



Published in final edited form as:

J Behav Med. 2002 October ; 25(5): 438–467.

Effects of Anonymity, Gender, and Erotophilia on the Quality of Data Obtained from Self-Reports of Socially Sensitive Behaviors

Lauren E. Durant, Michael P. Carey, and Kerstin E. E. Schroder

Center for Health and Behavior, Syracuse University

Abstract

This study examined the effects of anonymity, gender, and erotophilia on the quality of self-reports of socially sensitive health-related behaviors. A sample of 155 male and 203 female undergraduate students was randomly assigned to an anonymous and a confidential (i.e., non-anonymous) assessment condition. Gender, erotophilia, self-reports (on substance use, sexual behaviors, illegal activity), and perceived item threat were assessed by questionnaire. Data quality was strongly affected by experimental condition and gender. Thus, terminations were more frequent in the confidential condition and among women. In the confidential condition, women were significantly more likely to “prefer not to respond” to sensitive item compared to men. Both female gender and confidential condition were associated with lower frequency reports of sensitive health behaviors, and greater perceived threat of the assessment questions. Self-reported engagement in sensitive behaviors was positively related to both perceived question threat and erotophilia. Path analyses suggest that question threat mediates the effects of anonymity manipulations and gender on data quality (item refusal, termination), and that erotophilia mediates the effects of gender on incidence and frequency self-reports. The results indicate that anonymous assessments as well as male gender are associated with better data quality.

Keywords

Anonymity; self-report; gender; sexual behavior; substance use

Investigation of sexual behavior, substance use, and other socially sensitive behaviors typically relies upon self-report. Because self-report of such behaviors is vulnerable to self-presentation and demand biases, investigators often administer surveys anonymously (i.e., without any identifying information) to promote candid reporting and to minimize bias. However, some studies cannot be conducted anonymously. For example, intervention research and other longitudinal investigations need to link self-report over time. For these studies, investigators typically provide participants with assurances of confidentiality in order to optimize data quality (e.g., Carey, Carey, Maisto, Gordon, & Weinhardt, 2001). The need for data collection under non-anonymous conditions gives rise to the question: Are data collected under confidential conditions equivalent in quality to those collected under anonymous conditions?

Several investigations have provided reassurances that confidential conditions elicit self-reports equivalent to those obtained under anonymous conditions. For example, Malvin and Moskowitz (1983) found few differences between anonymous and confidential self-reports of drug use among 8th and 9th grade students. Zagumny et al. (1996) assessed HIV-related risk behavior of undergraduate students by questionnaire, using four variations of anonymity that requested: (a) no identifying information, (b) name only, (c) demographic information only, or (d) name and demographic information on a page attached to the front of the survey. No differences among groups on any of the dependent measures were found. Werch (1990) investigated whether varying levels of anonymity affect data quality of self-reported alcohol use in a sample of university students. Participants were offered anonymity or confidentiality

with collateral verification. Respondents assigned to the confidential condition were actually more likely to report instances of being drunk in the past year.

In contrast to these findings, several recent studies suggest that anonymity manipulations affects data quality. For example, Ong and Weiss (2000) assessed stigmatized behaviors such as cheating, stealing, and use of marijuana under anonymous and confidential conditions in a sample of 155 undergraduate students and found higher self-reports in anonymous administrations. For example, of those who had actually cheated, only 25% acknowledged having done so in confidential questionnaires whereas 74% admitted the behavior in anonymous questionnaires. In a similar study among 10th grade students, Bjarnason and Adalbjarnardottir (2000) found lower self-reports of substance use (cannabis) in the confidential condition; alcohol use and smoking were reported less often only by girls. Consistent with this finding, O'Malley, Johnston, Bachman, and Schulenberg (2000) found lower self-reports of illicit drug use among 8th grade students in a confidential, non-anonymous condition.

One moderating factor that may account for these apparently discrepant findings, is perceived question sensitivity. The pattern of results reported in the literature suggests that assurances of anonymity increase data quality, but only when the data requested are sensitive in nature. This assumption is supported by results of a recent meta-analysis evaluating the effects of confidentiality on self-report behavior (Singer, Von Thurn, & Miller, 1995). However, due to methodological limitations of the few studies contrasting anonymity with confidentiality, clear conclusions about their effects on data quality and on the possible moderators and mediators of these effects cannot be drawn.

Six methodological factors attenuate our confidence in the findings obtained in previous research. First, most studies testing the effects of confidentiality or anonymity involved school children and may not generalize to older adolescents or adults. Second, most studies have restricted the range of sensitive health-related questions to those addressing substance use. To our knowledge, no study has evaluated the effects of confidentiality and anonymity on a wide range of sexual behavior self-reports, an area of obvious relevance to health researchers interested in HIV or other sexually transmitted infections. Third, few studies have included manipulation checks to verify that participants believed in anonymity and confidentiality as intended. Fourth, the hypothesis that anonymity affects data quality only in response to sensitive questions has not been tested directly because no study has simultaneously assessed perceived question sensitivity among respondents. Fifth, with only a few exceptions (Bjarnason & Adalbjarnardottir, 2000), the range of possible indicators for the effects of anonymity versus confidentiality assurances has been limited to either response rates (Campbell & Waters, 1989), incidence reports (Ong & Weiss, 2000), or frequency reports (Malvin & Moskowitz, 1983). However, a full appreciation of the effects of anonymity versus confidentiality needs to include all of these indicators and consider the possibility of a trade-off between response refusal (termination, item non-response) and response bias (under- or over-reporting). For example, confidentiality and anonymity may elicit similar incidence or frequency reports because of a systematic dropout of individuals high in social anxiety and perceived question threat in the confidential condition. Also missing from the literature on anonymity effects is attention to individual differences that may influence susceptibility to question threat and to socially desirable responding. Individual differences may help to reconcile conflicting findings, and should be tested as possible moderators of the quality of self-reports, especially for reports of sexual behavior.

The current investigation was designed to investigate the effects of anonymity assurances on reports of sensitive behaviors and to address each of the methodological limitations of previous research (Weinhardt, Forsyth, Carey, Jaworski, & Durant, 1998). We assessed sensitive

behaviors in a sample of young adults (undergraduate students), which is under-represented in methodological studies on anonymity effects. Second, we asked participants to provide data about three sensitive content areas, namely, alcohol consumption and drug use (both of which are illegal among young adults under age 21), and sexual behaviors (Schaeffer, 2000). Third, we completed pilot work to confirm that the content of the questionnaire was considered sensitive and included manipulation checks in the study. Fourth, we assessed perceived question threat in order to examine the relationships between anonymity, perceived item sensitivity, and self-report behavior. Fifth, we used three outcome measures to determine data quality: termination (drop-out), item non-response, and self-reports of sensitive behaviors. These three outcome variables allow for direct comparisons with other studies that have employed at least one of these dependent variables. Sixth, we explored whether two characteristics of the respondent (*viz.*, gender, erotophilia) affected the quality of the data.

We selected gender and erotophilia based on previous evidence that they might affect self-reports of sexual behavior and substance use. First, because both sexual behavior and substance use vary between men and women, and because of the “double-standard” that operates to stigmatize sexual behavior and substance use in women, it is important to know whether the gender of a respondent influences the quality of self-report data about socially sensitive behaviors. Investigators have noted that males tend to report more sexual partners compared to females (Catania, Gibson, Chitwood, & Coates, 1990; Schroder, Carey, & Venable, *in press-a*, *in press-b*; Tourangeau, Smith, & Rasinski, 1997; Turner, Miller, & Rogers, 1997). However, the reason for this reporting bias has not been clarified. Males may be more likely to over-report, and females more likely to underreport these behaviors (Schaeffer, 2000). In the area of substance abuse (including alcohol and nicotine) a clear pattern of reporting by participant gender is less apparent, although there is some evidence that females are less likely to admit use and are more likely to be non-responders than males (Bjarnason & Adalbjarnardottir, 2000).

A second individual difference factor that may affect the quality of self-reported sexual behavior is individuals’ level of comfort about sexual matters. Erotophilia refers to a participants’ comfort with sexual themes, sexual behaviors, and erotica (Fisher, Byrne, White, & Kelley, 1988). Erotophilics show a positive approach to sexual topics and stimuli whereas erotophobic show a generalized avoidance of such stimuli (Gilbert & Gamache, 1984). Thus, participants’ responses may vary in quality as a function of comfort with sexual behaviors and the respective self-reports. Catania, McDermott, and Pollack (Catania, McDermott, & Pollack, 1986) also found that the erotophobia-erotophilia dimension may directly affect the quality of self-reports. Because some respondents may perceive questions about sexual behaviors as threatening (Fisher et al., 1988; Schaeffer, 2000), a direct measure of participants’ comfort with disclosing such material was included as a potential moderator of the experimental condition on data quality.

For this study, we refer to response rate and item refusal as direct indicators of data quality and to incidence and frequency self-reports as indirect indicators of response bias and self-report validity. Retrospective self-reports do not allow a direct estimation of true scores and error and thus can only serve as indirect indicators of data quality; they inform about “more” or “less” but not about the absolute magnitude of error in self-reports. Based on this distinction, the following hypotheses were tested. We predicted that:

1. compared to participants in the confidential condition, participants in the anonymous condition would terminate less often, leave fewer items unanswered, and provide higher incidence and frequency reports of sensitive behaviors;
2. participants’ ratings of perceived question threat would be greater in the confidential condition on the most sensitive behaviors;

3. both perceived question threat and erotophobia would be positively related to termination and item refusal;
4. participants who report engaging in critical behaviors would rate items as more threatening;
5. erotophilia would be positively related to self-reports and explore whether erotophilia would moderate the effects of condition and gender on data quality;
6. the effects of gender on self-reports would be mediated primarily through gender differences in erotophilia; and
7. perceived question threat would mediate the effects of the experimental condition on the number of PNRs but would appear as a second or final outcome in path analyses predicting frequency and incidence self-reports. (The latter assumption is a logical conclusion derived from Hypothesis 4 claiming that engagement in sensitive behaviors should enhance perceived question threat.)

Method

Participants

The participants were 155 male and 203 female undergraduates recruited from psychology courses at Syracuse University. Most were young (male $M = 20.1$, female $M = 20.4$), white (males 71%; females 85%), and in their second or third year of college (sophomores 47%; juniors 51%).

Measures

Demographic Questionnaire—A 7-item questionnaire requested information regarding participants' age, ethnicity, class standing, living circumstances, housing location, sexual orientation, and relationship status.

Sexual Opinion Survey (SOS)—The SOS is a 21-item measure that assesses respondent's disposition to respond to sexual cues along a negative-positive dimension of affective-evaluation from "erotophobic" to "erotophilic" (Fisher et al., 1988). Participants were instructed to rate on a seven-point scale (i.e., 1 = "strongly agree" to 7 = "strongly disagree") how they felt about erotica items. The SOS yields one global score with low scores indicating that the participant is more erotophobic. In college samples, the SOS demonstrates internal consistency (alphas in the current sample were .80 for males, and .89 for females) and test-retest reliability ($r = .80$; Fisher et al., 1988); convergent and discriminant evidence of its validity has been assembled (Tanner & Pollack, 1988).

College Experiences Behavioral Inventory (CEBI)—A 52-item questionnaire was developed to assess the frequency of alcohol, drug, and sex-related activities over the past four weeks and the past 4 years. Many of these items have been used previously in studies of alcohol use, drug use, and sexual behavior, including risk reduction trials (Carey et al., 1997, 2000; 2002; Jaworski & Carey, 2001). Eighteen items assessed current and past alcohol use and related behaviors, 14 items assessed current drug use and associated problems, 13 items assessed current and past frequency of sexual behaviors and related problems, and 7 items assessed frequency of "neutral" behaviors (i.e., behaviors that were not expected to be perceived as embarrassing, stigmatizing, illegal, or private). One-half of the items in each category used a reference time frame of 4 weeks whereas the other one-half referred to behaviors over the past four years. The four-year time period was included to increase the likelihood of participant involvement in sensitive behaviors, and, thus, the variance in the outcomes.

Participant Comfort Questionnaire (PCQ)—For each of the 52 items in the CEBI, participants were asked to rate how uneasy they felt in answering each of these questions. Question threat was assessed with 9-point Likert response scales, ranging from 1 (“not at all uneasy”) to 9 (“very uneasy”). These ratings served as indicators of “question threat.”

Perception of anonymity—Three items assessed participants’ perception of anonymity: “I felt personally identifiable on this survey,” “If someone wanted to know about my response on this survey it would be easy,” and “I felt my answers were anonymous on this survey.” Each answer was scored on a five-point scale (1 = strongly agree, 2 = agree, 3 = neither agree or disagree, 4 = disagree, and 5 = strongly disagree) and scored such that the higher scores indicated a greater perception of anonymity. The internal consistency for participant’s perceptions of anonymity was high (males’ alpha = .89; females’ alpha = .97).

Procedures

Recruitment—Undergraduates were recruited from introductory level psychology courses. With the instructor’s permission, and during regularly scheduled class time, we made a brief presentation regarding the study’s purpose and rationale. The study was introduced as a study of “College Experiences.” Participants were told that, if they chose to participate, they would be asked to respond to a battery of questionnaires inquiring about their attitudes toward, and engagement in, a set of health-related behaviors. They were also told that they would receive extra credit from their instructors commensurate with their time involvement, consistent with practices at the University. After the class presentation, we staffed an information and sign up table outside of the classes to provide students an opportunity to ask additional questions about the study and to volunteer if they were interested. Interested students were invited to attend a subsequent survey session.

Random assignment and the experimental manipulation—When participants arrived for the group session, each was randomly assigned to either the anonymous or the confidential condition using an alternating number system. The procedure then had four steps. First, in a common room, participants in both conditions received identical information regarding the study’s purpose and their rights as a research participant. They were then asked to provide written informed consent.

Next, participants moved to separate rooms to receive instructions that, unbeknownst to them, were unique to each condition. Two critical differences involved the use of identifying information, and the procedure for turning in their survey. In the anonymous groups, participants were told that they were to use a code name (rather than their actual name) when completing the survey. They were told that the purpose of the code name was to assure their anonymity, but that it would allow their data to be linked in a longitudinal (e.g., risk reduction) study. They were told that the code would be determined by the “context determined rule generated pseudonym” technique (Carifio & Biron, 1978; Durant & Carey, 2000). This technique uses the first letter of middle name (if no middle name write “z”); first letter of the month participant was born; first letter of participant’s mother’s first name (if unknown, write “y”); and first letter of participant’s father’s first name (if unknown, write “x”). Also, participants in the anonymous condition were instructed to place their completed survey in an envelope and place it in a large box at the rear of the auditorium. They were also told to take a written debriefing form, and to sign their name on a credit allotment sheet before exiting.

In contrast, participants assigned to the confidential administrations were told to provide explicit identifying information (i.e., the respondents’ name and date of birth) when completing the battery. They were assured that their responses were strictly confidential, and that in scientific reports all data would be reported with group averages rather than individual

responses. They were also instructed to give their completed battery to a research team member seated in the front of the room, who would record their name in order to assure that the participant received credit for his or her participation.

Third, participants from both groups then reassembled in a common room for the survey completion. Before beginning the battery, however, the investigator told all participants that some questions in the battery were personal but that they were intended to obtain necessary information about student health behaviors to facilitate intervention programs. However, the investigator recognized that some participants may feel uncomfortable responding to some questions. The investigator informed participants that, if they felt too uncomfortable with any item, they could select a “prefer not to respond” (PNR) option, and that selecting this option was recommended rather than providing misleading or inaccurate data. In addition, participants were told that they could check a box at the end of the questionnaire without penalty if they wished to withdraw from the study or terminate their participation, but felt uncomfortable doing so publicly. They were reassured that they would receive their course extra credit regardless of whether or not they completed the survey. However, all participants were asked to complete the PCQ even if they choose to withdraw from the study. Participants in both conditions then received a battery of questionnaires that included the demographics questionnaire, SOS, CEBI, and PCQ; unbeknownst to the participants, only the anonymous groups received a large manila envelope with their packets. All participants then completed the battery of questionnaires privately and independently.

Fourth, when participants in the anonymous condition completed the survey, they deposited it in the box in the return of the auditorium, picked up a debriefing form, and signed a list to receive credit for their participation. When participants in the anonymous condition completed the survey, they brought it to the front of the auditorium and gave the battery to a research team member, and were given a debriefing form to read. (The room was configured in such a way that these differential procedures were not apparent to participants.) While the participant read the debriefing form, which explained how the data would be used and the true purpose of the study, the RA checked the questionnaire to see if the participant had checked the “termination” box. If the participant had checked the “termination” box, he or she was asked if they would reconsider this choice now that they had read the debriefing form, and knew the true purpose of the study. All participants were thanked for their contributions.

Data Preparation and Analyses

Selection of target items—Prior to the analysis of the behavioral frequency questions, target questions for the primary and exploratory analyses were selected according to three criteria: (a) a high rating of perceived question threat/sensitivity, (b) a high non-response rate per question, and (c) a common item pool for both genders (i.e., the selected items were rated as most threatening by both men and women). In the initial step, judgments regarding what substantiated a “high” rating were determined separately by gender. For males, a high average threat rating was 2.0 or higher on a 9 point Likert scale whereas for females, a high average threat rating was 2.8 or higher. For both males and females the PNR criteria for inclusion was 5%, a rate that is considered significant in survey research (Turner, 1999). Based on the aforementioned criteria, the 12 most threatening items for each gender were selected. Nine of 12 questions were viewed by both genders as threatening and were selected for further analyses (see Tables 1 and 2).

Data preparation—Prior to data analyses, three preliminary steps were taken. First, outliers (i.e., cases greater than three standard deviations from the mean) were reduced to a raw score corresponding to $z = 3$. Second, skewed variables were re-expressed using the transformation that provided the best approximation to a Gaussian distribution. Third, in addition to (a)

participant refusal and (b) the number of PNR responses, three indicators were prepared from (c) the behavioral self-reports: (1) A sum score was computed using standardized scores of the outlier-treated but untransformed behavioral frequency self-reports (see Table 3 for descriptive statistics). (2) For some analyses, the data were analyzed on incidence level by dividing the sample into those who did and those who did not engage in each of the selected behaviors. (3) As a second behavioral frequency self-report measure, a sum score was computed for the number of behaviors in which subjects reported to have engaged (range 0 to 9).

Data analysis—Data were analyzed by Stata™ Version 7. Chi Square tests were used to determine the effects of experimental condition and gender on rate of participant refusal (i.e., the number of subjects denying use of their data) in two separate analyses. (Logistic regression could not be applied because of empty cells in the anonymous condition. For the same reason, interactions between condition and gender on participation refusal could not be tested.) The number of missing values was analyzed with Poisson regressions (using robust estimation of standard errors). Further, Poisson regressions were applied to behavioral frequency reports, and logistic regressions were used to analyze dichotomized behavioral self-reports on incidence level. Multivariate analysis of variance (ANOVA) was applied to the sum of the *z*-transformed behavioral self-reports, using SPSS™ 10.1. All posthoc planned comparisons were performed one-sided. In order to test whether SOS moderates the effects of the experimental condition on self-reports, hierarchical regressions were conducted. The moderator component (the product of the predictor variables) was entered in the last step of the regression procedure in order to test the incremental variance explained by the interaction term. Mediator effects of SOS and question threat were tested by path analyses, using LISREL 8.20.

Results

Preliminary Analyses

Preliminary analyses were performed to test the equivalence of the experimental groups, using Chi Square and *t*-tests. No differences emerged for any of the demographic variables (i.e., age, ethnicity, living circumstances, housing location, relationship status, sexual orientation, and year in college). Similar results were found for analyses conducted separately by gender.

Manipulation check—Mean comparisons (*t*-tests) were performed on the composite score of participants' perceptions of anonymity. Both men and women in the anonymous condition indicated a greater perception of anonymity than participants in the confidential condition (men: *M*s = 12.34 vs. 5.76, *t* = 23.94, *p* < .0001; women *M*s = 13.49 vs. 5.04, *t* = 44.72, *p* = .0001). This analysis confirms that participants perceived the level of anonymity as intended.

Primary Analyses

The results are divided into three parts: (a) analyses testing the effects of experimental condition and gender on behavioral self-reports and perceived question threat (Hypotheses 1 and 2), (b) analyses testing the links between perceived question threat and erotophilia to the outcomes (Hypotheses 3, 4, and 5), and (c) path analyses testing mediator effects (Hypotheses 6 and 7).

Effects of Experimental Condition and Gender

Termination—Nineteen male and nine female participants, all assigned to the confidential condition, indicated they did not want their data used by checking a box at the end of their survey. However, after debriefing, all 28 potential terminators consented to allow their data to be used. Chi Square tests were performed, showing significant effects for both condition and gender (condition: $Chi^2 = 30.04$, *p* < .0001; gender: $Chi^2 = 7.46$, *p* < .006). Participants in the confidential condition were more likely than those in the anonymous condition to terminate

(28 vs. 0, respectively), and males were more likely to terminate than females (19 vs. 9, respectively).

“Prefer not to respond” (PNR)—The number of PNRs (i.e. the summed and log-transformed number of PNR over the nine target behaviors) was analyzed by ANOVA. There was a main effect for condition on the number of PNR responses across the 9 most sensitive behaviors $F(1, 354) = 95.27, p < .0001, \text{partial } \eta^2 = .21$. Participants in the anonymous condition had fewer PNRs ($M = .044, SD = .25$) than participants in the confidential administration ($M = 2.06, SD = 3.3$). The main effect of gender was not significant [$F(1, 354) = 3.57, p < .06$]. However, the interaction between condition and gender was significant ($F(1, 354) = 6.31, p < .012, \text{partial } \eta^2 = .018$). Although men and women did not differ in the number of PNRs in the anonymous condition, women had significantly more PNR responses than men in the confidential condition ($M = .39$ vs. $M = .26, t = 2.30, p < .012$). In total, condition and gender accounted for 24.5% of the variance in the outcome.

Frequency reports—A multivariate ANOVA was performed to test the effects of condition and gender on behavioral self-reports across the nine target behaviors (see Table 4). The multivariate ANOVA revealed significant effects for condition ($F(9, 254) = 3.29, p < .001, \text{multivariate effect size} = .104$) and gender ($F(9, 254) = 14.57, p < .001, \text{multivariate effect size} = .340$).

Among the univariate tests, condition affected frequency reports for seven of the nine behaviors, with the strongest effects emerging for alcohol use ($F = 20.53, p < .001$, and $F = 18.55, p < .001$), drug use and drug trading ($F = 9.18, p < .003, F = 8.86, p < .003$), and use of fake ID ($F = 12.63, p < .001$). Among sexual behaviors, effects of condition were less consistent with significant effects for masturbation ($F = 10.93, p < .001$) and number of sexual partners ($F = 6.15, p < .014$). The effects for unprotected sex and the number of HIV tests were not significant. For all significant effects, the behavioral self-reports in the confidential condition were lower than in the anonymous condition (see Tables 1 and 4).

Univariate tests for gender revealed significant effects for three of the nine behaviors. The strongest effect emerged for masturbation ($F = 117.78, p < .001$), which alone had an effect size of $\eta^2 = .31$ and which was mainly responsible for the strong multivariate effect of gender. Gender also produced significant effects on self-reports of alcohol consumption ($F = 18.17, p < .001$, and $F = 12.57, p < .001$). For all variables, men’s frequency reports were higher than women’s. No significant main effects of gender were found for number of sex partners, smoking, trading drugs, using a fake ID, number of partners, and HIV tests.

Among the nine behaviors, the only significant condition-by-gender interaction emerged for masturbation ($F = 5.11, p < .025$). Condition had no effect on women’s self-reports, which were low in both conditions ($t = .96, p = .17$), but affected men’s self-reports, which were significantly higher in the anonymous condition ($t = 3.22, p < .002$).

In order to determine whether the effects of condition and gender apply to the composite index of behavioral frequency reports, a univariate ANOVA was performed with the sum of the nine z -transformed variables. In this ANOVA, highly significant effects were found for condition ($F(1, 262) = 13.39, p < .001$) and gender ($F(1, 262) = 15.81, p < .001$) but no interaction was found ($F(1, 262) = .08, p < .774$). η^2 for the model was .096.

Perceived question threat—Parallel to the frequency reports, a multivariate ANOVA was performed testing the effects of condition and gender on perceived question threat of the nine behaviors (see Table 4). The multivariate effects of condition and gender were significant

(condition: $F(9, 341) = 5.29, p < .001$, partial $\eta^2 = .123$; gender: $F(9, 341) = 4.12, p < .001$, partial $\eta^2 = .098$) but no interaction effect emerged ($F(9, 341) = 1.24, p = .269$).

Univariate tests for condition revealed seven significant effects. The strongest differences between anonymous and confidential assessment emerged for smoking marijuana ($F(1, 349) = 29.92, p < .001$) and use of a fake ID ($F(1, 349) = 21.09, p < .001$). All effects were in the predicted direction with subjects in the confidential condition giving higher threat ratings than subjects in the anonymous condition (see Table 2).

Gender produced significant effects on all threat ratings, with the strongest gender difference emerging for the number of partners during the past 4 years ($F(1, 349) = 23.75, p < .001$), and the weakest gender difference for the question on trading drugs ($F(1, 349) = 4.62, p < .032$). In all cases, women gave higher threat ratings than men.

A univariate ANOVA tested effects of condition and gender on the sum of the threat ratings. As with the multivariate ANOVA, a significant effect emerged for both condition and gender. Subjects in the confidential condition had higher threat ratings than participants in the anonymous condition ($M = 27.5, SD = 14.8$ vs. $M = 20.1, SD = 12.5, F(1, 349) = 25.45, p < .001$), and women had higher threat ratings than men ($M = 27.1, SD = 15.2$, vs. $M = 19.4, SD = 11.2; F(1, 349) = 28.49, p < .001$). Together, condition and gender explained 14% of the variance in perceived question threat.

Perceived Question Threat and Erotophilia as Predictors of Data Quality and Self-Reports

Hypotheses 3, 4, and 5 referred to the predictive value of perceived question threat and erotophilia on data quality and behavioral self-reports. We expected that (a) question threat would increase, and (b) erotophilia would decrease, the likelihood of termination and the number of PNR responses. For termination, *t*-tests were performed. Subjects who initially refused their participation gave significant higher threat ratings than those who consented to the use of their data ($M = 41.8$ vs. $M = 30.1, t = 3.14, p < .001$). Erotophilia was not related to termination ($M = 69.1$ vs. $M = 73.1, t = 1.08, p < .28$). Similarly, correlations with the number of PNRs indicated significant effects for question threat ($r = .36, p < .0001$) but not for erotophilia ($r = -.02, p = .66$).

With regard to incidence self-reports, we predicted that perceived question threat and erotophilia would both be positively related to engagement in risky and socially stigmatized behaviors, and would explain additional (i.e., unique) variance in the outcomes. A series of hierarchical logistic regressions was conducted with the dichotomized behavioral self-reports as outcomes. For each of the nine dichotomous outcome variables, condition and gender were entered as control variables in the first step, perceived question threat in the second step, and erotophilia in the last step of the logistic regressions. Table 5 shows the results. For every behavior, perceived question threat added significantly to the prediction of sensitive behaviors with ΔChi^2 ranging from 4.7 (HIV test) to 62.1 (smoked marijuana). Erotophilia explained additional variance when entered in the third step with ΔChi^2 values ranging from 6.8 (multiple partners) to 88.6 (masturbation). The only non-significant result emerged for the item on HIV tests. Across all nine behaviors, both question threat and erotophilia were positively related to engagement in the proscribed behaviors.

For frequency self-reports, Spearman rank correlations revealed significant positive relationships between perceived question threat for all but one of the behavioral self-reports, indicating that subjects perceived questions about a behavior as increasingly threatening the more often they engage in it (see Table 5). (The only non-significant correlation was found for the number of times a person was drunk during the past 4 years). Similarly, erotophilia was positively related to frequency self-reports, except for the item on HIV testing.

Moderator effects of erotophilia—In order to test the assumption that liberal attitudes towards sex, as measured with the SOS, would moderate the effects of condition and gender on behavioral self-reports, a series of hierarchical regressions was performed. The regressions included condition-by-erotophilia and gender-by-erotophilia interactions. For the number of PNR responses, a Poisson regression (robust) was performed and revealed a significant interaction between condition and erotophilia ($z = 1.97, p < .049$). The positive coefficient indicates that the tendency to produce more PNRs in the confidential and less PNRs in the anonymous condition was stronger among erotophilic subjects, whereas the number of PNRs depended less on anonymity among erotophobic subjects. Further hierarchical regressions were performed with the dichotomized behavioral self-reports, the z-score composite of behavioral frequencies, and the number of different behaviors subjects reported (ranging from 0 to 9). However, except for the number of PNRs, no further interactions of erotophilia with either condition or gender emerged, although the main effect of erotophilia on the outcomes was significant in all but one of the analyses.

Path analyses testing mediator effects—A series of three path analyses was performed in order to test mediator effects of erotophilia and perceived question threat on data quality, and the effects of these potential mediators on behavioral self-reports (Hypotheses 6 and 7). Path analyses were performed separately for each of three quantitative outcomes (i.e., the number of PNRs; the number of behaviors a person reported, ranging from 0–9; and the composite of the standardized and normalized behavioral frequency reports). The analyses were performed with LISREL 8.20 (Joereskog & Soerbom, 1996a), using matrices of polychoric and polyserial correlations provided by PRELIS 2 (Joereskog & Soerbom, 1996b) (see Table 6).

Figure 1a shows the preliminary assumptions used as a working model. Erotophilia was expected to mediate the effects of gender on self-report behavior. Direct causal paths were specified onto question threat, with experimental condition (positive, with anonymous = 0, confidential = 1) and gender (negative, with female = 1, male = 2) as predictors. Dependent on the fit, the model was modified in subsequent steps. If relationships between erotophilia and question threat were indicated by intermediate model test results, they were interpreted as a causal effect of erotophilia.

Path specification towards the behavioral outcome depended on the indicator. For the number of PNRs, which is regarded as a direct indicator of reduced data quality, question threat was specified as a positive predictor, and simultaneously as mediator of condition and gender effects. Figure 1b shows the results of the final model, which provided the best fit to the data. They indicate a strong direct effect of condition on the number of PNRs, in addition to an indirect effect mediated by question threat. Further, the effects of gender on the number of PNRs were fully mediated by perceived question threat, with men being less affected by question threat and accordingly less likely to refuse responses than women. Erotophilia was predicted by gender but unrelated to either question threat or the number of PNRs. In total, 24% of the variance in question threat and 37% of the variance in the number of PNRs were explained by the model. The fit was excellent with $\text{Chi}^2 = 3.62, df = 4, p = .46; \text{RMSEA} = .00, \text{RMSE} = .018, \text{GFI} = 1.00, \text{AGFI} = .99$.

The model predicting the number of behaviors is shown in Figure 1c. In this model, a positive path from self-report onto question threat was specified, indicating that engagement in a non-normative behavior would enhance perceived item threat. In order to enhance the fit of the model, a path leading from erotophilia to question threat was required. Both condition and gender had direct and indirect effects on question threat, with indirect effects being mediated through behavioral self-reports. Self-reports were predicted with direct paths by confidentiality (negative) and erotophilia (positive). Gender effects on self-report were mediated by

erotophilia (with men having higher scores in both). The final model had a satisfying fit with $\chi^2 = 10.67$, $df = 3$, $p = .014$; $RMSEA = .099$, $RMSE = .038$, $GFI = .99$, $AGFI = .96$. In both self-report and question threat, 22% of the variance could be explained.

A final model was tested for the behavior frequency composite. This outcome appeared unrelated to perceived question threat but was predicted directly by condition and gender, with an additional partial mediation of gender effects through erotophilia. Again, male sex and erotophilia appeared as positive predictors and confidentiality as a negative predictor of behavior frequency reports. The fit was satisfying with $\chi^2 = 11.26$, $df = 3$, $p = .010$; $RMSEA = .10$, $RMSE = .042$, $GFI = .99$, $AGFI = .96$. Sixteen percent of the variance in the frequency reports was explained.

Discussion

The present study examined the effects of anonymity and gender on self-report of socially sensitive health behaviors and tested question threat and erotophilia as their potential mediators. Perceived question threat served a twofold purpose in this study: First, threat ratings were used to select those questions that are experienced as most sensitive by the respondents. Previous research has suggested that anonymity affects self-report behavior only if the target behaviors are of sensitive nature (Singer et al., 1995), but failed to assess perceived question threat in order to test this assumption. Thus, it is possible that studies reporting no effects of anonymity on response frequency (e.g., Malvin & Moskowitz, 1983; Werch, 1990; Zagumny et al., 1996) failed to present items that were perceived as threatening by the participants. By focusing on items identified as most sensitive by the respondents, we were able to address the characteristic of the self-reports that should prompt greatest concerns about the quality of the data. Second, we analyzed question threat as an intermediate outcome of the assessment conditions, examined its relationships with self-reports, and tested its potential mediator effects on data quality.

Further, we included three major outcome criteria in our research: Termination and the number of item refusals served as direct measures of data quality, and variations in self-reports were analyzed as indirect indicators of response bias and self-report validity. Previous methodological studies often relied exclusively on incidence or frequency self-reports in order to derive conclusions about the quality of the data (Catania et al., 1990; Catania et al., 1993; Malvin & Moskowitz, 1983; Ong & Weiss, 2000; Tourangeau et al., 1997; Turner, Miller, Smith, Cooley, & Rogers, 1996; Weinhardt et al., 1998). However, retrospective self-reports do not allow direct evaluations of their validity, and the use of inferential heuristics such as “higher self-reports of sensitive behaviors are more accurate” is not a convincing “validation” strategy (Schroder et al., in press-b). Methodological studies testing mode effects on self-reports of sensitive behaviors indicate that inaccuracy may also occur in the form of over-reporting (Berk, Abramson, & Okami, 1995; Coxon, 1999; Downey, Ryan, Roffman, & Kulich, 1995; Leigh, Gillmore, & Morrison, 1998; Schroder et al., 2001). Therefore, we distinguished between (a) response refusal and termination as direct quality criteria, and (b) systematic variations in self-report as indirect indicators of self-report validity.

Seven key findings emerged from this research. First, the confidential condition displayed detrimental effects on data quality and reduced incidence and frequency reports. Second, women reported fewer sensitive behaviors and were more likely to terminate or to indicate a preference not to answer (the latter effect occurred only in the confidentiality condition). Third, parallel effects of condition and gender were observed in perceived question threat: confidential, non-anonymous assessment enhanced question threat ratings, and women perceived the items as more threatening than men. Fourth, question threat was negatively related to data quality (termination, response refusal) but positively related to incidence and

frequency reports. Fifth, in path models, question threat mediated the effects of condition and gender on non-response, but was modeled best as an outcome of engagement in sensitive behaviors. Sixth, erotophilia did not affect data quality directly; instead, erotophilia moderated the effects of condition on the number of PNRs and had independent positive effects on incidence and frequency self-reports. Seventh, erotophilia was lower in women and appeared as a mediator of gender effects on self-reports. We discuss each of these results in turn.

The present study provides strong support for the assumption that anonymity enhances the quality of self-report of socially sensitive behaviors. Both termination and the number of PNR responses were reduced in the anonymous condition; moreover, incidence and frequency reports of sensitive behaviors were higher in the anonymous condition. These results corroborate the pattern found in the literature, which suggests that anonymity manipulations affect responses to socially sensitive questions. Self-reports of alcohol and drug use were most strongly affected by confidentiality, which may be due to the fact that both alcohol and drug use are illegal among young adults below age 21. Self-reports of sexual behaviors were less affected by condition.

In addition, self-reports were affected by gender. Further, the results suggest that anonymity reduces the effects of gender on data quality. Women were more likely to choose termination, and had a higher number of PNRs in the confidential condition. In both the anonymous and confidential conditions, incidence and frequency reports of women were lower than men's. Women also perceived questions about sexuality, drinking behavior, and drug use as more threatening. Although the pattern of results might be interpreted to indicate that women are more motivated to bias their responses, the evidence is not sufficient to infer greater bias in women's self-reports. Several contradictory arguments have to be taken into account. First, men's and women's self-reports were equally affected by anonymity manipulations, and gender differences were equally strong in both conditions. If women would be more inclined to bias their responses, one would expect greater gender differences in the confidential condition (an effect that emerged only for the PNR variable). Second, women may choose to terminate or refuse responses as an alternative to biasing their answers. Third, the same gender-specific social norms that enhance question threat in women are likely to prevent frequent engagement in these behaviors. Thus, although women may experience questions as more threatening, they may have less to report. Fourth, question threat may be enhanced in women because they intend to provide valid answers but associate more negative consequences with candid self-report. In addition, there is no evidence in the literature indicating a higher response bias for sensitive behaviors among women. In sum, the results of this study indicate that women feel more threatened and are more likely to refuse or terminate, but they do not allow inferring higher self-report accuracy among men.

We included perceived question threat as a potential motivational factor that could explain the effects of anonymity on self-report behavior. The simultaneous effects of anonymity on self-report behavior and perceived question threat suggest a possible mediator role of question threat. With regard to data quality, we found support for the mediator hypothesis. Termination occurred only in the confidential condition, and those who withdraw felt more threatened. Further, the mediator model predicting the number of PNRs, which tested indirect effects of condition and gender through question threat (see Figure 1b), provided an excellent fit to the data. However, we found only a partial mediation of the effects of confidentiality on the number of PNRs. Thus, the effects of condition may be mediated by additional factors. Non-anonymous assessments may create additional barriers to open responding, such as a perceived violation of privacy or a perceived lack of justification for the requested information.

It seems reasonable to assume that the effects of anonymity on incidence and frequency self-reports are mediated by question threat as well. However, this contradicts the expectation that

involvement in prohibited and socially disapproved behaviors enhance perceived question threat. The former assumption requires a negative relationship, the latter a positive relationship between self-reports and perceived threat. We found positive correlations between perceived question threat and self-reports, which implies an effect of self-reports on threat perception and contradicts the idea that question threat mediates the effects of condition on self-report. Theoretically, we still may assume a bi-directional relationship between question threat and self-report behavior, as indicated in Figure 1a. However, our results suggest that a possible inhibiting effects of question threat on self-report is masked or superseded by the threat-enhancing effect of reporting about proscribed and disapproved behaviors. Accordingly, perceived question threat occurred as an outcome of engagement in the target behaviors (Figure 1c) or appeared unrelated to the composite of the behavioral frequencies (Figure 1d).

In addition to perceived question threat, the present study included erotophilia as a potential moderator and mediator of self-report behavior. Methodological research on self-report validity stresses the importance of social desirability and perceived social norms (Catania et al., 1990) but neglects other potentially significant motivational factors such as personal opinions. We assume that self-report behavior is affected by similar social-cognitive variables that energize other behaviors in everyday life. According to the Theory of Reasoned Action (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 1975) and its successor, the Theory of Planned Behavior (Ajzen, 1991; Ajzen & Madden, 1986), perceived social norms are only one source that affects behavior. In addition, self-report behavior, including dropout, item refusal, and response bias, may be affected by attitudes towards the target behaviors and the assessment situation.

Erotophilia may be regarded as an attitudinal measure of internalized liberal or restrictive norms, which we included to explore its effects on self-report behavior. In contrast to Catania et al. (1986), who found evidence for direct effects of erotophilia on item refusal, erotophilia appeared as unrelated to both the number of PNRs and termination in the present study. However, erotophilia was found to moderate the effects of anonymity of item refusal. Response refusal of erotophilic subjects was stronger affected by non-anonymity than response refusal of erotophobic subjects. Thus, it appears that for erotophobic subjects, decisions regarding open responding and item refusal are determined more by internal (attitudinal) factors and less by external factors (the assessment situation). Further, erotophilia was positively and independently related to incidence and frequency self-reports. It seems that erotophilic subjects are more likely to engage in a variety of hedonistic behaviors that are not socially approved. However, it is unclear whether erotophilia affects response bias in addition to affecting behavior. It is possible that erotophobic subjects do not only engage less in the target behaviors but also under-report their behavior.

The relationships among erotophilia, gender, and behavioral self-reports are difficult to interpret. Path analyses indicate that erotophilia mediates gender effects on self-reports, with men scoring higher in both erotophilia and behavioral self-reports. It appears that, in accordance with gender-specific social norms, women's attitudes inhibit substance use and sexual behavior. Whether and to what degree erotophilia contributes to this pattern by affecting response bias in addition to affecting behavior remains an open question. It is possible that erotophobia, similar to question threat, causes under-reporting among women. If it is true that, in response to sensitive questions, under-reporting occurs, the logical conclusion would be that self-reports of women are more biased than those of men because women provide lower self-reports and are more motivated to underreport. Again, such conclusions may be premature because there is no direct evidence that women provide less accurate self-reports than men. Further, methodological studies using diary data for comparison indicate that over-reporting of non-normative behaviors may affect the accuracy of retrospective self-reports as well (Coxon, 1999; Leigh et al., 1998).

The same caution seems warranted regarding the conclusion that confidential self-reports are less accurate than anonymous self-reports. This conclusion remains untested. Social norms in everyday life vary for different social groups. In some groups, substance use and risk taking receives approval from peers. Response tendencies may be influenced by perceived social norms of significant others such as peer groups, friends, and sexual partners who may provide the significant normative frame for self-evaluation and response bias in anonymous assessment conditions. More research is needed to test the preconditions, mediating, and moderating factors that affect response bias in anonymous and confidential assessments.

Although the current study improves upon prior efforts, we wish to acknowledge three limitations. First, we did not obtain a second source of information about subjects' behavior; without a validation criterion, we cannot determine whether over- or under-reporting occurred (cf. Carey et al., 2001). Second, the study focuses on a population of young adults under age 21 and may not generalize to other populations. Alcohol and drug use items are illegal in this population unlike sexual activities, which are neither prohibited nor associated with strict norms in this population. For that reason it is not surprising that questions regarding substance use were perceived as more threatening, and that the effects of anonymity occurred mainly in self-reports of substance use. Third, although a range of sexual risk behaviors were assessed in the present study, sexual behavior was represented with three items only in the set of most threatening questions. This does not allow a generalization of the results to the full domain of sexual risk behaviors that is targeted in HIV risk populations. The effects of anonymity, gender, question threat, and erotophilia need to be tested further with higher-risk groups and with a wider array of sexual behaviors.

In sum, the present study provides clear evidence that non-anonymous assessment conditions can lead to reduced data quality, and suggestive evidence that non-anonymous conditions may lead to under-reporting of socially sensitive health behaviors. Therefore, we recommend the use of anonymous assessments whenever feasible. In longitudinal research and other studies that require identification of subjects, fully anonymous assessments are not possible so we recommend the use of methods that optimize the degree of anonymity and privacy they offer. Using unique codes and visibly separating identifying information from item responses can enhance anonymity and may improve data quality. In general, we recommend use of assessment techniques that reduce the perceived threat, to enhance data quality and minimize under-reporting bias. We also recommend instructions that reduce the impact of social norms, create a neutral assessment situation, provide a convincing and persuasive justification for the request for private information, reduce concerns against open responding, and stress the importance of accurate self-reports. Instructions such as these will optimize data quality and reduce response bias. We also call for continued research on the determinants of data quality and accuracy, especially in the assessment of social sensitive health behaviors. We regard the inclusion of perceived question threat as an important step forward in methodological research on data quality and the validity of behavioral self-reports. Studies that do not include an independent measure of question threat or perceived item sensitivity miss a chance to directly test motivational hypotheses on self-report bias.

Acknowledgements

This study was supported by grants from the National Institute of Mental Health (R01-MH54929 and K02-MH01582) to the Michael P. Carey. Correspondence concerning this article should be addressed to Michael P. Carey, Ph.D., Director, Center for Health and Behavior, 430 Huntington Hall, Syracuse University, Syracuse, New York 13244-2340. Electronic mail should be sent via Internet to mpcarey@syr.edu.

References

- Ajzen I. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 1991;50:179-211.

- Ajzen, I.; Fishbein, M. Understanding attitudes and predicting health behavior. Englewood Cliffs, NJ: Prentice Hall; 1980.
- Ajzen I, Madden TJ. Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology* 1986;22:453–474.
- Berk, R.; Abramson, PR.; Okami, P. Sexual activities as told in surveys. In: Abramson, PR.; Pinkerton, SD., editors. *Sexual Nature, Sexual Culture*. Chicago: University of Chicago; 1995.
- Bjarnason T, Adalbjarnardottir. Anonymity and confidentiality in school surveys on alcohol, tobacco, and cannabis use. *Journal of Drug Issues* 2000;22:335–344.
- Campbell MJ, Waters WE. Does anonymity increase response rate in postal questionnaire surveys about sensitive subjects? A randomized trial. *Journal of Epidemiology and Community Health* 1989;44:75–76. [PubMed: 2348155]
- Carey MP, Braaten LS, Maisto SA, Gleason JR, Forsyth AD, Durant LE, Jaworski BC. Using information, motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban women: a second randomized clinical trial. *Health Psychology* 2000;19:3–11. [PubMed: 10711582]
- Carey MP, Carey KB, Maisto SA, Gordon CM, Weinhardt LS. Assessing sexual risk behavior with the timeline followback (TLFB) approach: Continued development and psychometric evaluation. *International Journal of STD and AIDS* 2001;12:365–375.
- Carey, MP.; Carey, KB.; Maisto, SA.; Venable, PA.; Schroder, KEE. Reducing HIV risk among psychiatric outpatients: Results from a randomized controlled trial. 2002. Manuscript submitted for publication
- Carey MP, Maisto SA, Kalichman SC, Forsyth AD, Wright EM, Johnson BT. Enhancing motivation to reduce the risk of HIV infection for economically disadvantaged urban women. *Journal of Consulting and Clinical Psychology* 1997;65:531–541. [PubMed: 9256553]
- Carifio J, Biron R. Collecting sensitive data anonymously: The CDRGP technique. *Alcohol and Drug Education* 1978;23:47–66.
- Catania JA, Gibson DR, Chitwood DD, Coates TJ. Methodological problems in AIDS behavioral research: influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin* 1990;108:339–362. [PubMed: 2270232]
- Catania JA, McDermott L, Pollack L. Questionnaire response bias and face-to-face interview sample bias in sexuality research. *Journal of Sex Research* 1986;22
- Catania, JA.; Turner, H.; Pierce, RC.; Golden, E.; Stocking, C.; Binson, D.; Mast, C. Response bias in surveys of AIDS-related sexual behavior. In: Ostrow, DG.; Kessler, RC., editors. *Methodological issues in AIDS behavioral research*. New York: Plenum Press; 1993. p. 133-162.
- Coxon AP. Parallel accounts? Discrepancies between self-report (diary) and recall (questionnaire) measures of the same sexual behaviour. *AIDS Care* 1999;11:221–234. [PubMed: 10474624]
- Downey L, Ryan R, Roffman R, Kulich M. How could I forget? Inaccurate memories of sexually intimate moments. *Journal of Sex Research* 1995;32:177–191.
- Durant LE, Carey MP. Self-administered questionnaires versus face-to-face interviews in assessing sexual behavior in young women. *Archives of Sexual Behavior* 2000;29:309–322. [PubMed: 10948721]
- Fishbein, M.; Ajzen, I. *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, Mass: Addison-Wesley; 1975.
- Fisher WA, Byrne D, White LA, Kelley K. Erotophilia as a dimension of personality. *Journal of Sex Research* 1988;25:123–151.
- Gilbert FS, Gamache MP. The sexual opinion survey: Structure and use. *Journal of Sex Research* 1984;20:293–309.
- Jaworski BC, Carey MP. Effects of a comprehensive STD-related information, enhanced motivation, and behavioral skills intervention on sexual behavior in female college students. *Journal of Adolescent Health* 2001;29:417–425. [PubMed: 11728891]
- Joereskog, K.; Soerbom, D. LISREL 8: User's reference guide. Chicago, IL: Scientific Software International; 1996a.
- Joereskog, K.; Soerbom, D. PRELIS 2: User's reference guide. Chicago, IL: Scientific Software International; 1996b.

- Leigh BC, Gillmore MR, Morrison DM. Comparison of diary and retrospective measures for recording alcohol consumption and sexual activity. *J Clin Epidemiol* 1998;51:119–127. [PubMed: 9474072]
- Malvin JH, Moskowitz JM. Anonymous versus identifiable self-reports of adolescent drug attitudes, intentions, and use. *Public Opinion Quarterly* 1983;47:557–566.
- O'Malley PM, Johnston LD, Bachman JG, Schulenberg J. A comparison of confidential versus anonymous survey procedures: Effects on reporting of drug use and related attitudes and beliefs in a national study of students. *Journal of Drug Issues* 2000;30:35–54.
- Ong AD, Weiss DJ. The impact of anonymity on responses to sensitive questions. *Journal of Applied Social Psychology* 2000;30:1691–1708.
- Schaeffer, NC. Asking questions about threatening topics: A selective overview. In: Stone, AA.; Turkkan, JS.; Bachrach, CA.; Jobe, JB.; Kurtzman, HS., editors. *The science of self-report: Implications for research and practice*. Mahwah, N.J.: Lawrence Erlbaum Associates; 2000. p. 105-121.
- Schroder KEE, Carey MP, Vanable PA. Methodological issues in the assessment and analysis of sexual risk behavior: I. Item content and scaling. *Annals of Behavioral Medicine*. in press-a
- Schroder KEE, Carey MP, Vanable PA. Methodological issues in the assessment and analysis of sexual risk behavior: II. Accuracy of self-reports. *Annals of Behavioral Medicine*. in press-b
- Singer E, Von Thurn DR, Miller ER. Confidentiality assurances and response: A quantitative review of the experimental literature. *Public Opinion Quarterly* 1995;59:66–77.
- Tanner WM, Pollack RH. The effect of condom use and erotic instructions on attitudes towards condoms. *Journal of Sex Research* 1988;25:537–541.
- Tourangeau R, Smith TW, Rasinski KA. Motivation to report sensitive behaviors on surveys: Evidence from a bogus pipeline experiment. *Journal of Applied Social Psychology* 1997;27:209–222.
- Turner, CF.; Miller, HG.; Rogers, SM. Survey measurement of sexual behavior: Problems and progress. In: Bancroft, J., editor. *Researching sexual behavior. Methodological issues*. Bloomington, IN: Indiana University; 1997.
- Turner, CF.; Miller, HG.; Smith, TK.; Cooley, PC.; Rogers, SM. Telephone audio computer-assisted self-interviewing (T-ACASI) and survey measurements of sensitive behaviors: Preliminary results. In J. F. In: Banks, R.; Gerrard, L., editors. *Survey and Statistical Computing 1996*. Chesham Bucks, U.K.: Association for Survey Computing; 1996.
- Turner HF. Participation bias in AIDS-related telephone surveys: Results from the National AIDS Behavioral Survey (NABS) non-response study. *Journal of Sex Research* 1999;36:52–58.
- Weinhardt LS, Forsyth AD, Carey MP, Jaworski BC, Durant LE. Reliability and validity of self-report measures of HIV-related sexual behavior: Progress since 1990 and recommendations for research and practice. *Archives of Sexual Behavior* 1998;27:155–180. [PubMed: 9562899]
- Werch CE. Two procedures to reduce response bias in reports of alcohol consumption. *Journal of Studies on Alcohol* 1990;51:327–330. [PubMed: 2359305]
- Zagumny MJ, Ramsey RJ, Upchurch MP. Is anonymity important in AIDS survey research? *Psychological Reports* 1996;78:270. [PubMed: 8839318]

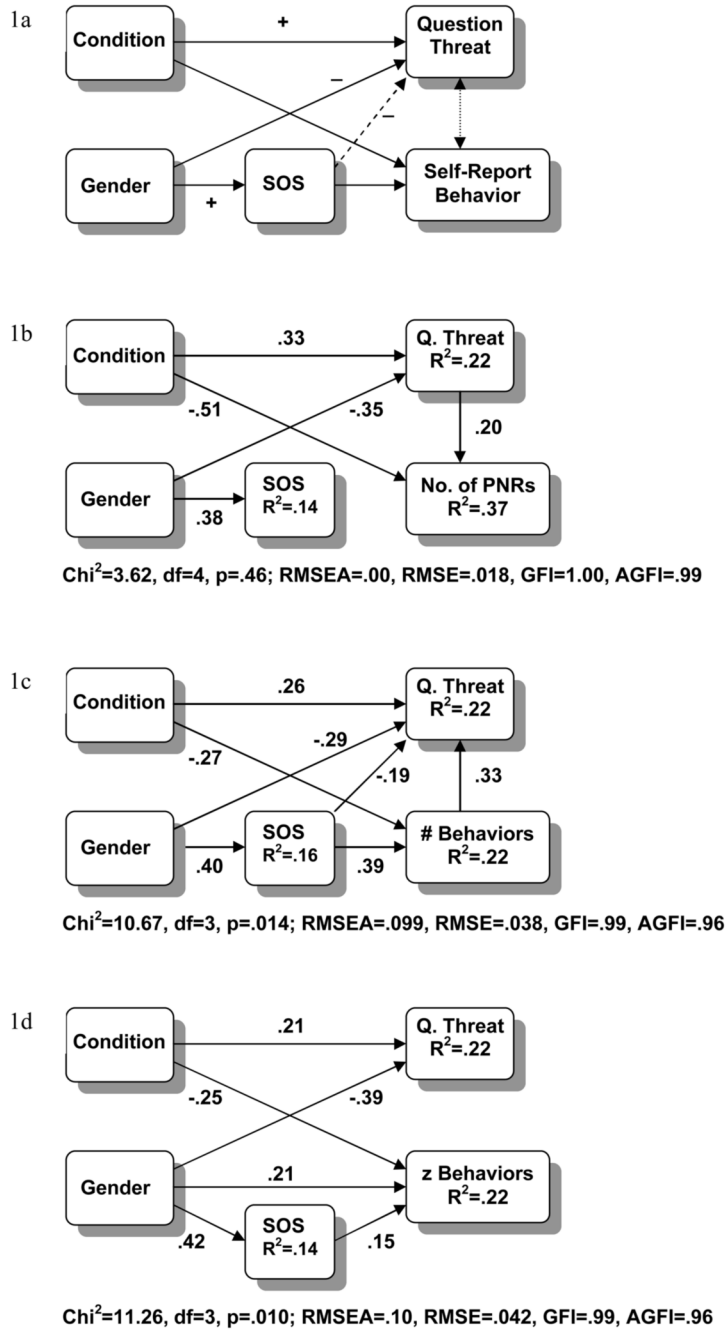


Figure 1.
 Mediator Models Linking Perceived Question Threat and Self-Report Behavior

Table 1
Descriptive Statistics for Behavioral Self-reports for the Nine Most Threatening Items by Gender and Condition

Variable	Gender	n	Anonymous (Control) Condition			Confidential Condition			%			
			M	SD	Range	M	SD	Range				
No. Sex Partners	Female	99	3.5	3.5	0-15	2.9	3.7	0-20	35%	35	74	80%
	Male	78	4.2	5.5	0-40	4.3	6.1	0-35	82%	64	52	74%
Unprotected Sex ¹	Female	99	2.7	6.1	0-28	3.0	6.0	0-30	35%	35	36	36%
	Male	79	2.8	6.4	0-40	1.8	3.1	0-15	33%	26	20	27%
Masturbation ¹	Female	97	1.8	3.1	0-15	1.4	3.2	0-20	41%	40	25	28%
	Male	76	12.2	11.2	0-60	9.7	13.8	0-60	92%	70	42	70%
Smoked Marijuana	Female	99	76.4	191.5	0-1000	7.6	16.7	0-100	59%	58	30	46%
	Male	79	150.7	556.1	0-4380	28.8	68.3	0-400	68%	54	38	60%
Traded Marijuana	Female	99	24.7	108.5	0-1000	3.1	13.2	0-100	37%	37	18	21%
	Male	78	20.2	65.3	0-500	8.8	23.9	0-100	37%	29	21	31%
No. of HIV Tests	Female	99	.4	.9	0-4	.5	1.0	0-6	23%	23	25	24%
	Male	79	.33	.86	0-4	.27	.8	0-4	17%	13	11	15%
Used Fake ID	Female	99	36.3	86.9	0-575	13.4	32.0	0-175	58%	57	41	52%
	Male	79	43.9	126.9	0-1000	28.8	80.9	0-500	60%	47	32	47%
Times >5 Drinks	Female	99	63.6	108.9	0-560	27.3	42.9	0-200	88%	87	71	81%
	Male	79	132.6	221.4	0-1000	76.7	148.0	0-850	91%	72	55	81%
Times Drunk	Female	98	73.7	116.0	0-576	51.7	123.7	0-1000	89%	87	72	83%
	Male	79	128.6	196.6	0-1000	78.2	146.1	0-800	91%	72	57	75%

¹ Behavior and threat ratings assessed for a reference period of 4 weeks, otherwise for a reference period of 4 years.

Table 2
Descriptive Statistics for Question Threat Ratings for the Nine Most Threatening Items by Gender and Condition

Variable	Gender	n	M	Anonymous Condition			Confidential Condition						
				SD	Range	PNR _s ²	% PNR ²	n	M	SD	Range	PNR _s ²	% PNR ²
No. Sex Partners	Female	99	3.31	2.81	1-9	0	0%	104	4.37	3.38	1-9	12	11.5%
	Male	79	2.08	1.96	1-9	1	1.3%	76	2.93	2.56	1-9	6	7.9%
Unprotected Sex ¹	Female	99	2.74	2.49	1-9	0	0%	104	3.23	3.00	1-9	3	2.9%
	Male	79	1.84	1.54	1-9	0	0%	76	2.26	2.14	1-9	1	1.3%
Masturbation ¹	Female	99	3.82	3.00	1-9	2	2.0%	104	5.20	3.53	1-9	16	15.4%
	Male	79	3.33	2.80	1-9	3	3.8%	76	4.14	2.79	1-9	16	21.1%
Smoked Marijuana	Female	99	2.37	2.38	1-9	0	0%	104	4.00	3.34	1-9	38	36.5%
	Male	79	1.58	1.53	1-9	0	0%	76	2.91	2.59	1-9	13	17.1%
Bought/Sold Marijuana	Female	99	2.35	2.52	1-9	0	0%	104	3.05	3.17	1-9	19	18.3%
	Male	79	1.87	1.96	1-9	1	1.3%	74	2.45	2.47	1-9	9	11.8%
Number of HIV Tests	Female	99	2.54	2.50	1-9	0	0%	104	3.17	2.94	1-9	1	1.0%
	Male	79	2.01	1.84	1-9	0	0%	76	2.05	2.10	1-9	2	2.6%
Used Fake ID	Female	99	1.92	1.95	1-9	0	0%	104	3.12	2.95	1-9	25	24.0%
	Male	79	1.32	1.06	1-9	0	0%	76	2.28	2.09	1-9	8	10.5%
Times Drank >5 Drinks	Female	99	1.87	1.76	1-9	0	0%	103	2.51	2.48	1-9	16	15.4%
	Male	79	1.33	.78	1-7	0	0%	75	1.84	1.57	1-7	8	10.5%
Times Drunk	Female	99	2.13	1.95	1-9	1	1.0%	103	2.17	2.10	1-9	17	16.3%
	Male	79	1.30	.85	1-7	0	0%	76	1.92	1.62	1-7	9	11.8%

¹ Behavior and threat ratings assessed for a reference period of 4 weeks, otherwise for a reference period of 4 years;

² PNRs and % PNR refer to the number of items of the behavioral self-reports that respondents preferred not to answer.

Table 3
Descriptive Statistics of Further Dependent Variables by Condition and Gender

Variable	Gender	Anonymous Condition					Confidential Condition				
		<i>n</i>	<i>M</i>	<i>SD</i>	Med.	Range	<i>n</i>	<i>M</i>	<i>SD</i>	Med.	Range
<i>z</i> - Score Behavior Composite	Female	96	-.06	5.37	-.53	(-5.2) -19.1	47	-2.75	2.32	-3.60	(-5.2) -2.7
	Male	74	2.45	6.07	.61	(-5.1) -19.2	49	-.15	5.99	-1.83	(-5.2) -20.2
Sum of missing values	Female	99	.11	.35	0	0-2	104	2.5	3.0	1	0-14
	Male	79	.18	.45	0	0-2	76	1.9	3.8	0	0-20
Terminators	Female	99	0	0	0	0	104	.09	.28	0	0-1
	Male	79	0	0	0	0	76	.25	.44	0	0-1
Perceived Question Threat	Female	99	23.1	14.1	20	9-71	102	31.0	15.3	30	9-68
	Male	78	16.3	8.7	13	9-52	74	22.6	12.6	18.5	9-54

Table 4
Multivariate ANOVAs Testing the Effects of Condition and Gender on Behavioral Self-reports (Upper Half) and Question Threat (Lower Half)

<i>n</i> = 266	Condition			Gender			Condition * Gender		
	<i>F</i>	<i>df</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>F</i>	<i>df</i>	<i>p</i>
Behavioral Self-Reports									
No. of partners	6.2	1	.014	3.1	1	.078	1.4	1	.242
Unprotected sex	0.0	1	.919	0.1	1	.749	0.3	1	.613
Masturbation	10.9	1	.001	117.8	1	.000	5.1	1	.025
Smoke Marijuana	8.9	1	.003	2.8	1	.095	0.1	1	.733
Bought/Sold Marijuana	9.2	1	.003	3.6	1	.053	2.2	1	.144
HIV tests	0.3	1	.572	1.6	1	.208	0.0	1	.825
Fake ID	12.6	1	.000	2.8	1	.094	1.2	1	.277
Freq. drink > 5 drinks	20.5	1	.000	18.2	1	.000	0.1	1	.778
Drunk	18.5	1	.000	12.3	1	.001	0.0	1	.851
Multivariate (<i>n</i> = 353)	3.3	9	.001	14.6	9	.000	1.7	9	.091
Question Threat Items									
No. of partners	10.1	1	.002	23.7	1	.000	0.2	1	.663
Unprotected sex	4.5	1	.035	14.7	1	.000	0.0	1	.974
Masturbation	10.9	1	.001	7.1	1	.008	0.9	1	.341
Smoke Marijuana	30.0	1	.000	13.3	1	.000	0.2	1	.670
Bought/Sold Marijuana	6.1	1	.014	4.6	1	.032	0.0	1	.881
HIV tests	2.0	1	.155	9.9	1	.002	1.4	1	.245
Fake ID	21.1	1	.000	8.9	1	.003	0.2	1	.667
Freq. drink > 5 drinks	9.0	1	.003	9.6	1	.002	0.1	1	.716
Drunk	3.3	1	.070	7.8	1	.005	2.5	1	.116
Multivariate (<i>n</i> = 353)	5.3	9	.000	4.1	9	.000	1.24	9	.269

Table 5
Statistical Effects of Question Threat and Erotophilia on Behavioral Self-report: Pearson Correlations and Logistic Regressions

Behavior	Spearman Correlations		Hierarchical Logistic Regressions controlling for Condition and Gender						
	PQT		Perceived Question Threat			Erotophilia			
	r	SOS	B	SE	p	B	SE	p	$\Delta\chi^2$ 3 rd Step
Multiple partners	.24***	.23***	.20	.05	.000	.02	.01	.010	6.8**
Unprotected sex	.36***	.20***	.34	.05	.000	.03	.01	.000	19.0***
Masturbation	.11*	.59***	.30	.06	.000	.09	.01	.000	88.6***
Snok Marijuana	.39***	.15***	.57	.09	.000	.02	.01	.003	9.2**
Trade Marijuana	.41***	.19***	.41	.06	.000	.03	.01	.000	16.2***
HIV tests	.13**	.04	.11	.05	.026	.010	.007	.185	1.8
Faked ID	.29***	.12*	.34	.07	.000	.02	.00	.004	8.6**
Times > 5 drinks	.09*	.25***	1.58	.67	.018	.04	.01	.000	18.0***
Times drunk	.07	.21***	.43	.17	.013	.04	.01	.000	17.4***

* $p < .05$;

** $p < .01$;

*** $p < .001$;

Note: Bivariate Correlations were performed with frequency reports; Regression coefficients refer to results in the last (3rd) step of the logistic regressions, $\Delta\chi^2$ refers to the contribution in the step in which the predictor was entered. PQT = Perceived Question Threat; SOS = Sexual Opinion Survey (Erotophilia)

Table 6
 Matrices of Polychoric and Polyserial Correlations Prepared by Prelis 2

Path Model 1	1	2	3	4	5
1	1.00				
2	No. of PNR Responses	1.00			
3	Question Threat	0.36	1.00		
4	SOS	-0.02	-0.16	1.00	
5	Treatment	0.59	0.34	-0.08	1.00
	Gender	-0.11	-0.34	0.38	-0.03
Path Model 2	1	2	3	4	5
1	1.00				
2	No. of behaviors enacted	1.00			
3	Question Threat	0.14	1.00		
4	SOS	0.38	-0.18	1.00	
5	Treatment	-0.27	0.17	-0.09	1.00
	Gender	0.16	-0.31	0.40	0.11
Path Model 3	1	2	3	4	5
1	1.00				
2	Behavior Freq. Composite	1.00			
3	Question Threat	0.02	1.00		
4	SOS	0.25	-0.18	1.00	
5	Treatment	-0.25	0.17	-0.09	1.00
	Gender	0.27	-0.31	0.40	0.11