviors on college campuses. We are not interested in any individual response, but rather what proportion of individuals engage in certain behaviors. Some of the questions will be of a sensitive nature. The percentage of students responding “Yes” and “No” in the previous survey is given next to each item. Please answer each question honestly. All answers will remain confidential. To receive credit for participating, please print your entire Social Security number and your full name in the space provided on the questionnaire. When you have completed the questionnaire, raise your hand and I will come by to pick up your questionnaire. Are there any questions? [Pause for questions.]