7-25-2013

Measuring the impact of inquiry mode above and beyond situational characteristics and experimenter contact in research relating to self-reported sexual attitudes and behaviors

Ethan B. McCallum
University of Missouri-St. Louis

Electronic Dissertation Project Title Page

Measuring the impact of inquiry mode above and beyond situational characteristics and experimenter contact in research relating to self-reported sexual attitudes and behaviors

Ethan B. McCallum
M.A., Clinical Psychology, University of Missouri – St. Louis, 2010
B.A., Psychology, College of St. Benedict / St. John's University, 2007

A Dissertation Submitted to The Graduate School of the University of Missouri – St. Louis in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Clinical Psychology

June 2013

Advisory Committee

Zoë Peterson, Ph.D.
Chairperson

Kristin Carbone-Lopez, Ph.D.

Steven Bruce, Ph.D.

Matthew Taylor, Ph.D.

Table of Contents

Acknowledgments

I would like to first thank my advisor, Dr. Zoë Peterson, for everything that you have done to help me to get to this point. Your guidance, mentorship, caring, and support over the course of my time at UMSL has truly been a defining experience for me. Not only have you helped me stay on track with the many hurdles and pitfalls which have arisen, but you have shown me so much about being a researcher, a psychologist, and a person. I am extremely fortunate to have had the opportunity to work with you, and I know that I will continue to benefit from this experience for the rest of my career.

I would also like to thank my committee members, Dr. Carbone-Lopez, Dr. Bruce, and Dr. Taylor for your thoughtful consideration of the many complexities of this project, and the supportive suggestions you provided. You have all helped me to move from a vague idea to a finished product.

Thank you also to the faculty, staff, and students of the UM – St. Louis psychology department. You all have made meaningful contributions to my development over the last four years, and I could not have done this without you. I would also like to give a special thanks to my cohort-mates, Brianna Mann, Shannon McNeil, Tiffany Mueller, and Diane Rosenbaum. You have all been right there with me every step of the way, even after we spread out all over the country. I am sure that I will continue to lean on you for wisdom, support, and comic relief in the years to come as well.

Much thanks as well to my research assistants, William Dement, Sarah Gamblin, Krystal Genteman, and Sarah Kirchhofer. Obviously, none of this would have been

possible without your help. I am particularly grateful to William for repeatedly going "above and beyond" in order to keep things running smoothly.

Last, but certainly not least, I must extend a special thank you to my wife Maggie and my daughter Liesl. The sacrifices that you have both made in support of my graduate work and this manuscript in particular have been nothing short of tremendous. Thank you both for putting up with all of those nights and weekends when I was either gone or distracted. Thank you both for your patience and support throughout this long, long process. You have both been my link to the real world, and you've always keep me grounded. Words cannot express how appreciative I am for everything that you have done.

Abstract

Factors related to the research context such as inquiry mode, experimenter contact, and setting may affect participants' comfort with and willingness to admit to engaging in sensitive sexual behaviors or to disclose certain sexual attitudes. Three-hundred-and-thirty-seven undergraduates (261 female, 41% non-White) completed a survey containing measures of sexual behavior, attitudes, sexual victimization, and sexual perpetration history. The level of experimenter contact (high vs. low contact), setting of completion (in lab vs. out of lab), and inquiry mode (pencil-and-paper vs. computer) were manipulated and participants were randomly assigned to one of eight conditions

I hypothesized that low contact, out of lab, computer conditions would be associated with the highest rates of reported sexual behaviors (including higher frequencies, a wider variety of behaviors, and higher rates of reported victimization and perpetration). I also predicted that these same experimental conditions would be associated with more liberal attitudes towards sex and sexuality. Further, I hypothesized that these effects would be moderated by race, such that differences across conditions would be greater for non-White participants than for White participants because non-White participants might fear that reporting socially undesirable sexual behavior will fuel racial stereotypes.

For female participants, a general pattern emerged across sexual behavioral measures suggesting that mode interacts with race to impact responding: Non-White women tended to report more sexual behaviors on pencil-and-paper surveys than on

computers. White women either demonstrated no mode-related differences or reported more sexual behaviors in computer conditions than in paper-and-pencil conditions. One exception was sexual victimization, with White women reporting more victimization on pencil-and-paper measures than on computer. For attitudinal measures, experimenter contact tended to be the most important experimental variable, though effects were again moderated by race. White women endorsed more liberal attitudes towards sex in high contact conditions, and non-White women endorsed more liberal attitudes in low contact conditions. Evaluation of differences for men was hampered by a small sample of male participants. Overall, these results suggest that methodological factors such as experimenter contact and mode have a significant impact on sexual self-report and the direction and magnitude of impact is often moderated by race.

Measuring the impact of inquiry mode above and beyond situational characteristics and experimenter contact in sexual behavior self-report research

INTRODUCTION

Findings from modern sex research are frequently applied to a wide range of efforts, including prevention and prosecution of sex crimes and controlling the spread of sexually transmitted infections (STIs) such as HIV. Sexual behavior research also provides important insights into positive sexual functioning across the lifespan. An awareness of private sexual behaviors being engaged in and the attitudes being espoused is vital information for researchers, practitioners, and policy makers alike. However, the extent to which sex research can truly inform the aforementioned efforts is largely dependent on the degree to which the data being collected reflect the behaviors and attitudes which occur outside of plain sight. Obtaining accurate information is a central challenge within the field of sex research.

Historically, sex research has been collected predominately through various retrospective self-report methods (e.g., face-to-face interview, pencil-and-paper questionnaire). One major drawback to reliance on self-report is that the quality of the data is dependent on participants' ability and willingness to accurately report their past sexual behavior and current attitudes. Further complicating matters is the inability of researchers to verify the information provided. Many efforts to improve the quality of sexual behavior research have been aimed at identifying alternatives to self-report, such as biomarkers (e.g., semen presence in women's urine samples; Langhaug, Sherr, & Cowan, 2010). However, most of the alternatives identified have limited applicability, are

invasive, and often result in high rates of false positives. Similarly, direct observations of sexual behavior as a means of verification are impossible due to ethical and practical restrictions. Further, biomarkers and direct observation do nothing to capture information related to the attitudes being held by individuals as they engage in various sexual behaviors. As a result, self-report remains the primary means for measuring sexual behavior and attitudes, leaving researchers with the task of evaluating the accuracy of self-reports while acknowledging the bias likely to be present within all responses.

Accuracy and Bias: Evaluating Sexual Behavior Research

One of the most challenging issues in measuring sexual behavior is addressing the issue of accuracy. Researchers often operate on the "more is better" principle, assuming that measures or techniques that elicit higher rates of reported sexual behavior are getting closer to the actual rates at which the behavior took place (Tourangeau & Yan, 2007). This assumption is based on the observation that sexual behavior is personal, private, and sometimes socially unacceptable or embarrassing, and as such participants are presumably more likely to underreport than over-report behaviors (e.g., Catania, Gibson, Chitwood, & Coates, 1990; Gillmore, Leigh, Hoppe, & Morrison, 2010). However, it is important to note that any comparisons made between retrospective reports are estimates and cannot establish true accuracy.

As previously noted, self-reports of sexual behavior are dependent on participants' ability to accurately remember those behaviors and their willingness to accurately report them to researchers, dependencies which introduce the potential for bias. In the absence of a well-established point of comparison, the degree of bias present

within self-reported sexual behavior is ambiguous and debated. Researchers have employed a range of different strategies to evaluate the accuracy of self-reports, and in turn have drawn a range of conclusions.

One method for evaluating retrospective self-reports of behavior is to compare them to daily diaries. Daily diaries differ from traditional self-reports in that participants are asked to complete daily reports on specific behaviors rather than waiting an extended period of time before reporting. Daily diaries are often used as a point of comparison based on the assumption that immediate recall of specific behaviors will be more accurate than trying to think back over periods of months or years. In one study using daily diaries as a point of comparison for retrospective self-reports, there was a 31% mean difference observed for lifetime number of sexual partners (e.g., McAuliffe, DiFranceisco, & Reed, 2007). The size of this discrepancy led the study's authors to question the degree of confidence typically placed on traditional retrospective self-reports.

However, based on a systematic comparison of seven large scale US population-based studies of sexual behavior, a very different conclusion was drawn. Given that each of the seven studies tapped into the same population, consistent results across studies would lend support to the reliability of self-report measurement. In spite of substantial variability in the methodology (e.g., question wording, mode of inquiry) implemented across studies, the authors found "remarkable levels of consistency" between studies in self-reported lifetime sexual partners, leading them to conclude that concerns related to bias in self-reports are largely overblown (Hamilton & Morris, 2010).

The central driving force behind such polarized assessments of self-report data may be absence of an agreed upon method for assessing the quality of self-report data being collected. For example, even if retrospective self-report data was found to closely correspond to daily diary data, this would only provide evidence that research participants are able to remember past behavior with relative accuracy; it would not provide evidence that they are reporting those behaviors honestly. Similarly, though consistency can be observed across studies, there is no way to evaluate the degree to which participants are being consistently inaccurate.

Though eliminating cofounds entirely is impossible, steps may be taken to minimize their impact. One effort to improve the quality of self-report data has focused on identification of factors within a study's methodology that may contribute to biased responding. One such methodological factor may be inquiry mode, or the means by which researchers query participants about their sexual behavior.

Inquiry mode

Self-report modes of inquiry elicit responses from participants about thoughts, behaviors, or experiences that have happened in the past. However, the means by which these responses are elicited varies greatly. Classically, researchers had few self-report modes of inquiry from which to choose (Knapp & Kirk, 2003). Perhaps the first option considered for self-report data collection was a face-to-face interview, in which a researcher would sit down with a participant and ask them questions. For a long time, the only available alternative to the face-to-face interview was a pencil-and-paper

questionnaire, in which participants were prompted by the text on a page to provide responses.

Modern technology has since provided a number of alternative iterations of the classic inquiry modes. Interviews can now be conducted over the phone, removing a participant from a face-to-face interaction. Further, pencil-and-paper surveys are now frequently replaced by computer-based questionnaires, which might still be completed on site at a research facility or alternatively in a location of the participant's choosing, using the internet as a means of data collection. Another option is the computer assisted self-interview (CASI), which allows for a computer to replace the human interviewer and actively question participants, often using a prerecorded audio component to negate any literacy requirements (Knapp & Kirk, 2003). More recently, mobile technologies such as Personal Digital Assistants (PDAs) and smartphones have provided researchers with another convenient alternative to the pencil-and-paper questionnaire (e.g., Vannier & O'Sullivan, 2008). All of these modern inquiry modes share many common elements with the more traditional forms of self-report, but they also each contain unique elements as well. These unique elements introduce a great deal of methodological variability into self-report research.

It is possible that the methodological variability, which exists across the diverse range of available inquiry modes, may introduce systematic bias into sexual behavior research, and participants may be more likely to provide accurate information in some conditions than in others. This potential can be framed as a problem, in that interpretation of self-report research collected across a variety of inquiry modes may become even more unclear and difficult to interpret. More optimistically, the potential for differences

in bias across inquiry mode also implies that a better understanding of the factors that promote accurate reporting may provide researchers with an opportunity to identify a methodology that minimizes bias. There has been some indication that several important sources of bias may indeed be impacted by the mode of inquiry through which responses are obtained. However, results have been mixed, and in spite of decades of research, little consensus has been reached as to the degree, if any, that mode of inquiry plays in impacting self-reported sexual behavior and attitudes.

Sources of Bias in Self-Reported Sexual Behavior in Relation to Mode of Inquiry

In spite of the lack of a consensus as to the degree to which distortion within self-report data can be attributed to mode of inquiry, it has been linked to several important and well-studied sources of bias. Factors such as imperfect recollection, social desirability, and participation bias all have the potential to skew the quality of the data being obtained, and all have been linked to inquiry mode. Recollection or the ability of a participant to accurately remember the frequencies of behaviors over a variable span of time is one of the major potential sources of bias. Another important source of bias is social desirability or an individual's motivation to be viewed in a favorable light. A final issue is participation bias. Much like social desirability, participation bias relates largely to the social pressures regarding sexual behavior. Many individuals may avoid specific questions about sexual behavior or may avoid studies relating to sexual behavior entirely.

Accurate Recollection of Sexual Behavior

Sex researchers routinely ask participants to look back over long periods of time, sometimes a participant's entire lifespan; a challenging task that requires participants to compute the frequency of specific behaviors, such as unprotected vaginal intercourse or unique sexual partners. Inevitably, some participants make mistakes in recollection. However, the degree of this memory bias is difficult to determine without a more reliable point of comparison.

Though some degree of memory bias is unavoidable, there is some indication that the mode of retrospective self-report can modestly impact consistency between diaries and later reporting of sexual behavior. One such example comes from a study examining adult sexual behavior using several different modes of retrospective self-report, including pencil-and-paper questionnaire, CASI, and audio-enhanced CASI and comparing them to reports made using a daily dairy technique (McAuliffe, et al., 2007). The results from this study indicated that participants in the CASI conditions made retrospective reports that were somewhat more consistent with daily diaries than participants in the pencil-and-paper questionnaire condition, suggesting that participants in the computer conditions may have been less impacted by memory bias than those in the pencil-and-paper condition. More research is needed to confirm these findings and expand them to other modes of self-report data collection.

Research to date provides strong support for recall as an important factor in the degree to which sexual behavior is accurately reported but has not advanced far enough to establish specific methodological guidelines to minimize memory bias. There is some

indication that computer based modes of inquiry increase motivation and that motivated recall may ultimately yield modest improvements in the accuracy of participants' reporting. Clearly, more research is needed to make any conclusions about the impact of inquiry mode on recall and to expand the recall literature to examine other modes of administration (e.g. internet-based survey).

Social Desirability and Self-Reported Sexual Behavior and Attitudes

Social desirability generally refers to an effort by participants to be evaluated favorably. Researchers have long been concerned about the impact of social desirability on the content participants are willing to report, with concerns dating as far back as the early 1930's when personality assessors questioned participants' efforts to portray themselves in an overly favorable light (Bernreuter, 1933 as cited in Meston, Heiman, Trapnell, & Paulhus, 1998). Social desirability is a particular concern for the measurement of sexual behaviors and attitudes, which are typically kept private and are rarely disclosed to strangers. The concept of social desirability has been further broken down by researchers recognizing that participants, not only aim for positive evaluation by others, but strive to protect their own self-image as well (e.g. Paulhus, 1984). Impression management refers to participants' efforts to tailor their responses in such a way as to maintain or project a pro-social image to others who may be viewing the results. This is contrasted with self-deception, which is conceptualized as an unconscious effort by participants to respond in an overly favorable way in order to protect or inflate their self-image (Paulhus, 1984). Much of the research relating to the impact of self-report mode has focused on impression management as a possible motivation for editing responses to match societal expectations, particularly in modes which directly involve interaction with

a human experimenter in the data collection process (e.g. Richman, Weisband, Kiesler, & Drasgow, 1999; Testa, Livingston, & VanZile Tamsen, 2005; Wood, Nosko, Desmarais, Ross, & Irvine, 2006). Though self-deception is less frequently studied, and conceptually less clearly linked to mode of inquiry, it is also possible that self-deception plays a role in participants' responding. As impression management is the source of social desirability bias most well-studied to date, it requires a thorough review to establish possible mode of inquiry related considerations that might be made.

Impression management. The vast majority of research relating to the relationship between mode and social desirability has been focused on impression management. Theoretically, individuals reporting details about their sexual behavior or attitudes to an interviewer sitting across from them may be more likely to engage in impression management than those completing a pencil-and-paper or computer-based questionnaire in private. This theoretical expectation has been supported by an accumulation of research suggesting that participants' responding to behavioral measures score higher on measures of socially desirable responding in face-to-face interviews than those completing computer-based questionnaires (Richman, et al., 1999). Findings have been less consistent when computer-based surveys are compared with pencil-and-paper surveys, as some studies have suggested differences in social desirability, and others have suggested equivalence (Weigold, Weigold, & Russell, 2013).

There are a number of factors that have been identified which make impression management efforts more likely. One important factor is the type of question being asked. Questions that relate to sensitive information and tap into gender or cultural roles or some form of stigma are more likely to elicit motivated "editing," or impression

management efforts (Tourangeau & Yan, 2007; Kays, Gathercoal, & Buhrow, 2012). This is further amplified when participants view questions as intrusive or have concerns about possible negative repercussions for disclosing sensitive information to researchers. The relationship between these factors clearly supports the possibility of an inquiry mode dependent effect in sexual behavior research. Sexual behaviors and attitudes are considered to be private and are typically tied to both gender and cultural values. Further, many people do see questions about sexual behavior and attitudes to be somewhat intrusive and may be concerned about their responses being made public. Some researchers have suggested that the perceived level of intrusiveness and or threat of disclosure may vary across self-report mode.

Research related to sexual behavior and inquiry mode is mixed but does show support for a possible mode-dependent effect which may impact the reporting of some behaviors but not others. For example, in the aforementioned population-based review, little variation was observed across inquiry mode for questions related to lifetime partners, suggesting that perhaps the question may not be as sensitive as previously thought or that sensitivity of the question may not be the only or primary determinant of impression management (Hamilton & Morris, 2010). In contrast, one study examining a wider range of behaviors provides some indication that mode-dependent differences may exist for some behaviors, such as unprotected oral sex and recent sexual partners, with more of these specific behaviors being reported via anonymous CASI conditions than self-administered pencil-and-paper questionnaires; this relationship was not found for other behaviors, such as lifetime sexual partners (Brown & Vanable, 2009). It is possible that topics such as multiple sexual partners may be too commonly discussed or widely

experienced to evoke significant impression management efforts. It appears that there is a threshold of sensitivity that, when crossed, leads participants to engage in impression management at higher rates for some inquiry modes than others.

Environmental factors that vary across mode present another possible mechanism through which social desirability might operate. Differences in inquiry mode may impact important factors such as proximity to the experimenter and degree of anonymity (or the participants' perception of anonymity). These factors may in turn be an important source of systematic variability in responding. Accumulating research suggests that mode itself is not sufficient to predict socially desirable responding but may interact with other factors such as question content or presence of others to impact distortion efforts (Richman, et al., 1999). Such findings are in keeping with social desirability theory and support the possibility of inquiry mode-dependent distortion effects in sexual behavior research. One finding which has consistently emerged is that computer-dependent modes of collection yield lower rates of distortion than face-to-face interviewing (Brown & Vanable, 2009; Langhaug, et al., 2010; Richman, et al., 1999). Though a meta-analysis of impression management and mode of inquiry research revealed no significant overall difference between computer-based and pencil-and-paper questionnaires, with consideration of moderators, participants completing computer administered questionnaires scored significantly lower on measures of socially desirable responding, than those completing pencil-and-paper questionnaires, suggesting less distortion in their responses (Richman, et al., 1999). Specifically, when participants were alone and were able to skip questions and backtrack, they showed less distortion in computer based conditions. Other studies have supported the finding that participants are more candid

when responding to computer-based questionnaires than face-to-face interviewing or pencil-and-paper formats (Feigelson & Dwight, 2000). However, it has been argued that observed differences between various modes of inquiry may be due to in part to methodological variability across conditions. One study sought to compare pencil-and-paper surveys with computer-based surveys after carefully ensuring equivalence in experimenter contact and setting of completion for both conditions (Weigold, et al., 2013). In this study, no differences were observed between pencil-and-paper and computer based surveys on measures of personality or social desirability, suggesting that differences that have previously been observed may be related to differences in level of experimenter contact, or setting of survey completion.

Researchers also have sought to identify the mechanisms contributing to inquiry mode-dependent distortion. One study examining participants' responding through pencil-and-paper, onsite computer-based, and internet questionnaires revealed a number of notable differences in participant's perceptions (Bates & Cox, 2008). Participants reported a variable perception of anonymity across conditions, tending to report higher rates of perceived anonymity in computer based administration conditions, and they also reported a belief that the accuracy of their responses varied across inquiry modes as well, with higher rates of perceived anonymity being positively associated with perceived accuracy. Interestingly, in spite of their self-perceived distortion in some conditions, no significant differences were observed in the behaviors participants reported across conditions. This inconsistency highlights the complexity of the mixed findings relating to mode of inquiry and impression management. Participants themselves seem unclear of the degree to which inquiry mode impacts their responding.

Though impression management has received a fair amount of attention in mode of inquiry research, the majority of studies specifically related to sexual behavior have focused on sexual partners, masturbation, and vaginal intercourse (Catania, et al., 1990). Reviews of sexual behavior methodology have also suggested that the majority of research has been conducted with college students, and more diverse community populations have been largely ignored (Weinhardt, Forsyth, Carey, Jaworski, & Durant, 1998). This is a notable limitation of current literature in that research focusing largely on common sexual behaviors engaged in by majority populations (i.e., White, middle-class populations) is far less likely to capture mode-dependent differences than research on less common behaviors or a minority population's behaviors, as the social pressure for conformity is less in the former than the latter. Further research is needed in order to determine the impact of these factors on impression management in sexual behavior reporting.

Self-deception. There is some indication that certain types of sex-related questions are more likely to activate self-deceptive efforts than others. For example, it has been found that participants who score highly on measures of self-deception also are likely to provide an overly positive view of their sexual adjustment, likely in an effort to maintain the belief that they are sexually well-adjusted (Meston, et al., 1998). However, this study did not find any relationship between reports of specific sexual behaviors and self-deception efforts.

At this point, very little research exists regarding the relationship between mode of inquiry and self-deception efforts. However, there is some conceptual justification for such a relationship. It has been suggested that individuals who have more perceived

control over a situation may be less motivated to protect themselves with deceptive efforts (Fox & Schwartz, 2002). Such a relationship would predict less self-deception efforts in more independent collection modes, which afford participants a greater deal of control. An examination of this hypothesis using pencil-and-paper surveys along with computer-based questionnaires found no significant differences across mode for a measure of self-deception (Fox & Schwartz, 2002). However, other studies have found differences between group-administered pencil-and-paper surveys and computer-based or individually-administered questionnaires, with the individually-based administration yielding higher rates of self-deception (Lautenschlager & Flaherty, 1990). This seems inconsistent with the idea that self-deception should be lower in situations involving greater perceived control. One possible way to interpret these results is that self-deception is more likely when questionnaires are completed independently of social contact and plays less of a role when other participants or evaluators are immediately present, as this latter condition may shift an individual's focus from self-evaluation (i.e., self-deception) to social evaluation (i.e., impression management).

Participation Bias

Participation bias refers to the systematic decision by certain types of individuals to seek out or avoid participation in a study (Catania, et al., 1990). This type of self-selection can be global (i.e., unit response bias), in that certain individuals may avoid participation entirely, or localized (i.e., item response bias), with participants declining to provide responses to specific items. These two forms of participation bias are problematic in sexual behavior research due to the private nature of such behaviors, the social connotations of "sexual research," and the perceived intrusion of specific questions about

sexual behavior. Ideally, sexual behavior researchers successfully recruit and retain highly representative samples of a target population and achieve high response rates from those participants in order to minimize concerns about biased results as a function of participant self-selection.

Unit response rates. Survey-based research, such as that most commonly conducted in sexual behavior studies, is largely dependent on contacting individuals to solicit participation in a study. Depending on the mode of data collection, participation may involve showing up on-site to complete participation, mailing back a survey, logging on to a computer, or answering a telephone. The unique demands of various modes may give certain types of participants various incentives or disincentives to participate and may have an impact on which individuals ultimately agree to participate.

There are indications that modes of inquiry have a substantial impact on the degree of unit non-response. An accumulation of research suggests that web-based surveys yield 11% lower response rates than other modes of data collection such as on-site interviews or pencil-and-paper questionnaires (Manfreda, et al., 2008). There are also indications that these disparities grow larger when web-based participation is solicited through non-computer based methods such as postal mail.

Concerns with lower rates of participation in web-based research are tempered by research reviews indicating that web-based samples are typically more representative than traditional samples in respect to gender, socioeconomic status, geographic location, and age (Gosling, Vazire, Srivastava, & John, 2004). Further, web-based samples appear to be relatively equivalent to traditional samples in regard to race. It is also worth noting

that though early critics of web-based research suggested that participants may have been particularly psychologically dysfunctional or maladjusted, a number of studies provide evidence which counters this assumption (Gosling, et al., 2004).

More pertinent to sexual behavior research is the relationship between response rates to internet surveys and the type of question being asked. It would be problematic for online sex research if a certain subset of the population refused to participate in online sex research due to concerns about providing information about their private behaviors through such a medium. However, if people decide not to participate in online survey research regardless of question content, due to disinterest or some other nonsystematic factors external to the survey topic, the concerns for sex researchers, specifically, might not be as great. There is some indication that response rates to web-based surveys are not significantly related to question sensitivity, suggesting that individuals do not appear to be self-selecting out of studies to avoid answering sensitive questions (Cook, Heath, & Thompson, 2000).

Though participants do not seem deeply concerned with question sensitivity, there are several factors which impact a decision to participate in web-based research. Two factors that appear to be particularly important in the decision-making process are saliency and confidentiality (Tourangeau & Yan, 2007). In other words, participants are more motivated to participate in a web-based study if they believe the topic is relevant to them as individuals and are reasonably confident that their confidentiality will be protected. It is not clear to what extent these factors are specific to participation in web-based research; relevance and confidentiality are likely to be important considerations for individuals who are invited to participate in other modes of data collection, too. Though it

seems unlikely that mode of inquiry would have much of an impact on perceived relevance, it is reasonable to expect differences across mode in terms of perceived confidentiality (or even anonymity).

Item response rates. Though obtaining a diverse group of participants is a challenge in sexual behavior research, it is equally challenging to ensure that participants answer questionnaires completely. Participants who begin surveys frequently omit responses to certain items or discontinue prior to completion. If missed items or discontinuation points are systematic, this type of behavior can lead to biased results.

There is some indication that participants are more likely to omit answers or to provide "zero" or "never" responses to questions about atypical sexual behaviors (e.g. sexual violence, extramarital sex) than questions about more common sexual behaviors (Catania, et al., 1990). This latter tendency is particularly problematic as zero responses, unlike omitted responses, leave researchers with a difficult decision regarding the interpretation of the data. Whereas some participants may have genuinely never engaged in an infrequent or uncommon behavior, others may endorse a "never" response in an effort to comply with social demands or to protect sensitive personal information.

There appears to be some evidence of the impact of inquiry mode on item non-response rates in sexual behavior research. Pencil-and-paper questionnaires containing items relating to specific sexual behaviors have been shown to yield significantly more omissions than otherwise identical online questionnaires (e.g., Wood, et al., 2006, Kays, et al., 2012). Further, participants tend to skip more items towards the end of pencil-and-

paper conditions, suggesting that participant fatigue may be a greater concern in pencil-and-paper modes than in computer-based collection modes (Wood, et al., 2006).

Another factor which plays a role in item non-response is motivation level. Participants with low levels of motivation are not likely to answer surveys completely or carefully. Variable levels of motivation across inquiry modes are likely to lead to greater levels of item nonresponse in some modes than others, particularly under certain conditions such as exposure to sensitive item content. More research is needed to determine whether systematic differences exist in motivation levels across various modes of inquiry.

Overall, it is clear that the individuals who agree to participate in research and the questions these individuals agree to answer play a central role in the type and quality of the data obtained. In sexual behavior research, there are always concerns surrounding who is agreeing to participate and the degree to which these participants answer all questions fully and honestly. Such concerns call into question the generalizability of sexual behavior data to the wider population. There is some evidence that internet-dependent studies have lower response rates than more traditional studies, increasing the possibility for a selection bias. However, this concern is offset to some degree by indications that internet studies are highly representative. Research also suggests that participants are more likely to omit responses to sensitive questions about sexual behavior and that these omissions are more frequent in some data collection modes than others. This finding presents another possible route for inquiry mode dependent differences to emerge. More research is needed to understand why differences exist

between inquiry modes and what specific motivations participants have for opting out of a given study or omitting a response.

Problems with Current Literature

Much of the sex research related to inquiry mode has sought to compare the responses of two or more groups on identical questions while varying the mode of inquiry through which the questions are presented (e.g., McAuliffe et al., 2007; Morrison-Beedy, Carey, & Tu, 2006; Wood et al., 2006). Any observed differences in these studies are typically concluded to be the result of differential impact across inquiry modes. However, this type of design is problematic in that much of the variability presented across condition may not be the direct result of the inquiry mode itself but may instead relate to the methodological variability existing across conditions (e.g., Weigold et al., 2013). For example, when comparing an internet based inquiry to a traditional pencil-and-paper survey completed in a lab, researchers are varying mode (i.e., internet versus paper), but are also varying the degree of experimenter contact and the environment in which participants complete the survey. It is possible that methodological variability related to factors such as experimenter contract may be equally or more impactful on participants' responses as the inquiry mode by which questions are presented. Though it could be argued that experimental contact and environmental factors truly are components inherent to specific modes of inquiry, these factors are rarely held constant across studies and may be a contributing factor in the lack of consistency observed in inquiry mode research.

It is also important to note that studies focused only on examining mode effects across specific inquiry modes will quickly become obsolete as a result of rapid advances

in computers, software, and telephone-based survey technology. This can already be seen in early inquiry mode research aimed at evaluating emergent computer technology, which lacked the customizability and ease of use of modern computers. Such drastic technological changes make comparisons across modes of data collection highly problematic. In order to truly advance the understanding of inquiry mode effects, researchers must begin to be more mindful of the common underlying factors that are present but variable across inquiry modes.

Experimenter contact

Conceptually, the degree to which a participant must interact with a human experimenter is likely to play an important role in their motivation for impression management. In keeping with this expectation is the consistent finding that participants typically report lower rates of sexual behavior in face-to-face interview than other modes of inquiry (e.g., Brown & Vanable, 2009; Feigelson & Dwight, 2000; Langhaug et al., 2010). Outside of face-to-face interview, the picture is less clear what role experimenter contact plays in participants' responding, partly because it is a variable that is rarely manipulated systematically. One study attempting to examine experimenter presence compared participants' responses to sensitive questions (including questions related to sexual behavior) in two different conditions (Wood et al., 2006). In the first condition, participants completed a measure online in a university computer lab in the presence of other participants and a supervising experimenter. In the second condition, participants were free to complete an equivalent measure online whenever and wherever they chose. Though the study did not find any differences in reporting between conditions, the design

prevented direct interpretation of experimenter effects independent of the presence of other participants, or the differing situational characteristics between conditions.

Though lab based conditions present the clearest, observable interaction between participant and experimenter, there is arguably some degree of interaction in all conditions. Even within internet research, participants are exposed to an informed consent page which typically lists identifying information such as the investigator's name and the institution to which they belong. Further, participants are generally aware that, though they are providing responses to a computer, these responses are ultimately received and processed by another human being. This "virtual experimenter" effect is often overlooked within internet based survey research, but may play some role in participants' responding (e.g., Ollesch, Heineken, & Schulte, 2006).

Situational characteristics (e.g., setting)

Another factor not often accounted for within inquiry mode research is the situational characteristics which vary across inquiry mode. Modes such as phone-based or internet based survey leave researchers with little control over situational characteristics, and little ability to measure those characteristics. This lack of control was demonstrated in a recent study which allowed participants to complete pencil-and-paper or computer based surveys in any location of their choosing, and followed up by asking them where they had elected to do so (Hardré, Crowson, & Xie, 2012). Participants reported completing surveys in a wide range of settings, with a range of potential distractions. The study found that participants completing computer based surveys were more likely to do so in the presence of social distractions (e.g., friends, family, in the midst of a lecture).

Factors such as the setting in which participants elect to complete a survey or the presence of others when completing the survey are likely to impact participants' responding. This is contrasted with lab based conditions in which experimenters have nearly complete control over the situational characteristics. Studies that have sought to place more control over setting of completion have not found any significant differences across modes (e.g., Weigold et al., 2013). However, situational characteristics also represent a portion of the methodological variability that is naturally present across inquiry modes, and rarely has a study sought to separate out the unique influence of this variability.

Limitations in the types of sexual behaviors assessed

Currently, there is limited and mixed research related to mode of inquiry and questions regarding nonconsensual sex, with existing studies generally focused on female victimization. There is some indication that women are more likely to disclose sexual assault related to alcohol use through a web survey than they are in a phone interview (Parks, Pardi, & Bradizza, 2006). Another study indicated higher rates of sexual assault disclosure by participants through pencil-and-paper inquiry than those observed in CASIs (Testa, et al., 2005). However, the latter study was limited by low response rates in the computer condition, which exemplify concerns about inquiry mode-dependent participation bias, as only 61.4% of contacted participants showed up for the computer condition in comparison to the 87.6% of participants who completed and returned pencil-and-paper surveys. Yet another study of sexual victimization disclosure found no difference in disclosure rates via computer-based, face-to-face, or pencil-and-paper

modes of inquiry, though participants did indicate a preference for computer-based reporting (DiLillo, DeGue, Kras, Di Loreto-Colgan, & Nash, 2006).

A number of studies have looked at intimate partner violence or sexual abuse perpetration as part of a larger battery of "sensitive topics" aimed at determining any inquiry mode related differences in disclosure rates (e.g. Hines, Douglas, & Mahmood, 2010; Reddy, et al., 2006; Rosenbaum, Rabenhorst, Reddy, Fleming, & Howells, 2006). None of these studies found significant differences in disclosure rates. However, two of the studies (Reddy, et al., 2006; Rosenbaum, et al., 2006) draw their samples from an undergraduate population at the same university and fail to provide statistics on actual disclosure rates short of mean scores for topic areas. Based on those mean scores, the participants in these studies appeared to have very low rates of disclosure for physical abuse perpetration and almost no disclosure of sexual abuse perpetration (Reddy, et al., 2006). Further, the Hines et al. study (2010) compared men's responses to an online questionnaire to responses obtained through a phone interview, preventing any conclusions regarding on-site inquiry modes.

Existing results indicate that mode may play a role in individuals' decisions to participate in research on sexual victimization and respond accurately to questions which prompt victimization disclosure responses. More attention is needed in this area in order to accumulate the literature necessary to make recommendations for both research and applied settings regarding methodological factors which maximize accurate disclosure. Further, more research is needed to understand the relationship between inquiry mode and disclosure of perpetration, as this topic has been all but ignored in existing literature,

with the handful of studies which have considered the topic being limited by design and far from conclusive.

Possible Moderating Factors for Inquiry Mode and Accompanying Methodological Covariates

Currently, there has been very little research aimed at identifying demographic factors that might sway the impact of inquiry mode or other methodological variables on self-reports of sexual behavior and attitudes. There are a number of factors which have the potential to impact participants' responding based on inquiry mode, experimenter contact, or situational factors. A brief overview of some these factors and the theoretical means by which they may impact responding will help to highlight the importance of considering mediating and moderating factors in this area.

Gender

Conceptually, there are a number of reasons why gender is a possible moderator that should be considered when examining inquiry mode-dependent impact on self-reported sexual behavior and attitudes. As previously mentioned, there is a long standing assumption within sexual behavior research that the more behaviors being reported, the closer researchers are to tapping into the "true" number of participants' behaviors. However, the different cultural expectations for men and women regarding sex often challenge this assumption. For example, due to a "sexual double standard" in many Western cultures, men reporting higher numbers of sexual partners or more casual sexual partners may be seen as more attractive or sexually accomplished than men reporting lower numbers of partners, whereas women reporting higher numbers of partners may be

seen as immoral or promiscuous (e.g., Crawford & Popp, 2003). These differing social expectations may lead women to underreport sexual behaviors and men to over-report them in an effort to be socially desirable (e.g., Schroder, et al., 2003; Smith, 1992).

Another related factor may be concordant vs. discordant gender in data collection modes which require interaction between participants and researchers. Participants who are interacting with a same gendered researcher may be more or less likely to edit their responses, depending on the type of question being asked, than participants interacting with a researcher of a different gender. This effect has been well-established for face-to-face interviews, with concordant gender pairs yielding higher rates of reported sexual behavior than gender discordant pairs (Catania, et al., 1996). Conceptually, any experimenter effects would likely be stronger in modes with higher rates of interaction, such as a face-to-face interview, and be less pronounced in modes with limited interaction, such as a web-based survey. There is some research to suggest that mode impacts men and women differently, with significantly less item non-response amongst men on web-based surveys, but no difference on pencil-and-paper administrations (Kays, et al., 2012). More research is needed to understand the moderation of inquiry mode effects by gender.

Race

Open discussions about sexual behavior and attitudes are viewed very differently by different racial groups (e.g., Langhaug, et al., 2010). While White Americans may be somewhat uncomfortable sharing information about their sexual behavior with strangers, the additional overlay of minority status may lead non-Whites to feel even less

comfortable sharing sensitive information, particularly if such information relates to atypical or socially undesirable behaviors.

Much like gender, the concordance or discordance of the examiner's and participant's respective race may play an important role in the degree of socially desirable editing in which the participant engages, and perhaps the degree of motivation behind remembering and thoroughly answering questions as well. Decreasing social distance (i.e., matching the researcher and the participant in terms of race) may lead to more candid reporting. Cultural allegiance associated with matched race also may increase participants' motivation, which has been linked to more accurate recall and higher completion rates (Morrison-Beedy, et al., 2006). However, this is contrasted by cultural conformity as a motivation, which may lead participants to edit their responses to be more in line with traditional cultural values when interacting with a researcher from a similar background. There is also a possibility that a White examiner will lead non-White participants to conform to the values of the majority culture, or cue racial stereotypes, leading participants to modify their responses to be more in line with majority values or stereotypes.

Race related factors are closely linked to sexual behavior and also have the potential to impact the interactions between participant and examiner in a number of different ways. This complex interaction likely leads participants to both over- and under-report behaviors depending on a number of factors. Though there is insufficient research to indicate a direction, it seems likely that minority groups experiencing higher levels of discrimination and prejudice will be more likely to edit their responses to increase social desirability.

Present study

The primary aim of this study was to advance the current understanding of mode of inquiry effects within sex research by isolating the unique impact of situational characteristics, experimenter presence, and mode of inquiry. As noted above, these three variables have never before been considered in parallel. As a result, any observed differences within mode of inquiry research have typically been problematically credited directly to the mode by which self-report information has been collected. By separately manipulating each of these variables researchers can better understand the degree to which inquiry mode impacts self-reports of sexual behavior directly. This study also aimed to examine the potential moderating role of gender and race on the relationship between mode of inquiry and self-report sexual behavior.

In order to assess the impact of experimenter presence, the level of experimenter contact was manipulated (high versus low contact). In order to assess the impact of setting, the place in which the questionnaire is completed was manipulated (in lab versus out of lab). Participants assigned to the in lab condition who also were assigned to high experimenter contact were greeted by the experimenter, who obtained informed consent and verbally oriented the participant to the pencil-and-paper or internet measure respectively. Participants assigned to the "out of lab" condition who had also been assigned to high experimenter contact condition were required to contact the experimenter by phone before beginning the survey and subsequently contact them again upon completion. During the initial conversation, the experimenter greeted the participant, obtained informed consent and oriented the participant to their assigned

measure verbally using the same script utilized in the in lab condition. This allowed for out of lab contact closely resembling the level of contact experienced in the lab.

For participants assigned to low experimenter contact conditions, interaction was limited to scheduling conducted via email. For the in lab condition/low contact condition, participants scheduled an appointment time via email contact with a nameless lab email account and arrived at a lab space that was not monitored by an experimenter. In all low-contact conditions, participants received printed instructions, which greeted them, provided them an informed consent form, and oriented them to the survey. These forms were adapted from the script used with participants in the high contact condition and contained the same content.

All participants were invited to complete two trials in order to allow for each participant to complete both the internet based form and the pencil-and-paper measure. These trials were separated by a period of roughly two weeks. Half of the participants completed pencil-and-paper measures during the initial trial, while the other half of the sample completed the internet based survey during the first trial. Thus, there were eight possible conditions (two experimenter conditions, two location conditions, and two orders of mode of inquiry) to which participants could be randomly assigned. This arrangement resulted in a 2 (high vs. low experimenter contact; a between subject variable) x 2 (lab vs. home completion; a between subject variable) x 2 (paper vs. internet; a within subject variable) mixed factorial design.

Hypotheses:

The most clear and consistent finding related to inquiry mode is that the experimenter contact demanded by face-to-face interviewing yields lower reports of sensitive behavior than more private modes such as computer based or pencil-and-paper inquiries (e.g., Richman, et al., 1999). Moving data collection out of the lab may give participants an even greater sense of anonymity, further facilitating open responding (e.g., Bates & Cox, 2008). Finally, there is some indication that participants feel more anonymous in computer based inquiry modes than more traditional modes such as pencil-and-paper (Bates & Cox, 2008). As such, I hypothesized that inquiry mode, experimenter contact, and setting would all uniquely impact participants' responding to questions related to disclosure of sexual behavior, sexual attitudes, and sexual victimization/perpetration. Further, I anticipated that these three methodological factors would impact participants' responding to a measure of social desirability. Specifically, I predicted that low contact, internet based inquiry mode, and out of lab completion would promote the reporting of more sexual behaviors, less conservative attitudes towards sex, higher rates sexual victimization and perpetration, and less social desirability than high contact, pencil-and-paper mode, and in lab completion, respectively. I also hypothesized that participants' perceptions of anonymity, confidentiality, and accuracy would mirror the predicted direction of reported attitudes and behaviors, in that participants would experience low contact, internet based inquiry mode, and out of lab completion as more confidential and anonymous than high contact, pencil-and-paper mode, and in lab completion, respectively, and in turn would believe their own responses to be more accurate in the former conditions than in the latter. Previous studies have supported this

pattern of perceptions, independent of real differences in reported behavior (Bates & Cox, 2008).

Additionally, I predicted that gender and race would moderate the degree of impact observed for each of the three methodological factors being studied. In terms of gender, I hypothesized that observed differences across conditions would be larger for women than for men, as there is some evidence that women are more prone to engage in socially desirable responding regarding sexual behaviors and attitudes than men (e.g., Alexander & Fisher, 2003). Similarly, I anticipated that White individuals would have less motivation to edit responses than other racial groups regardless of condition, because non-White individuals may fear that they will fuel prejudice or stereotyping due to race if they admit to socially unacceptable sexual behavior; therefore, I predicted that White participants would demonstrate less difference across conditions than individuals who are racial minorities. Specific hypotheses were as follows:

1. Low experimenter contact, out of lab completion, and internet based inquiry mode would be associated with higher rates of reported sexual behaviors (i.e., wider range of reported sexual activities and more behaviors associated with STI risk) than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode, respectively.

    1a. These effects would be moderated by gender, such that the differences across condition would be greater for women than for men.

1b. These effects would be moderated by race, such that the differences across conditions for non-White participants would be greater than for White participants.

2. Low experimenter contact, out of lab completion, and internet based inquiry mode would be associated with higher rates of reported sexual victimization (child sexual abuse [CSA], adult sexual assault) and perpetration disclosure than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode, respectively.

    2a. These effects would be moderated by gender, such that the differences across condition would be greater for women than for men.

    2b. These effects would be moderated by race, such that the differences across conditions for non-White participants would be greater than for White participants.

3. Low experimenter contact, out of lab completion, and internet based inquiry mode would be associated with more permissive or liberal attitudes towards sex than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode, respectively.

    3a. These effects would be moderated by gender, such that the differences across condition would be greater for women than for men.

3b. These effects would be moderated by race, such that the differences across conditions for non-White participants would be greater than for White participants.

4. Low experimenter contact, out of lab completion, and internet based inquiry mode would be associated with lower rates of socially desirable responding than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode, respectively.

    4a. These effects would be moderated by gender, such that the differences across condition would be greater for women than for men.

    4b. These effects would be moderated by race, such that the differences across conditions for non-White participants would be greater than for White participants.

5. Low experimenter contact, out of lab completion, and internet based inquiry mode would be associated with higher rates of perceived anonymity, confidentiality, and accuracy than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode, respectively.

Methods

Participant Recruitment

Recruitment began during the fall semester of 2011. Students enrolled in psychology courses during this semester had the opportunity to log in to a university subject pool portal with multiple studies that students could complete in order to get various amounts of extra credit for the courses in which they were enrolled. During the fall 2011 semester, 254 university students expressed interest in this study by providing basic contact information through the subject pool portal. Of the 254 students who expressed interest, 170 responded to emails requesting them to schedule an appointment to complete the initial portion of the study. In the spring 2012 semester, 247 students expressed interest in the study, 160 of them scheduled an initial appointment, and 130 attended those appointments to complete the survey. In an effort to bolster the sample, I also attempted to recruit students from psychology courses during the summer 2012 term as well. Though there was no formal subject pool running during the summer term, researchers went to several classes to provide students with information about the study, and instructors offered their students extra credit in these courses for participation. During the summer semester, 26 students expressed interest in the study, 15 scheduled an appointment, and 11 completed the initial survey. Participants were also recruited during the fall 2012 semester, in which 149 students expressed interest in the study, 101 scheduled appointments, and 73 completed the initial survey. Across the four collection periods, 667 students indicated interest in the study, 446 scheduled appointments, and 337 completed initial appointments (see table 1).

Table 1
Percentage Interested/Scheduled/Completed based on collection period

	Interested	Scheduled	%	Completed	%
Fall 2011	254	170	67	130	76
Spring 2012	247	160	65	123	77
Summer 2012	26	15	58	11	73
Fall 2012	149	101	68	73	72
Total	676	446	66	337	76

The low rates of scheduled appointments (66%) amongst individuals who initially expressed interest in the study were somewhat unexpected. Anecdotally, a number of students contacted researchers to withdraw from the study after discovering that they could not simply complete an online survey from their computer right after they logged into the subject pool online portal. Limitations on the categorical options for the subject pool portal required me to list the study as an "online study," which may have promoted this expectation. It seems plausible that many students who discovered additional steps to completion simply decided not to follow through with the study. Students who did take the additional step of scheduling an appointment attended their initial appointments at a reasonably high rate (76%).

Beyond contact information, no information or consent was collected from prospective participants until they attended their first appointment. As such, for the purposes of this study, I did not consider interested individuals as participants until they attended their first appointment. Additionally, the description for this study indicated that participants were expected to complete two surveys, approximately two weeks apart. Of the 337 participants who completed the initial survey, 113 returned to complete the follow up survey (34%). This percentage was also much lower than expected. Participants were asked to schedule follow up appointments immediately after

completing their initial surveys and were provided with up to two reminder emails if they did not schedule or attend follow up appointments. In spite of this, the majority of participants who completed the initial survey did not return to complete the follow up.

Participants

Participants for this study were 337 men and women between the ages of 18 and 62 (M = 23.43, SD = 6.80) recruited from a psychology subject pool at a Midwestern urban public university over the course of four semesters and a summer term (see Table 1). Seventy-two participants identified as male (21.4%), 261 identified as female (77.7%), and three identified as transgendered or "other" (0.9%). The participants who did not identify as male or female were removed from all analyses. Participants were provided with course credit as compensation for participation in the study.

Most of the participants identified themselves as White/European American (62.8%) or Black/African-American (31.8%). Participants also identified as Native American/Alaskan Native (3.6%), Asian/Asian American (5.7%), and "other" (3.6%). See table 2. Note that these categories were not exclusive, and participants were able to identify with multiple racial groups. For the purposes of analyzing racial differences, participants who identified with anything other than exclusively White/ European American were classified as non-White participants. Within this classification, participants who identified as bi-racial and multi-racial were considered non-White, even if White was one of the racial groups with which they identified; this was based on the assumption that a bi-racial participant would have similar concerns about negative stereotypes as other non-White participants. Based on these criteria, 58.6% of participants were classified as White and 41.4% were classified as non-White.

Table 2
Demographic information

		N	%
Gender	Male	72	21.4
	Female	261	77.7
	Transgendered/other	3	0.01
Race*	White/European American	211	62.8
	Black/African-American	107	31.8
	Asian/Asian American	19	5.7
	Native American/Alaskan Native	12	3.6
	Other	12	3.6

*Racial categories are not mutually exclusive

White and non-White participants were also compared based on a number of demographic factors (age, income, economic status). No significant differences were observed between the two groups on any of the factors examined (see table 3).

Table 3
Demographic factors as a function of racial group

	White	Non-White
Age; Mean (SD)	23.4 (6.5)	23.6 (7.3)
Income; N (%)		
$14,999 or less	50 (26%)	40 (29%)
$15,00 to $29,000	43 (22%)	26 (19%)
$30,000 to $59,999	43 (22%)	35 (26%)
$60,000 to $99,999	30 (15%)	19 (14%)
$100,000+	30 (15%)	17 (12%)
Relationship Status; N (%)		
Monogamous relationship	107 (55%)	74 (53%)
Non-monogamous / Dating	48 (24%)	37 (27%)
Not dating	41 (31%)	28 (20%)

Note. White and non-White participants did not differ significantly on any of the demographic factors examined.

One participant was removed from all analyses due to a failure to complete beyond the first page of the survey packet. Five additional cases were removed from all analyses due to experimenter errors which prevented proper identification of participants'

condition. Between zero and three cases were removed from specific analyses due to participants' failure to complete the majority of items on the specific measures being used in these analyses. This left between 328 and 331 participants available for each of the analyses.

Measures

Demographic measure. Basic demographic information was collected using a 15-item Demographic Questionnaire developed for this project which included basic questions about sex, age, year in school, religious affiliation, income, race, ethnicity, country of birth and relationship status.

Sexual behavior. Sexual behavior was assessed using two measures aimed at tapping into multiple domains. A general measure of sexual behaviors (SBM) was used to assess a variety of behaviors ranging from very socially acceptable behaviors (e.g., "kissing someone of the other sex on the mouth") to less socially acceptable behaviors (e.g., anal sex, group sex, and use of dominance or mild consensual force with same and other sex partners). This 29-item measure was designed for this study, but items were based on existing measures of sexual behavior (e.g., Browning, Hatfield, Kessler, Levine, 2000; Cowart-Steckler, & Pollack, 1998). Participants indicated if they had ever engaged in a specific activity at any point in their lifetime. Total scores were calculated as the number of behaviors endorsed by each participant.

A second measure was used to assess sexual risk behavior, or sexual behaviors that have been associated with an increased likelihood of transmitting STIs. Risk-taking was assessed using items from the Sexual Risk Survey (SRS), a measure of sexual risk

taking created specifically for use with college populations (Turchik & Garske, 2009). The SRS is a 23-item survey focused on sexual behavior participants have engaged in over the last 6 months. The survey consists of items such as, “How many partners have you had sex with?” and “How many times have you had sex with someone you just met?” and provides detailed definitions of subjective terms such as “sex.” The SRS has demonstrated good internal consistency ($\alpha = .88$) and 2-week test-retest reliability (.93). Higher scores on the SRS indicate greater risk taking behavioral disclosures.

Sexual victimization history. Sexual victimization history in adulthood was assessed using the short form of the Sexual Experiences Survey Short Form Victimization (SES-SFV; Koss et al., 2007). The SES-SFV is the most recent version of a well-established measure of sexual victimization. The SES-SFV presents participants with seven different unwanted or forced sexual experiences (e.g., "Someone had oral sex with me or made me have oral sex with them without my consent by:"). Each of these experiences is followed by a description of 5 types of coercion (i.e., verbal pressure, verbal manipulation, intoxication, threat of physical harm, physical force), allowing participants to indicate how they were coerced into that specific sexual experience. Participants provide an indication of the number of times each form of coercion has taken place over the past 12 months and since they were 14 years old; for the sake of this study, I only asked about victimization since the age of 14. The published version of the SES-SFV allows participants to indicate that they have experienced each act *0, 1, 2,* or *3 or more times*; for the sake of this study, I simply asked participants to write in a number indicating how many times they experienced each act. Based on responses to the SES-SFV, participants were classified as having experienced adult sexual victimization (i.e., a

response greater than 0 on any item) or as having not experienced adult sexual victimization (i.e., 0 on all items).

Participants' history of CSA victimization was assessed using a modified version of a CSA measure first developed by David Finkelhor (1981). The measure consists of questions about specific sexually abusive behaviors experienced prior to a participant's 14th birthday (e.g., "When you were 13 years old or younger, how many times did an older person [at least 5 years older than you] fondle you in a sexual way?"). Participants were asked to write in a number indicating how many times they experienced each act. Participants were also presented with the opportunity to indicate "no response." Similarly to the SES-SFV, participants were grouped into two categories---those who reported one or more act of CSA and those who reported no acts of CSA. A continuous measure of CSA was also created by summing all response on the measure.

Sexual perpetration history. Perpetration history was assessed using the Sexual Strategies Scale (SSS; Peterson et al., 2010). The scale is based on the Post-refusal Sexual Persistence Scale, an established measure of coercive behaviors (Stuckman-Johnson, Struckman-Johnson, & Anderson, 2003). This scale contains a list of coercive strategies used to obtain sex from an unwilling potential partner "who initially said no." Levels of coercion range from enticement (e.g., "Continuing to touch and kiss them in the hope that they will give in to sex"), to verbal coercion (e.g., "Telling them lies"), to use of intoxication (e.g., "Getting them drunk/high in order to convince them to have sex."), to use of physical force (e.g., "Tying them up"). Items are written to sound relatively innocuous in order to encourage honest endorsement. Participants are instructed to check all applicable strategies which they have used after a potential sexual partner initially said

"no." In a study with men, the SSS demonstrated convergent validity with another commonly-used measure of sexual perpetration; yet, men endorsed higher rates of perpetration on the SSS than the other measure, suggesting that the SSS may be capturing instances of coercion that were missed by the other measure (Strang, Peterson, Hill, & Heiman, in press). Participants who endorsed one or more strategies on the SSS were classified as having used coercion; participants who endorsed no strategies were classified as not having used coercion.

Sexual attitudes. The Sociosexual Orientation Inventory (SOI) is an 8-item measure that assesses attitudes towards casual, uncommitted sexual relationships (Fisher, Davis, Yarber, & Davis, 2010). The SOI consists of items such as "How often do you fantasize about having sex with someone other than your current partner?" and "with how many different partners have you had sex within the past year?" The measure also asks participants to rate the degree to which they agree with statements such as, "Sex without love is OK." on a scale ranging from 0 "I strongly disagree" to 8 "I strongly agree." Notably, the SOI includes items assessing both uncommitted sexual behavior and attitudes toward uncommitted sex. The SOI has demonstrated strong 2-month test-retest reliability (r = .94) and adequate internal consistency (Simpson & Gangestad, 1991). Higher scores are associated with stronger acceptance of and willingness to engage in uncommitted sex.

The Sexual Opinion Survey (SOS) was used to provide further measurement of sexual attitudes (Fisher, Byrne, White & Kelley, 1988). The SOS is designed to measure erotophobia-erotophilia, a personality dimension characterized by an affective reaction to sexual stimuli ranging from negative to positive (Fisher et al., 1988). The SOS consists of

21 items, each of which pairs a sexual stimulus with an affective response (e.g., "Masturbation can be an exciting experience," "I would not enjoy seeing an erotic movie."). Participants rate their degree of agreement with each item on a scale ranging from 1 (*strongly agree*) to 7 (*strongly disagree*). The scale is well studied, and has demonstrated high internal consistency (α = .76 - .89) and adequate test-retest reliability over several weeks (*r* = .61; Fisher et al., 1998). Higher scores on the SOS are indicative of greater erotophilia or more sex-positive attitudes.

Social desirability. Social desirability was assessed using the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1991). This is a well-established scale consisting of 40 items, which measure two components of desirable responding—self-deceptive positivity (i.e., positively-biased but relatively honest responses; e.g., "I never regret my decisions") and impression management (i.e., deliberately positive self-presentation aimed at favorable evaluation by others; e.g., "I never cover up my mistakes"). Participants rate items on a scale from 1 (*Not true*) to 7 (*Very true*). Participants receive one point for each item endorsed with a 6 or a 7. The scale has demonstrated adequate reliability (α = .68-.86) and test-retest correlations over a 5 week period (*r* =.65-.69; Paulhus, 1991). Higher scores on the BIDR indicate more socially desirable responding.

Participant Perceptions. Twelve additional items were included to assess participants' perceptions of anonymity, confidentiality, and the accuracy of their own responses. There were four items included for each of these factors (e.g., I feel this study is completely confidential."), and each item asked participants to rate the degree to which they agreed with the statement on a seven point likert-type scale ranging from 1 (*strongly*

disagree) to 7 (*strongly agree*). Participants were only presented with these items at the end of the time 2 survey administration in order to avoid introducing subject-expectancy effects. These three four-item scales demonstrated adequate internal consistency reliability in the current sample (for anonymity, $\alpha = .96$; for confidentiality, $\alpha = .96$; and for accuracy, $\alpha = .99$).

Context of Questionnaire Completion. Several questions were also included to assess the context in which the participant completed the questionnaire (e.g., "Was anyone else in the room when you completed this questionnaire?"). These questions were presented to participants across all conditions during both the time 1 and the time 2 survey administrations.

Procedure

Participants indicated their interest in the study by logging in to a campus subject pool web portal and completing a pretest questionnaire in which they provided contact information and indicated their gender. The study was advertised as "a research study about sexual experiences and attitudes." Interested participants were randomly assigned to one of the eight conditions and were contacted via email with instructions for participation in the study. Instructions provided in the emails were dependent on the condition to which participants were assigned. Participants in lab-based conditions were required to set up appointment times through online scheduling software in order to avoid confounding the experimenter contact variable. Participants who were completing surveys outside of the lab were simply provided with instructions outlining how they could complete the survey.

The instructions emailed to participants also reflected experimental contact and the first mode by which participants were to complete the survey. Participants in the high experimental contact condition were informed of the interactions which would take place with the experimenter, and these interactions were presented as a means to "ensure that the participant feels comfortable completing the survey." Similarly, participants in the low contact condition were informed of the lack of interaction with the experimenter, which was also presented as a means to "ensure that the participant feels comfortable completing the survey."

Participants in high experimenter contact conditions were matched with an experimenter of the same gender in order to minimize the potential for gender-related experimenter effects. All experimenters were White; thus matching race of the experimenter and participants was not possible. The experimenters for this study were trained undergraduate research assistants who used a detailed research protocol to ensure standardized administration across all trials.

Participants in high contact, in-lab conditions were greeted face-to-face by an experimenter and were verbally guided through informed consent and study procedures. Participants in the high contact, out of-lab conditions were greeted by an experimenter via phone and verbally guided through informed consent and procedures. Participants in all low contact conditions were provided with written instructions and did not interact verbally with experimenters at any point during the study. In-lab participants completed surveys in the lab space. Out-of-lab participants completed surveys in a setting of their own choosing. In the out-of lab, pencil-and-paper condition, participants were asked to return surveys to a predetermined drop-off point on campus.

In the initial email by which participants were informed of the instructions for participation, they were also provided information for the procedure related to the second portion of the study. During the second portion of the study, participants completed the measures in the alternate mode (e.g., if they completed pencil-and-paper in the first portion, they completed online in the second portion); the setting and experimenter contact conditions remained the same for the first and second portion of the study. Participants who were assigned to the in lab conditions were asked to schedule their follow-up appointment at the same time they scheduled their initial visit. All participants were contacted via their provided email address and were reminded about the second portion of participation. Participants who failed to return or respond within two weeks of the recommended completion date were provided with a reminder email requesting their completion of the required measures. Perceptions of anonymity, confidentiality, and accuracy were assessed only at the end of the second portion of the study in order to minimize expectancy effects. See table 4 for a description of each of the experimental conditions.

Table 4
Summary of Procedures for Each Experimental Condition

Condition	Description of procedures
High contact, in lab, computer	Participants scheduled appointments online, presented to lab space, were greeted by experimenters, were guided through informed consent by experimenters, and completed surveys via computer.
High contact, in lab, pencil-paper	Same as above, except participants completed surveys via pencil-and paper format.
Low contact, in lab, computer	Participants scheduled appointments online, presented to lab space with no experimenter present, read through informed consent alone in lab, and completed surveys via computer.
Low contact, in lab, pencil-paper	Same as above, except participants completed surveys via pencil-and paper format.
High contact, out of lab, computer	Participants scheduled appointments online, contacted experimenters via phone, were guided through informed consent by experimenter, were provided a web link and completed the survey via computer. Participants also followed up with experimenters via phone after completing the survey.
High contact, out of lab, pencil-paper	Participants picked up a pencil-and-paper survey from campus, scheduled appointments online, contacted experimenters via phone, were guided through informed consent by experimenter, and completed the pencil-paper survey. Participants also followed up with experimenters via phone after completing the survey. Participants then returned the surveys to campus.
Low contact, out of lab, computer	Participants were emailed a web link to the computer based survey, they read through informed consent, and completed the computer based survey.
Low contact, out of lab, pencil-paper	Participants picked up a pencil-and-paper survey from campus, read through the informed consent, and completed the survey at a location of their own choosing. Participants then returned the surveys to campus.

Protections Against Risk

Following the completion of questionnaires, participants were provided with a list of area resources that provide mental health services. This was done in order to acknowledge and account for any potentially negative reactions that participants might have to thinking about and responding to questions of a sensitive nature. Participants were provided with these resources across all conditions after completion of each of the two data collection sessions.

Due to the sensitive nature of the data being collected and the lack of anonymity within the design of this study, precautions were taken to protect participants' personal information and to keep their information confidential In terms of information collected via pencil-and-paper, with the exception of a coded identifying number, participants' identifying information was not included on any questionnaire or record. This identifying number was linked to a list of contact information, and that list was held securely in a locked office. All identifying information was destroyed once data collection had been completed. The electronic data collected was maintained on a secure server. Further, all information collected in this study was used solely for research purposes. This report and any subsequent reporting of results from this study will be done in such a way that no individual participant can be identified.

It is important to note that my assessment of sexual perpetration may have led some participants to disclose illegal activities. Given that this study is not anonymous, I have been careful to approach this assessment with caution. As such, I intentionally avoided questions relating to criminal behaviors which would have required me to file a

police report (e.g., elder abuse, child abuse, future intent to harm). Language was included in the informed consent statement informing participants of the unlikely possibility that a court could subpoena my data prior to the de-identification process. However, now that data have been deidentified, this risk has been eliminated.

Results

Preliminary analyses

Although I did not hypothesize systematic differences in response rate or item completion across modes of inquiry, such differences certainly seemed like a possibility. Thus, I examined differences in response rates and rates of incomplete data across all conditions in order to determine if there were differences in participation bias associated with experimenter contact, setting, or inquiry mode. Logistic regression analyses were used to compare participant response rates as a function of high or low experimenter contact, in or out of lab setting, and pencil-and-paper or computer based modes of inquiry. I first looked at the degree to which potential participants who indicated interest in the study actually followed through on scheduling appointments, and I then looked at the degree to which those who scheduled appointments followed through on completing them.

First, a logistic regression was run with potential participants' decision to schedule an initial appointment (yes/no) serving as the dependent variable. Setting, mode, and experimenter contact, along with their interactions, were included as the independent variables. The logistic regression revealed a main effect for mode ($\beta = 1.92$, odds ratio = 6.80, $p < 0.05$) with participants being more likely to schedule an appointment in

computer-based than paper-and-pencil conditions. However, there was also a significant interaction between mode and setting ($\beta = -1.04$, odds ratio = 0.34, $p < 0.01$). An examination of the interaction suggested that potential participants in the computer and paper modes scheduled appointments at the same rate in lab based conditions. Whereas in out of lab conditions, potential participants were significantly more likely to schedule appointments in computer based conditions (70.5%) than in paper-and-pencil conditions (51.5%). For those in the out of lab conditions, online completion meant that participants could complete the study right away rather than taking the additional step of coming onto campus to pick up a pencil-and-paper survey.

A second logistic regression was run for prospective participants who scheduled appointments, with attendance (yes/no) serving as the dependent variable, and setting, mode, experimenter contact, along with the interaction terms being entered as independent variables. There were no main effects for scheduling as a function of any of the experimental variables. However, there was a significant interaction between mode and setting ($\beta = -1.55$, odds ratio = 0.21, $p = 0.01$). Again, this interaction suggested that while scheduled participants in the computer and paper modes attended their appointments at approximately the same rate in lab, participants in out of lab conditions were significantly more likely to attend computer based appointments (84%) than paper-and-pencil based appointments (67%). This again suggests that participants who scheduled appointments in out of lab conditions were more willing to click through an email link than they were to come to campus to pick up a physical copy of the survey in order to participate. For a complete overview of participation rates across conditions, see table 5.

Table 5
Percentage of Participants Interested/Scheduled/Completed Based on Condition

	Interested	Scheduled	%	Completed	%
High, in, comp	75	57	76	40	70
Low, in, comp	86	57	66	36	63
High, out, paper	100	49	49	27	55
High, out, comp	98	65	66	53	82
Low, in, paper	86	63	73	50	79
Low, out, paper	96	52	54	41	79
High, in, paper	70	53	76	40	75
Low, out, computer	65	50	77	44	88
Total	676	446	66	331	74
Range	65-100	49-65	49-77	27-53	55-88

Due to the small number of men in my final sample, I also looked at the rates of prospective participants who expressed interest, scheduled, and completed appointments as a function of gender. The percentage of prospective participants who scheduled an initial survey after indicating interest did not differ based on gender, χ^2 (1, N = 676) = 1.07, p = 0.29. Similarly, the percentage of prospective participants who completed their initial appointment after scheduling did not differ based on gender, χ^2 (1, N =446) = 0.107, p = 0.73. Thus, the lower number of men in my final sample as compared to women seemed to reflect a lower number of men in the psychology subject pool and/or men's lower interest in this particular study topic rather than gender differences as a function of experimental variables.

Missing data were visually inspected for patterns related to condition and specific items. No obvious patterns emerged. Further, missing data were minimal across all of the outcome variables within the study, with less than 3% of participants missing datum on any individual item. Missing data were handled on a measure by measure basis. When

missing items could be considered non-endorsement of a particular behavior, I treated missing data as non-endorsement. Otherwise, missing items were assumed to be random, and I used the multiple imputation function included in the missing values add-on for SPSS version 20 (IBM Corp., 2012). When multiple imputation was used, five imputations were completed for missing values, and analyses were run separately on each imputed data set. The results from each imputation were averaged and this average is reported in my results. This approach to missing data is recommended because it allows for acknowledgement and incorporation of uncertainty within replaced values (e.g., Acock, 2005; Schafer, 1999).

For the Sexual Behavior Measure (SBM), less than 1% of responses were missing for each individual item. Missing items were treated as non-endorsement of the behaviors and were coded as "no" responses. Additionally, one case was removed from SBM analyses due to the participant's failure to complete more than 50% of the measure.

For the Sexual Risk Survey (SRS), less than 2% of responses were missing for each individual item. "How many times have you given fellatio without a condom?" was the most commonly omitted item, skipped by 1.8% of participants. Missing items for the SRS were treated as random and were accounted for using multiple imputation.

The Sociosexual Orientation Inventory (SOI) had less than 1% of omitted responses for each individual item. Missing items for the SOI were treated as random and were accounted for using multiple imputation. One case was removed from the SOI analyses due to the participant's failure to complete more than 50% of the measure.

Less than 2% of responses were missing for individual items on the Sexual Opinion Survey (SOS). Missing responses were treated as random and were accounted for with multiple imputation. The most commonly omitted item on the SOI was "When I think about seeing pictures showing someone of the same sex as myself masturbating it nauseates me (1.5% missing)."

On the Sexual Experiences Survey (SES-SFV), less than 2% of responses were missing for each individual item. Missing items were treated as non-endorsed experiences and were replaced with "no" responses. One case was also removed from SES-SFV analyses due to the participant failing to complete more than 50% of the measure.

There were fewer than 2% of omitted responses to individual items on the Child Sexual Abuse Measure (CSAM) as well. Missing items on the CSAM were treated as non-endorsed experiences and were replaced with "no" responses. Three cases were also removed from the CSAM data set, due to these participants failure to complete more than half of the measure.

For the Sexual Strategies Scale (SSS), all "non-checked" items were treated as non-endorsed strategies, and were replaced with "no" responses. On this measure, participants are instructed to only check relevant items, so non-endorsement of other items is expected.

On the Balanced Inventory of Desirable Responding (BIDR), there were less than 3% of responses omitted for each individual item. The most commonly omitted item on the BIDR was "I always declare everything at customs," with 2.7% of participants

skipping this item. Non-responses on the BIDR were treated as random, and were replaced using multiple imputation.

Main analyses

Modification of proposed analyses became necessary due to difficulties with participant recruitment and the distribution of participants' responses. The low response rate for time 2 data collection (34%) prevented me from conducting within-subject analyses of mode related effects as part of the hypothesis testing. Instead only time 1 data was used in the main analyses, and mode during time 1 was treated as a between-subject variable.

Further, as my recruitment of male participants fell well below my targeted sample size, I did not have adequate power to examine gender main effects or interactions as was initially proposed. Instead, male and female participants were assed separately and analyses using male participants were considered exploratory given the limited sample size. Because of the lack of power to detect differences among the male sample, I attended to effect sizes for all analyses involving male participants; still, no clear conclusions can be drawn given the low numbers of men in many cells of the analyses. See table 6 for a complete overview of the distribution of male participants across conditions.

Table 6
Distribution of Male Participants Across Conditions

In vs. out of lab	Paper vs. computer	Race	Experimenter contact	N
Out of lab	Paper	White	High	3
			Low	9
		Non-White	High	2
			Low	2
	Computer	White	High	9
			Low	7
		Non-White	High	3
			Low	3
In lab	Paper	White	High	4
			Low	6
		Non-White	High	6
			Low	0
	Computer	White	High	4
			Low	4
		Non-White	High	6
			Low	4
Totals				
Out: 38	Paper: 32	White: 46	High:37	
In:34	Computer: 40	Non-White: 26	Low:35	72

Finally, examination of the distribution of behavioral count variables (SRS, SOI, SSS, SES, CSAM), revealed severe positive skewness resulting from high rates of zero count responses, violating assumptions of normality. As such, a negative binomial regression model was used for the SRS, as it is more able to account for this type of distributions (e.g., Hutchinson & Holtman, 2005) and the SOI was transformed as described below. Additionally, because I was primarily interested in whether participants reported sexual aggression perpetration or sexual victimization rather than the frequency with which they reported these experiences, a logistic regression was used for dichotomous versions of the SSS, the SES, and CSAM. A summary of the descriptive data for the dependent variables can be found in table 7.

Table 7
Descriptive Data for Dependent Variables Examined in this Study

Cont. DV's	Mean scores (SD)			Min	Max	Skewness	Kurtosis
	Women	Men	Total				
SBM	13.35 (6.11)	12.74 (4.89)	13.21 (5.87)	0	28	0.214	-0.193
SRS	53.23 (72.96)	41.88 (56.8)	50.64 (69.61)	0	404	2.50	7.17
CSAM	3.43 (7.59)	1.03 (4.42)	2.92 (7.09)	0	44	3.30	11.50
SOI	33.23 (23.59)	49.82 (30.76)	36.83 (26.19)	2	155	1.75	4.268
SOS	74.76 (25.32)	75.66 (21.52)	74.93 (24.52)	9	126	-0.397	-0.288
BIDR SD	6.38 (4.10)	6.39 (3.64)	6.38 (4.10)	0	22	0.618	-0.147
BIDR IM	5.66 (3.26)	5.52 (3.30)	5.61 (3.27)	0	20	0.627	0.790
Dich. Dv's	Percentage of endorsement						
	Women	Men	Total				
SSS	38.7	52.8	41.7				
CSAM	34	15.5	30				
SES-SFV	57.1	29.6	51.2				

Hypothesis 1. Participants in low experimenter contact, out of lab completion, and internet based inquiry mode conditions were expected to report significantly higher rates of sexual behaviors than participants in high experimenter contact, in lab completion, and pencil-and-paper inquiry mode conditions. Further, I predicted that these effects would be moderated by Race with larger differences between conditions demonstrated by non-White participants than by White participants.

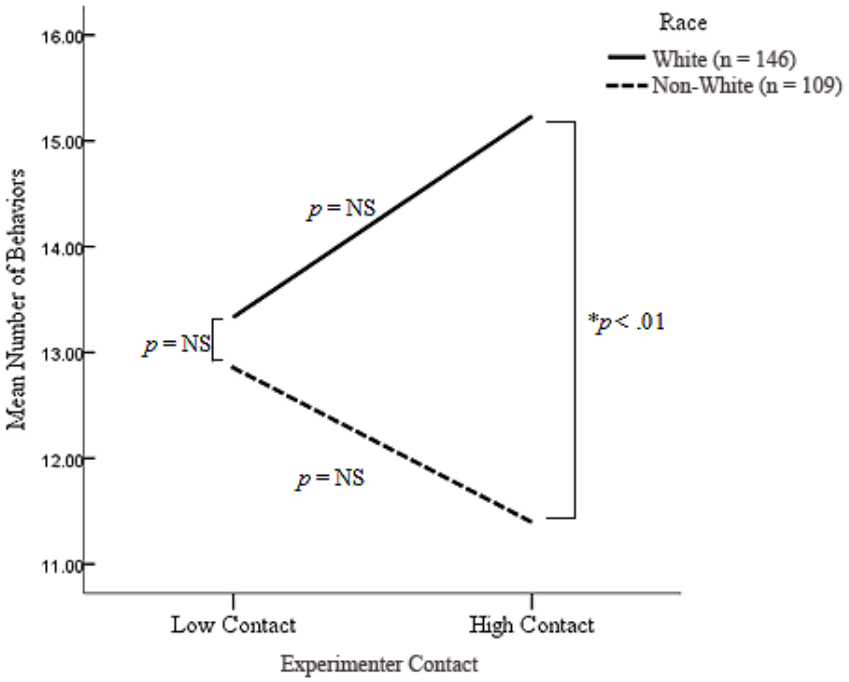
The measure of sexual behavior used to test this hypothesis was the SBM. Overall, women across conditions endorsed an average of 13.35 behaviors out of a possible 28 (*SD* = 6.11) whereas men endorsed an average of 12.74 behaviors (*SD* =

4.89) out of a possible 29. The most frequently endorsed behavior for men (91.7%) and women (96.9%) was "kissing someone of the other sex on the mouth." The least frequently endorsed item for both men (6.9%) and women (3.1%) was "Mouth contact with the anus/butt of someone of the other sex ('rimming')."

For the SBM, Hypothesis 1 was tested using a 2 (Race; White versus non-White) X 2 (Experimenter Contact; high versus low) X 2 (Setting; in lab versus out of lab) X 2 (Mode; computer versus pencil-and-paper) univariate ANOVA with the number of sexual acts endorsed as the dependent variable. This analysis was repeated separately for male and female participants.

For women, the only significant main effect was for Race, with White participants reporting more sexual behaviors (M = 14.28; SE = 0.50) than non-White participants (M = 12.13; SE = 0.58), F (1, 239) = 7.92, p = .005, η_p^2 = 0.03. Contrary to my hypotheses, there were no significant main effects of experimenter contact and setting on behaviors reported; mode had an effect which approached significance, F (1, 239) = 3.60, p = .06, η_p^2 = 0.02, and suggested that participants may have reported more behaviors on paper based surveys (M = 13.93; SE = 0.56) than those presented via computer (M = 12.48; SE = 0.53).

Figure 1
Total Number of Behaviors Reported by Women on SBM as a Function of Race and Experimenter Contact

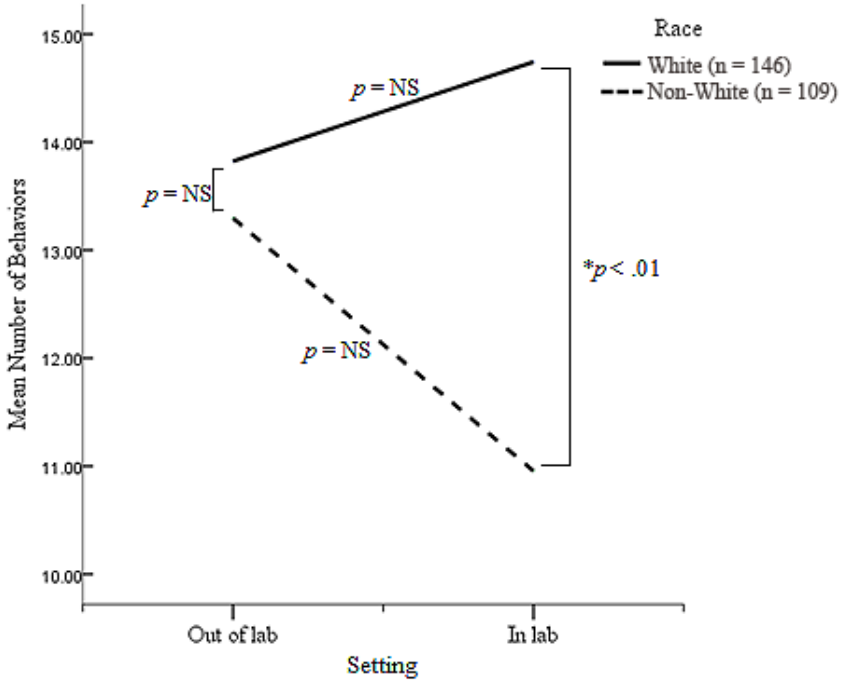


There was a significant interaction between the effects of experimenter contact and race on the number of sexual behaviors endorsed, F (1, 239) = 4.80, p = .03, η_p^2 = 0.02 (see figure 1).

However, contrary to the hypothesized direction of the interaction, for Non-White participants there was no significant difference in the number of behaviors reported in high contact (M = 11.40; SE = 0.85) versus the low contact (M = 12.86; SE = 0.79) conditions, p = .12. Further there were no significant differences observed between White participants in high contact (M = 15.23: SE = 0.73) versus low contact (M = 13.33; SE = 0.68) conditions, p = .07. The difference emerged between White (M = 15.23; SE = 0.73) and non-White participants (M = 11.40; SE = 0.85) in the high contact condition.

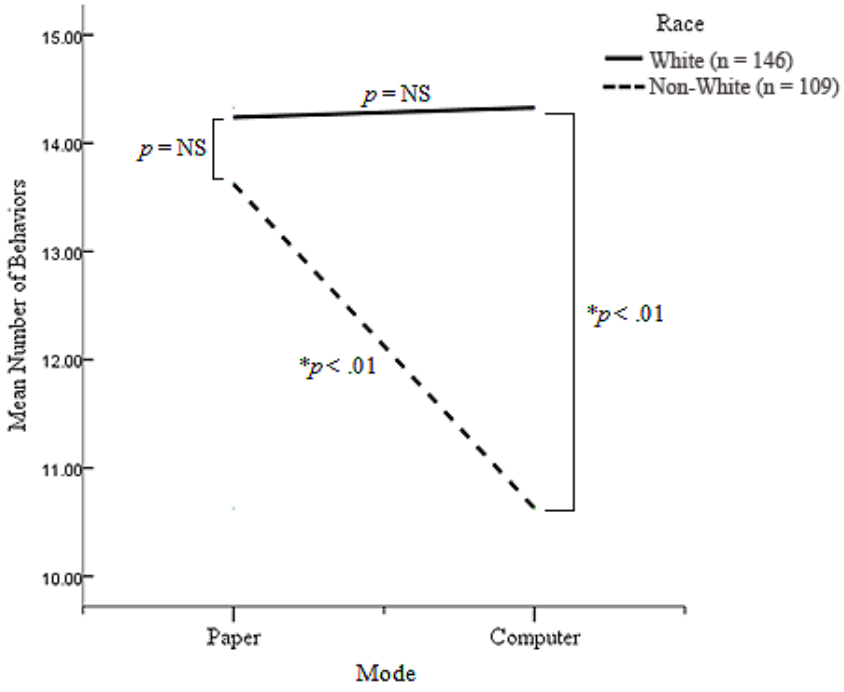
There was also a significant interaction between the effects of setting and race on the number of sexual behaviors endorsed, F (1, 239) = 4.52, p = .04, η_p^2 = 0.02 (see figure 2). However, again contrary to the hypothesized direction of the interaction, for Non-White participants there was no significant difference in the number of behaviors reported in lab based conditions (M = 10.96; SE = 0.79) versus out of lab conditions (M = 13.30; SE = 0.85), p = .11. Further there were no significant differences observed between White participants in lab based conditions (M = 14.74; SE = 0.72) versus out of lab conditions (M = 13.82; SE = 0.70), p = .24. Again, the difference emerged between White (M = 14.74 SE = 0.72) and non-White participants (M = 10.96; SE = 0.79) in the high contact condition.

Figure 2
Total Number of Behaviors Reported by Women on SBM as a Function of Race and Setting



Finally, there was a significant interaction between the effects of mode and race on the number of sexual behaviors endorsed, F (1, 239) = 4.05, p = .045, η_p^2 = 0.02 (see figure 3). For Non-White participants there was a significant difference in the number of behaviors reported on pencil-and-paper surveys (M = 13.62; SE = 0.84) and computer based surveys (M = 10.63; SE = 0.81) conditions, p < .01. However, there were no significant differences observed between White participants on pencil-and-paper surveys (M = 14.24 SE = 0.73) and computer based surveys (M = 14.33; SE = 0.68) conditions, p = .99.

Figure 3
Total Number of Behaviors Reported by Women on SBM as a Function of Race and Mode



The same analyses were repeated on an exploratory basis for the male participants in the sample. No significant main effects were observed for race, experimenter contact, setting or mode. Further, no significant two-way interactions were observed for men. An

examination of effect sizes revealed no strong effects (all η_p^2s < .02), suggesting that the non-significant results may not simply reflect insufficient power. Thus, I did not find support for my hypotheses with men (see table 8 for a summary of men's and women's scores).

Table 8
Estimated Marginal Mean Scores on SBM as a Function of Gender

	Men	Women
	Estimated Marginal Mean ± SEM	Estimated Marginal Mean ± SEM
Experimenter Contact		
Low	11.83 ± 0.9	13.10 ± 0.5
High	13.09 ± 0.9	13.31 ± 0.6
Setting		
Out of lab	12.06 ± 0.9	13.56 ± 0.6
In lab	13.01 ± 0.8	12.85 ± 0.5
Mode		
Paper	13.16 ± 1.0	13.93 ± 0.6
Computer	11.93 ± 0.8	12.48 ± 0.5
Race		
White	13.00 ± 0.8	14.28 ± 0.5
Non-White	11.93 ± 1.0	12.13 ± 0.6
Mode*Race		
Paper/White	12.61 ± 1.1	14.24 ± 0.7
Paper/Non-White	13.89 ± 1.7	13.62 ± 0.8
Computer/White	13.39 ± 1.0	14.33 ± 0.7
Computer/Non-White	10.46 ± 1.2	10.63 ± 0.8
Setting*Race		
Out of lab/White	12.57 ± 1.0	13.82 ± 0.7
Out of lab/Non-White	11.54 ± 1.5	13.30 ± 0.9
In lab/White	13.44 ± 1.1	14.74 ± 0.7
In lab/Non-White	12.44 ± 1.2	10.96 ± 0.8
Experimenter contact*Race		
Low/White	13.08 ± 1.0	13.33 ± 0.7
Low/Non-White	10.17 ± 1.7	12.86 ± 0.8
High/White	12.92 ± 1.2	15.23 ± 0.7
High/Non-White	13.25 ± 1.3	11.40 ± 0.9

Hypothesis 1 was also tested using the Sexual Risk Survey (SRS). Due to the aforementioned concerns related to positive skewness, a negative binomial regression model was used, with race, experimenter contact, setting, and mode as predictor variables and the total score on the SRS as the outcome. Table 9 shows the results of the negative binomial regression model used to determine the significant predictors of the number of sexual experiences reported by women on the SRS. Contrary to my prediction, but consistent with SBM findings, female participants completing pencil-and-paper based surveys reported significantly more sexual risk behaviors than participants completing computer based surveys [*B* (95% confidence intervals) = 0.725 (0.317, 1.133); $p<0.001$], odds ratio = 2.07. Additionally, the model revealed a significant interaction between Mode and Race for the number of sexual experiences being reported [*B* (95% confidence intervals) = -0.814 (-1.341, -0.287); $p<0.01$], odds ratio = 0.44. The overall model had a satisfactory goodness-of-fit (1.98) as defined by deviance/df statistics. Follow up analyses revealed that for Non-White participants there was a significant difference in the number of risk behaviors reported on pencil-and-paper survey conditions (*M* = 69.11; *SE* = 9.95) and computer based survey conditions (*M* = 33.47; *SE* = 4.72), [*B* (95% confidence intervals) = 0.83 (0.19, 0.45); $p<0.001$], odds ratio = 2.30. However, there were no significant differences observed between White participants on pencil-and-paper survey conditions (*M* = 48.59; *SE* = 6.10) and computer based survey conditions (*M* = 53.1; *SE* = 6.09), *B* (95% confidence intervals) = -0.08 (0.17, 0.41); $p = 0.67$, odds ratio = 0.93.

Table 9
Negative Binomial Regression Model of Sexual Risk Survey (SRS) for Women

Variable	*B* (95% CI)	OR	Est. Marginal Mean ± SEM
Experimenter Contact	0.245 (-0.164, 0.65)		
Low		1.28	53.01 ± 4.8
High		1	46.09 ± 4.5
Setting	-0.206 (-0.598, 0.185)		
Out of lab		0.814	49.64 ± 4.7
In lab		1	49.22 ± 4.5
Mode	0.725 (0.317, 1.133)**		
Paper		2.07	57.94 ± 5.5
Computer		1	42.17 ± 3.8
Race	0.341 (-0.170, 0.852)		
White		1.41	50.80 ± 4.3
Non-White		1	48.10 ± 4.7
Mode*Race	-0.814 (-1.341, -0.287)*		
Paper*White		0.44	48.59 ± 6.1
Paper*Non-White		1	69.11 ± 9.9
Computer*White		1	53.11 ± 6.1
Computer*Non-White		1	33.48 ± 4.7
Setting*Race	0.430 (-0.093, 0.953)		
Out of lab*White		1.54	56.81 ± 6.8
Out of lab*Non-White		1	43.38 ± 6.4
In lab*White		1	45.43 ± 5.7
In lab*Non-White		1	53.33 ± 7.0
Experimenter contact*Race	-0.210 (-0.747, 0.328)		
Low*White		0.81	51.07 ± 6.1
Low*Non-White		1	54.36 ± 7.4
High*White		1	49.92 ± 6.3
High*Non-White		1	42.56 ± 6.3

**$p < .001$, *$p < .01$

The same analysis was repeated for men, and the results are summarized in table 10. Male participants completing pencil-and-paper based surveys reported significantly fewer sexual risk behaviors (M = 18.02; SE = 3.85) than male participants completing computer based surveys (M = 46.97; SE = 7.95), B (95% confidence intervals) = -1.39 (0.44, -2.26); p<0.01), odds ratio = 0.25. Further, male participants in low contact

conditions reported significantly fewer sexual risk behaviors (*M* = 25.51; *SE* = 5.4) than male participants completing surveys in high contact conditions (*M* = 33.18; *SE* = 5.7), *B* (95% confidence intervals) = -1.28 (-2.15, -0.41); $p<0.01$, odds ratio = 0.28. Thus, there was mixed support for my hypothesized main effects. Additionally, the model revealed a significant interaction between Experimenter Contact and Race for the number of sexual risk behaviors being reported by men, *B* (95% confidence intervals) = 2.04 (0.98, 3.10); $p<0.001$, odds ratio = 7.67. The overall model had a satisfactory goodness-of-fit (1.83) as defined by deviance/df statistics. Follow up analyses revealed that for White participants there was a significant difference in the number of risk behaviors reported in high contact conditions (*M* = 25.8; SD = 5.88) and low contact conditions (*M* = 50.84; *SE* = 10.06), *B* (95% confidence intervals) = 0.67 (0.09, 1.27); $p = 0.25$, odds ratio = 1.97. Additionally, there was a significant difference in the opposite direction observed between non-White participants on high (*M* = 57.8; *SE* = 14.15) and low contact (*M* = 21.56; *SE* = 7.35) conditions, *B* (95% confidence intervals) = -0.99 (-1.81, -0.16); $p = 0.02$, odds ratio = 0.37.

Table 10
Negative Binomial Regression Model of Sexual Risk Survey (SRS) for Men

Variable	*B* (95% CI)	OR	Est. Marginal Mean ± SEM
Experimenter Contact	-1.28 (-2.15, -0.41)*		
Low		0.28	25.51 ± 5.4
High		1	33.18 ± 5.7
Setting	0.49 (-0.78, 0.88)		
Out of lab		1.05	26.90 ± 5.2
In lab		1	31.47 ± 5.9
Mode	-1.39 (-2.26, -0.52)*		
Paper		0.25	18.02 ± 3.9
Computer		1	46.97 ± 7.9
Race	-0.87 (-1.77, 0.03)		
White		0.42	35.06 ± 5.5
Non-White		1	24.14 ± 5.5
Mode*Race	0.86 (-0.20, 1.92)		
Paper*White		2.36	26.92 ± 6.2
Paper*Non-White		1	12.06 ± 4.3
Computer*White		1	45.67 ± 9.5
Computer*Non-White		1	48.29 ± 12.9
Setting*Race	-0.41 (-1.44, 0.62)		
Out of lab*White		0.66	29.25 ± 5.7
Out of lab*Non-White		1	24.73 ± 8.2
In lab*White		1	42.03 ± 10.2
In lab*Non-White		1	23.56 ± 6.73
Experimenter contact*Race	2.04 (0.98, 3.310)**		
Low*White		7.67	51.14 ± 10.3
Low*Non-White		1	12.71 ± 4.7
High*White		1	24.04 ± 5.7
High*Non-White		1	45.81 ± 11.5

**$p < .001$, *$p < .01$

I found partial support for my hypotheses related to the impact of methodological factors on self-reported sexual behavior. For the SRS, both women and men demonstrated a main effect for mode on the number of sexual experiences being reported. As predicted, men reported significantly more experiences on computer based surveys, whereas, contrary to predictions, women reported significantly more experiences on pencil-and-paper surveys. Although the main effect of mode runs counter to my hypothesis for women, it is consistent with the trend towards significance ($p = .059$)

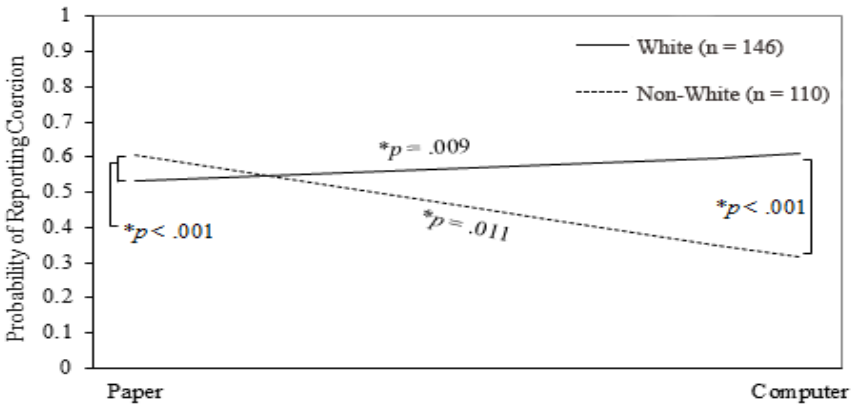
observed on the SBM which also suggested that women reported more behaviors in pencil-and-paper conditions.

I also found a race by mode interaction for women on both behavior measures. Non-White female participants were significantly less likely to report sexually risky behaviors on computer based surveys than they were on pencil-and-paper surveys; this difference was not found for White participants. The SBM also revealed significant race dependent interactions for setting and experimenter contact, with non-White participants reporting significantly fewer behaviors than White participants in high contact conditions but not in low contact conditions and in lab based conditions but not out of lab conditions. I also found a race by experimenter contact interaction for men on the SRS, which suggested that male White participants were significantly more likely to report sexually risky behaviors in low contact conditions as compared to high contact conditions, and non-White participants were significantly more likely to report risky behaviors in high contact conditions as compared to low contact conditions. Overall, for both measures of sexual behavior tested in Hypothesis 1, there was support for mode as a factor which independently influences reported sexual behavior and race as a moderator for the impact of experimental variables.

Hypothesis 2. Low experimenter contact, out of lab completion, and computer based inquiry mode conditions were predicted to be associated with greater likelihood of reported sexual victimization (CSA, adult sexual assault) and perpetration (adult sexual assault) than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode conditions. Further, I predicted that these effects would be moderated by race, with observed effects being greater for non-White than White participants.

The Sexual Strategies Scale (SSS) was used to examine the impact of methodological variables on disclosure of sexual perpetration. Participants were asked to indicate which of 23 coercive strategies they had used to convince someone to have sex. The majority of female participants (61.3%) denied having ever used any of the strategies, with the remaining participants acknowledging having used between 1-14 strategies (M = 2.94, SD = 2.41). A logistic regression was conducted with a dichotomous version of the SSS, which separated participants based on whether or not they had reported any history of coercive behavior. Race, Contact, Setting, and Mode were entered in step 1, and interactions between race and each of the experimental variables were entered in step 2. The regression revealed that female participants were more than three times as likely to acknowledge coercive behaviors on pencil-and-paper based surveys (44%) than on computer based surveys (35%; $\beta = 1.20$, odds ratio = 3.32, $p < 0.01$). The regression also indicated that White participants were more likely to report a history of coercive behaviors (39%) than non-White participants (38%; $\beta = 1.18$, odds ratio = 3.27, $p < 0.05$). However, these main effects were qualified by a race by mode interaction (β = -1.52, odds ratio = 0.22, $p < 0.01$). Follow-up analysis shows that White participants in computer based conditions (41%) were more likely to report coercive behaviors than White participants in pencil-and-paper conditions (37%), χ^2 (1, N = 146) = 6.92, p = .009. Non-White participants were more likely to report coercive behaviors in pencil-and-paper conditions (52%) than they were in computer based conditions (26%), χ^2 (1, N = 113) = 6.451, p = .011 (see figure 4).

Figure 4
SSS logistic regression, women, mode by race interaction effect



The same logistic regression analysis was repeated with male participants. Over half of men (52.8%) acknowledged having engaged in some form of coercive behavior in an attempt to convince someone to have sex with them. Of the men who reported having used some form of coercive strategy, 1-12 strategies were reported (M = 3.97, SD = 2.41). For men, none of the predictor variables in the regression model were significant. Given the small sample of men, I examined the odds ratios as measures of effect size. There was a moderate effect for experimenter contact by race interaction (odds ratio = 4.19); however, the huge confidence interval (.46 – 38.27) suggested that the results were unstable due to the small sample size.

Overall, I found partial support among women for my hypotheses related to the impact of methodological variables on self-reported sexual coercion. Main effects were observed for mode and race, suggesting that White participants and participants completing pencil-and-paper surveys were significantly more likely to report sexually coercive behavior than non-White participants and participants in the internet conditions, respectively. Though the direction of the mode related effect is not what I had

hypothesized, it is consistent with the effects observed related to the sexual behavior measures. I also observed a race related moderation of mode in that non-White participants were more likely to report coercive behaviors on pencil-and-paper measures, and White participants were more likely on computer based surveys. No significant or interpretable effects were observed for male participants.

The Child Sexual Abuse Measure (CSAM) was used to examine the impact of methodological variables on disclosure of childhood sexual victimization. The CSAM was used to create both a continuous and dichotomous measure of CSA. The majority of female participants reported no history of CSA (66%), resulting in a clustering of responses at zero. However, scores on the CSAM ranged from 0-44, with higher scores indicating acknowledgement of more instances of CSA, suggesting that some women experienced substantial sexual abuse in childhood (M = 3.43, SD = 3.43). Due to the distribution of the continuous variable, a negative binomial regression model was used for my analyses. The overall model had a satisfactory goodness-of-fit (2.6) as defined by deviance/df statistics.

Table 11 shows the results of the negative binomial regression model used to determine the significant predictors of the number of CSA experiences reported by women on the CSAM. Female participants in low contact conditions, [B (95% confidence intervals) = 0.506 (0.020, 0.992); $P < 0.05$, odds ratio = 1.66], out of lab conditions [B (95% confidence intervals) = 0.738 (0.274, 1.201); $P < 0.01$, odds ratio = 2.09], and pencil-and-paper based conditions [B (95% confidence intervals) = 0.950 (0.490, 1.410); $P < 0.001$, odds ratio = 2.59], reported significantly more behaviors than their counterparts in high contact, in lab, or computer based conditions, respectively.

Additionally, there was a interaction between setting and race which significantly impacted female participant's reporting of CSA [*B* (95% confidence intervals) = -0.852 (-1.460, -0.244); $P < 0.01$, odds ratio = 0.43], suggesting that, in out of lab conditions, non-White participants report significantly more CSA experiences than White participants, with no significant differences observed for in lab conditions. No significant differences were observed between White and non-White participants in either setting.

Table 11
Negative Binomial Regression Model of Child Sexual Abuse Measure (*CSAM) for Women*

Variable	*B* (95% CI)	OR	Est. Marginal Mean ± SEM
Experimenter Contact	0.506 (-0.355, 0.640)*		
Low		1.66	3.74 ± 0.39
High		1	2.48 ± 0.28
Setting	0.738 (0.274, 1.021)**		
Out of lab		2.09	3.56 ± 0.38
In lab		1	2.61 ± 0.28
Mode	0.950 (0.490, 1.410)***		
Paper		2.59	4.77 ± 0.49
Computer		1	1.95 ± 0.22
Race	0.324 (-0.339, 0.989)		
White		1.38	2.69 ± 0.27
Non-White		1	3.45 ± 0.38
Mode*Race	-0.109 (-0.712, 0.494)		
Paper*White		0.90	4.09 ±0.56
Paper*Non-White		1	5.55 ± 0.85
Computer*White		1	1.76 ± 0.25
Computer*Non-White		1	2.15 ± 0.36
Setting*Race	-0.852 (-1.460, -0.244)**		
Out of lab*White		0.43	2.54 ± 0.35
Out of lab*Non-White		1	4.99 ± 0.80
In lab*White		1	2.84 ± 0.41
In lab*Non-White		1	2.39 ± 0.39
Experimenter contact*Race	-0.193 (-0.819, 0.433)		
Low*White		0.83	3.14 ± 0.42
Low*Non-White		1	4.45 ± 0.70
High*White		1	2.30 ± 0.34
High*Non-White		1	2.68 ± 0.47

*$p < .05$, **$p < .01$, ***$p < .001$

Analyses were repeated for male participants. The majority of men in my sample denied any history of CSA (84%). For male participants who did indicate having experienced some form of CSA, scores ranged from 1-34 ($M = 1.03$, $SD = 4.42$). Notably, only 11 men indicated any form of CSA, and interactions between setting and race, and experimenter contact and race could not be estimated due to insufficient data. The model

had a satisfactory goodness-of-fit (1.62) as defined by deviance/df statistics. Non-White male participants reported more CSA experiences than White participants [*B* (95% confidence intervals) = -2.709 (1.183, -5.028); $P < 0.05$, odds ratio = 0.07], with no other significant predictors identified. Notably, a comparison of the pattern of men's and women's mean CSAM scores across conditions shows some signs of convergence that might have been better captured for men with a larger sample (see table 12).

Table 12
Comparison of men's and women's estimated marginal mean scores ± SEM on CSAM

	Men	Women
Experimenter Contact		
Low	1.21 ± 0.41	3.74 ± 0.39
High	0.26 ± 0.12	2.48 ± 0.28
Setting		
Out of lab	0.98 ± 0.33	3.56 ± 0.38
In lab	0.32 ± 0.13	2.61 ± 0.28
Mode		
Paper	1.08 ± 0.39	4.77 ± 0.49
Computer	0.29 ± 0.11	1.95 ± 0.22
Race		
White	0.23 ± 0.11	2.69 ± 0.27
Non-White	1.32 ± 0.41	3.45 ± 0.38
Mode*Race		
Paper*White	0.58 ± 0.28	4.09 ±0.56
Paper*Non-White	2.00 ± 1.09	5.55 ± 0.85
Computer*White	0.10 ± 0.06	1.76 ± 0.25
Computer*Non-White	0.87 ± 0.35	2.15 ± 0.36
Setting*Race		
Out of lab*White	0.44 ± 0.20	2.54 ± 0.35
Out of lab*Non-White	2.17 ± 1.10	4.99 ± 0.80
In lab*White	0.12 ± 0.08	2.84 ± 0.41
In lab*Non-White	0.81 ± 0.40	2.39 ± 0.39
Experimenter contact*Race		
Low*White	0.60 ± 0.24	3.14 ± 0.42
Low*Non-White	2.40 ± 1.29	4.45 ± 0.70
High*White	0.09 ± 0.07	2.30 ± 0.34
High*Non-White	0.72 ± 0.36	2.68 ± 0.47

I also looked at the CSAM as a dichotomous measure of CSA. Participants who reported any history of CSA on the CSAM were coded as "CSA positive," and those who did not report any experiences of CSA were coded as negative. A logistic regression was run, using the dichotomous version of the CSAM as the dependent variable. Race, Contact, Setting, and Mode were entered in step 1, and interactions between race and each of the experimental variables were entered in step 2. The regression revealed that female participants were more likely to acknowledge CSA history in low contact (37%) than high contact conditions (31%; $\beta = 0.88$, odds ratio = 2.41, $p < 0.05$). There was also a race by mode interaction ($\beta = -0.26$, odds ratio = 0.33, $p < 0.05$). Follow up analyses suggested that White participants were less likely to disclose CSA in computer based conditions (22%) than in pencil-and-paper conditions (32%), χ^2 (1, $N = 110$) = 31.67, $p < .001$. Non-White participants demonstrated no significant differences in disclosure likelihood across mode, χ^2 (1, $N = 146$) = 1.99, $p = .158$. Using the dichotomous version of the CSAM, there were no significant main effects for race, setting, or mode, and no other significant interactions. The logistic regression was repeated with male participants, and no significant main effects or interactions were observed.

The final measure used to test hypothesis 2 was the Sexual Experiences Survey (SES-SFV). Scores on the SES-SFV were totaled and frequency scores ranged from 0-3 across 29 different types of sexual victimization. Higher scores reflect more instances of sexual victimization, and scores for women ranged from 0-71, with 42.9% reporting no history of sexual victimization ($M = 6.35$, $SD = 11.82$). SES-SFV scores were converted into a dichotomous variable which distinguished between participants reporting any

history of adult sexual victimization and those who denied an adult sexual victimization history.

A logistic regression was conducted with the dichotomous SES-SFV variable serving as the dependent variable. Race, Contact, Setting, and Mode were entered in step 1, and interactions between race and each of the experimental variables were entered in step 2. The regression revealed a race by mode interaction ($\beta = 1.08$, odds ratio = 2.94, $p < 0.05$), with White participants being less likely to report sexual victimization history (58%) in computer based conditions than they were in pencil-and-paper conditions (69%), χ^2 (1, N = 146) = 9.76, p = .002. There were no significant mode differences for the non-White participants.

For male participants, 61.7% reported no history of sexual victimization. The logistic regression was repeated with male participants, and no significant main effects or interactions were observed. An examination of effect sizes revealed a moderate odds ratio for the interaction between race and setting; however, a huge confidence interval again suggested that results could not be trusted.

Across both measures of sexual victimization, effects were observed which provide partial support for my hypotheses related to the main effects of mode, experimenter contact, and setting. Further, race was found to be a moderator for a number of effects. Consistent with findings from the sexual behavior measures, it appears that pencil-and-paper surveys are more likely than computer based surveys to lead to reporting of sexual victimization amongst women. Few effects were observed for male

participants, though this is almost certainly a consequence of attempting to measure uncommon experiences within a small sample.

Hypothesis 3. I predicted that low experimenter contact, out of lab completion, and computer based inquiry mode would be associated with more permissive or liberal attitudes towards sex than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode. Further, I expected that these effects would be moderated by race, such that the differences across condition would be greater for non-White participants than for White participants. I tested these hypotheses using the Sociosexual Orientation Inventory (SOI), and the Sexual Opinion Survey (SOS).

The SOI posed something of a challenge for statistical interpretation. A total score was calculated for the SOI as outlined by the measure's authors (Webster & Bryan, 2006). Higher scores on the scale indicated more permissive attitudes towards sex. Because the measure contains several count variables, it had a concentration of responses around zero and outliers on the high end of the distribution. However, because the measure also contained likert-type items, it could not be run through a Poisson or negative binomial model. A log transformation was used to bring the distribution of the total score on the SOI into normality in order to allow for a 2 (Race; White versus non-White) X 2 (Experimenter Contact; high versus low) X 2 (Setting; in lab versus out of lab) X 2 (Mode; computer versus pencil-and-paper) univariate ANOVA with the total score on the SOI serving as the dependent variable. This analysis was repeated separately for male and female participants.

There were no significant main effects found for women, but there were a number of significant interactions. The interaction between experimenter contact and setting was significant, F (1, 239) = 6.02, p = .017, η_p^2 = 0.024, and suggested that female participants in low contact, in lab conditions (M = 1.48; SE = 0.27) reported being significantly more accepting of casual sex than those in high contact, in lab conditions (M = 1.37; SE = 0.31), p = 0.04. No significant difference was observed for out of lab conditions as a function of experimenter contact (see figure 5). Additionally, the interaction between experimenter contact and race was also significant, F (1, 239) = 11.04, p = .001, η_p^2 = 0.04, such that White participants in low contact conditions reported being more accepting of casual sex (M = 1.40; SE = 0.04) than those in high contact conditions (M = 1.52; SE = 0.04), p = .04. The reverse was observed amongst non-White participants, who reported more acceptance in the low contact conditions (M = 1.51; SE = 0.05) than in the high contact conditions (M = 1.35; SE = 0.05, p = .003. (See figure 6).

Figure 5

Women's Total SOI Score (Log Transformed) as a Function of Experimenter Contact and Setting

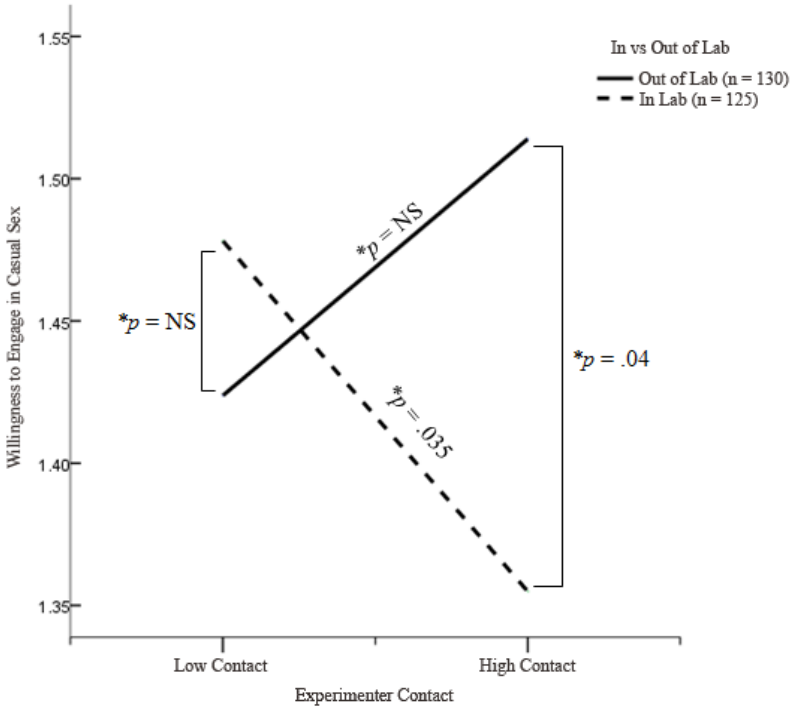
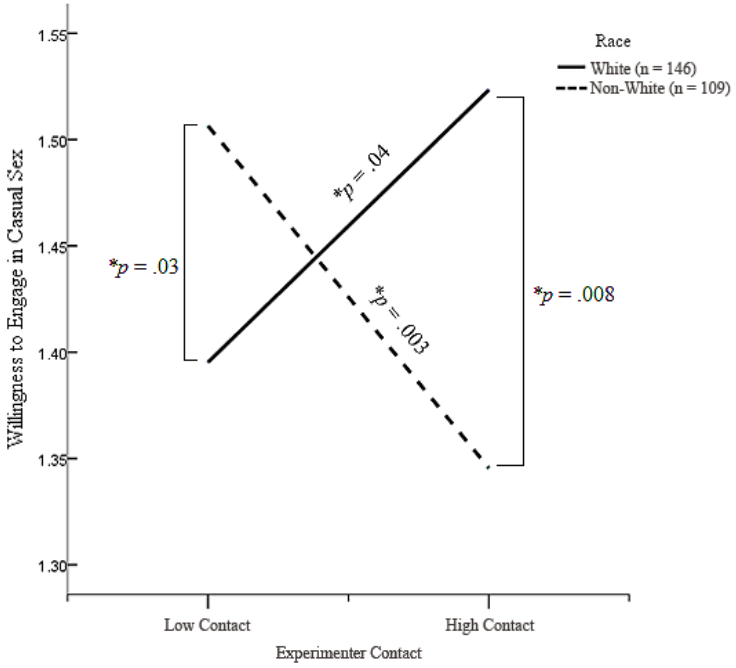


Figure 6

Women's Total SOI Score (Log Transformed) as a Function of Experimenter Contact and Race



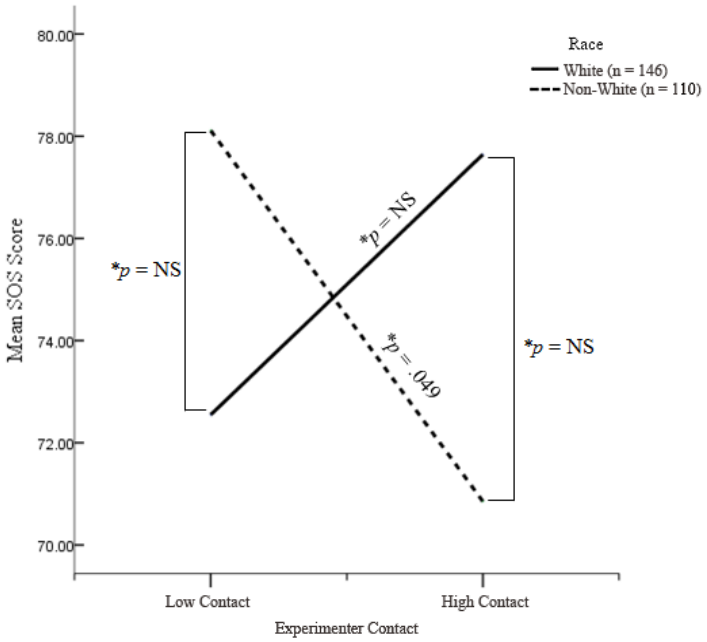
I ran the same analysis with the men in my sample, but the distribution of male responses could not be normalized by the logistic transformation. For the purposes of exploration, I can report that the model did not reveal any significant main effects or interactions for men.

The Sexual Opinion Survey (SOS) consists of 21 statements of which participants can indicate their degree of agreement on a 7-point scale ranging from "strongly agree" to "strongly disagree." Ten items on the SOS measure erotophobia, or negative attitudes about sex, and 11 items measure erotophilia, or positive attitudes towards sex. By subtracting the erotophobia score from the erotophilia score, I calculated a total score, with higher values indicating more positive attitudes related to sex. I then ran a 2 (Race; White versus non-White) X 2 (Experimenter Contact; high versus low) X 2 (Setting; in lab versus out of lab) X 2 (Mode; computer versus pencil-and-paper) univariate ANOVA with the total calculated score on the SOS serving as the dependent variable. This analysis was repeated separately for male and female participants.

For women, a significant main effect was observed for mode, F (1, 240) = 13.13, $p < .001$, $\eta_p^2 = 0.052$, with significantly more erotophilic views being reported on pencil-and-paper based surveys (M = 80.49; SD = 23.63) than surveys completed via computer (M = 69.46; SD = 26.10), $p < .001$. There was also an interaction between experimenter contact and race that approached significance (see Figure 7), F (1, 240) = 3.80, p = 0 .052, $\eta_p^2 = 0.016$, and trended towards suggesting that non-White participants (M = 78.11; SE = 3.27) reported more erotophilic views of sex in low contact conditions as compared to high contact conditions (M = 70.85; SE = 3.51), p = .05. There were no significant differences as a function of experimenter contact for White participants.

Figure 7

Women's Total SOS score as a Function of Experimenter Contact and Race



The same analysis was run for male participants, though no significant main effects or interactions were observed for any of the variables. However, moderate effect sizes were observed for race as a main effect ($\eta_p^2 = 0.03$), for the interaction between experimenter contact and race ($\eta_p^2 = .03$), and for the interaction between setting and race ($\eta_p^2 = .03$), suggesting that non-significant results for men may have reflected inadequate power rather than a lack of experimental effect.

Overall, though the main effect for mode of inquiry on participants' reported attitudes towards sex was opposite the direction predicted in my hypothesis, with more liberal attitudes in the paper and pencil than in the computer condition, it is consistent with the effects which have been observed across other measures. I also observed an interaction which trended towards significance suggesting that experimental contact may have differential impact on participants' reported attitudes towards sex based on race.

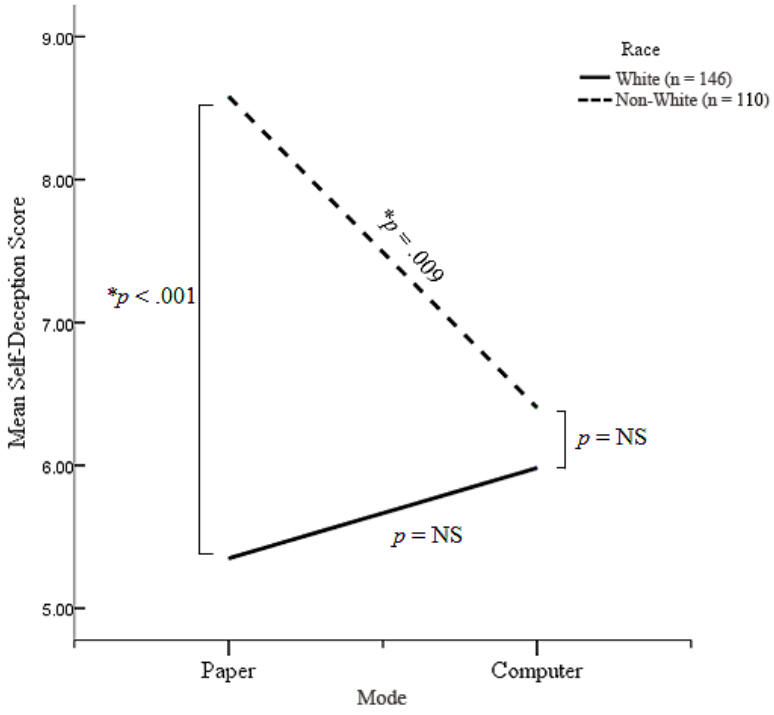
Taken together, my two measures of sexual attitudes provided partial support for my hypotheses among female participants. A race by experimenter contact interaction was observed for women for both measures, suggesting that non-White participants endorsed more positive or liberal views about sex in low experimenter contact conditions than they did in conditions of high contact. A main effect of mode was also observed on the SOS, which provided further support for greater levels of openness on pencil-and-paper surveys.

Hypothesis 4. I predicted that low experimenter contact, out of lab completion, and internet based inquiry mode conditions would be associated with lower rates of socially desirable responding than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode conditions. Further, I predicted that these effects would be moderated by race, such that the differences across condition would be greater for non-White participants than for White participants. The Balanced Inventory of Desirable Reporting (BIDR) was used to test this hypothesis. The BIDR can be scored to measure both Self-deception and Impression Management, two separate forms of social desirability. Lower scores on each index are suggestive of lower levels of that form of social desirability.

Both forms of social desirability were examined using a 2 (Race) X 2 (Experimenter Contact) X 2 (Setting) X 2 (Time 1 Mode) univariate ANOVA, and both ANOVAs were run separately for men and women. For self-deception, women displayed a main effect of race, F (1, 248) = 10.12, $p < 0$ .01, $\eta_p^2 = 0.039$, such that non-White participants engaged in significantly more self-deception (M = 7.39; SD = 4.49) than White participants (M = 5.59; SD= 3.86), p =.001. However, this main effect was

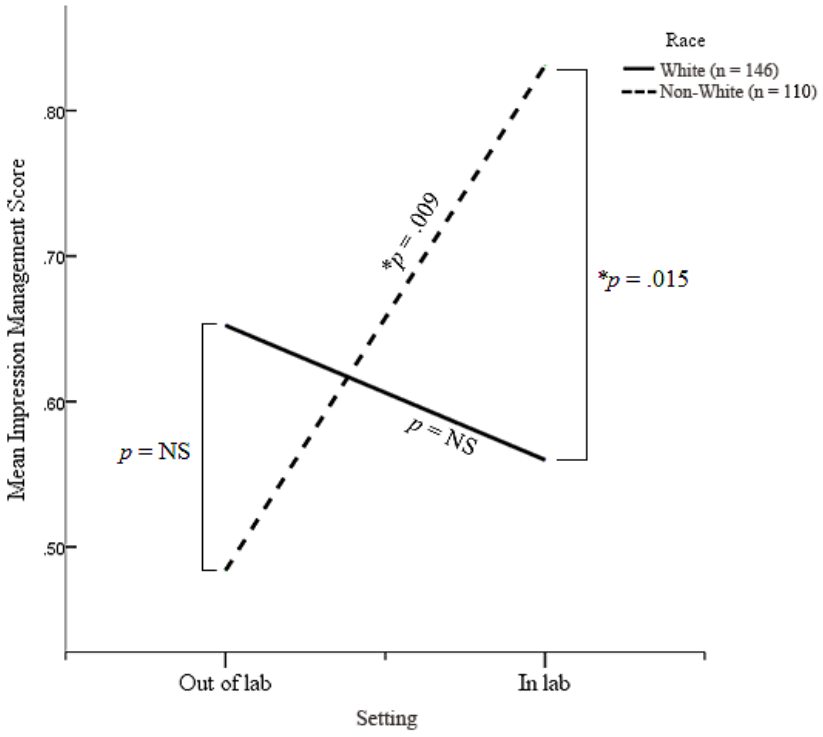
moderated by mode, F (1, 248) = 6.40, $p < 0$ .05, $\eta_p^2 = 0.025$, suggesting that non-White participants (M = 8.49; SD = 4.73) engaged in more self-deception on pencil-and-paper surveys than they did on computer based surveys (M = 6.27; SD = 3.95), p = .009. There were no significant mode differences for White participants. Follow up analyses suggest that the main effect for race is largely explained by differences observed in pencil-and-paper conditions (see figure 8). The same analyses were run for men, though no significant main effects or interactions were found. However, moderate effect sizes were observed for mode as a main effect ($\eta_p^2 = 0.024$), suggesting that non-significant results for men may reflect a sample that is inadequately powered to capture mode related effects.

Figure 8
Women's Total Score on the Self-Deception Index of the BIDR as a Function of Mode and Race



Another 2 X 2 X 2 X 2 univariate ANOVA was run with impression management as the dependent variable. No significant main effects or interactions were observed for women. A log transformation was required in order to bring the male distribution into normality. No significant main effects were observed for men using the transformed total score on the impression management index. However, an interaction between setting and race was observed, F (1, 63) = 7.78, p < 0 .01, η_p^2 = 0.11, which suggested that non-White participants engaged in more impression management in lab based conditions (M = 0.82; SD = 0.20) than they did out of lab conditions (M = 0.51; SD = 0.39), p = .01. There were no significant differences in setting for White participants. (See figure 9).

Figure 9
Men's Total Score on the Impression Management Index of the BIDR (*Log Transformed) as a Function of Setting and Race*



Overall I found partial support for my hypothesis, with setting (for men) appearing to have some effect on socially desirable responding through interactions with race. Further, Mode (for non-White women) was found to have a moderating effect which was consistent with the direction that I had initially predicted, with greater rates of self-deception observed on pencil-and-paper based surveys than those observed via computer. However, while this is consistent with my initial hypothesis, it is somewhat inconsistent with the mode related effects seen elsewhere, with participants appearing to be generally more open on pencil-and-paper surveys than they were on computers. I did not see any support for my hypotheses related to main effects of any of the independent variables, and observed effects were inconsistent across genders and across the respective forms of social desirability.

Hypothesis 5. I predicted that low experimenter contact, out of lab completion, and computer based inquiry mode would be associated with higher rates of perceived anonymity, confidentiality, and accuracy than high experimenter contact, in lab completion, and pencil-and-paper inquiry mode. The measure used to assess this hypothesis was only included for "time 2" surveys. I had 75 participants complete the "time 2" follow up surveys (57.3% White). The majority of the follow up sample was also female (74.7%). No significant differences were observed based on race or gender for any of the perception measures (see table 13).

Table 13
Participants' Scores on Perception as a Function of Race and Gender

	non-White M (SD)	White M (SD)	Male M (SD)	Female M (SD)
Perceived Anonymity	18.52 (7.9)	19.14 (9.26)	19.11 (8.86)	18.80 (8.76)
Perceived Confidentiality	19.71 (9.94)	19.65 (10.52)	20.05 (10.02)	19.55 (10.36)
Perceived Accuracy	19.31 (8.34)	19.93 (9.50)	19.89 (9.50)	19.59 (9.19)

I also compared the mean scores of each of the nine dependent variables for time 2 completers and non-completers using an independent-samples t-test in order to determine if there were any significant differences in self-report based on participant's completion of a second survey at time two. No significant differences were observed between the mean scores on any of the dependent variables (see table 14). This suggests that the time 2 completers do no differ systematically from non-completers on any of the primary variables of interest.

Table 14
T-test Comparing Time 2 Completers and Non-completers on Dependent Variables

	Mean (SD)		t	P
	Complete (N = 75)	Non-Complete (N = 261)		
Sexual Behavior Measure	12.99 (5.16)	13.31 (6.09)	0.42	0.67
Sexual Risk Survey	55.01 (68.01)	49.07 (69.47)	-0.66	0.51
Sexual Strategies Scale*	0.51 (0.50)	0.39 (0.49)	-1.74	0.09
Child Sexual Abuse Measure*	0.32 (0.47)	0.29 (0.46)	-0.48	0.63
SES-SV (Adult Victimization)	7.89 (13.58)	4.65 (10.07)	-1.92	0.06
Sexual Orientation Inventory	40.76 (36.98)	46.90 (86.52)	0.60	0.55
Sexual Opinion Survey	73.48 (20.43)	75.61 (25.49)	0.75	0.45
BIDR: Impression Mgmt.	5.57 (3.30)	5.64 (3.17)	0.18	0.86
BIDR: Self-Deception	5.79 (3.64)	6.53 (4.20)	1.37	0.17

*Mean score reflect the dichotomous versions of these variables

Because of the small sample size, for the purposes of exploratory analysis of hypothesis 5, men and women were combined and race effects were not considered. Each perception construct consisted of four likert-type items, each with a seven point scale ranging from "strongly disagree" to "strongly agree." Totaling the four items for each construct yielded scores ranging form 0-28. Visual inspection of the combined scores for

anonymity, confidentiality, and accuracy revealed clustering around the low and high ends of the distributions. As such, a dichotomous variable was calculated for each of the three perception variables with total scores of 0-14 indicating general disagreement and scores of 15-28 indicating general agreement. This allowed for me to examine the impact of mode, setting, and experimenter contact on high versus low perceptions of anonymity, confidentiality and accuracy of self-report.

Three separate logistic regressions were run in order to examine these relationships. For perceived anonymity, the regression revealed that participants were much more likely to perceive surveys as anonymous in pencil-and-paper based conditions (92.5%) as compared to computer-based conditions (40%; $\beta = 3.86$, odds ratio = 47.41, $p < 0.001$), and less likely to perceive them as anonymous in high contact conditions (18.9%) as compared to low contact conditions (44.7%; $\beta = -2.61$, odds ratio = 0.72, $p < 0.01$). Setting was not found to be significantly related to perceived anonymity.

In terms of perceived confidentiality, the logistic regression revealed that participants were much more likely to perceive surveys as being confidential in pencil-and-paper based conditions (68.5%) as compared to computer based conditions (0.05%; $\beta = 4.46$, odds ratio = 86.22, $p < 0.001$), and less likely to perceive them as confidential in high contact conditions (24%) as compared to low contact conditions (42%; $\beta = -2.20$, odds ratio = 0.11, $p < 0.5$). Again, setting was not found to be significantly related to perceived confidentiality.

For perceived accuracy, the logistic regression revealed no significant relationships between mode, experimenter contact, or setting and levels of perceived accuracy. Overall, the finding that participants perceived pencil-and-paper based conditions to be much more confidential and anonymous than computer based conditions was not surprising, given that participants consistently reported more sexual behavior and more permissive or liberal attitudes in pencil-and-paper conditions. Similarly, it is not surprising that participants found conditions in which they interacted directly with experimenters to be less anonymous and less confidential than low contact conditions in which they did not see or hear from an experimenter directly. Though the mode related effects ran counter to my initial hypothesis, they are consistent with the effects that I would expect to see given participant's approach to other measures in the study.

Discussion

The central impetus for this project was to clarify the impact of inquiry mode on participants' self-report of sexual behaviors, experiences, and attitudes with consideration for other methodological variables, specifically setting of completion and level of experimental contact. Five out of the eight sex-related dependent measures that were evaluated demonstrated main effects for inquiry mode, and all but two measures demonstrated mode related main effects or interactions (see tables 15, 16); thus, this study provides further support for the notion that inquiry mode is a methodological variable to which researchers need to attend. However, a detailed examination of my findings also highlights the complicated relationship between methodological and demographic factors, and the variable impact that these factors have on self-report of sensitive sexual information. While the results of this study cannot be distilled into a

"perfect formula" for assessing sexual behavior or attitudes, they draw attention to the importance of factors that are often taken for granted in sex research.

Sexual Behavior

Of all the sex related content focused on in this study, sexual behavior may be the most well studied in relation to inquiry mode. Many studies have sought to determine what impact various modes of inquiry have on participants' willingness to endorse various forms of sexual behavior. I looked at a range of socially acceptable and unacceptable sexual behaviors with the SBM and frequencies of sexual risk behavior with the SRS. For women, both measures suggested an effect for mode, though the effect for the SBM only trended towards significance. For both sexual behavior measures, there were indications that female participants were more willing to acknowledge behavior on pencil-and-paper surveys than they were on surveys which were completed on the computer. This finding was surprising, as it ran counter to my hypothesized direction of effect and counter to what has been observed in a number of other studies (e.g., Brown & Vanable, 2009, Feigelson & Dwight, 2000). However, in the Brown and Vanable (2009) study, mode differences were only found for two individual questions related to specific behaviors (unprotected oral sex and recent sexual partners), whereas my measures served to capture a much wider range of behaviors. Further, the Feigelson and Dwight (2000) meta-analysis of studies that were conducted prior to 2000 likely reflects an outdated view of technology. At the time, the authors suggested that participants' fear or

Table 15
Summary of the Significant Findings Related to the Impact of Experimental Variables on Women's Self-Report

Measure	Statistically significant main effects or interactions
Sexual Behavior Measure	Main: Race (White more) Mode (paper more; $p = 0.06$) Interaction: Race * Experimenter contact (non-White lower than White in high contact) Race * Setting (non-White lower than White in lab) Race * Mode (non-White lower than White on computer; non-White lower on computer than pencil-paper)
Sexual Risk Survey	Main: Mode (paper more) Interaction: Race * Mode (Non-white significantly lower on computer)
Sexual Strategies Scale	Main: Race (White more) Mode (paper more) Interaction: Race * Mode (Non-White lower on computer; White lower on pencil-paper)
Child Sexual Abuse Measure	Main: Mode (paper more) Experimenter contact (low contact more) Setting (out of lab more) Interaction: Race * Setting (non-White higher out of lab than in lab) Race * Mode (White higher on pencil-paper than computer)
SES-SFV	Interaction: Race * Mode (White higher on pencil-paper than computer)
Sexual Orientation Inventory	Interaction: Experimenter contact * Setting (Low contact, in lab more liberal than high contact, in lab) Race * Experimenter contact (non-White more liberal in low contact; white vice versa)
Sexual Opinion Survey	Main: Mode (paper more positive) Interaction: Race * Experimenter contact (non-White more positive in low contact, $p = .052$)
BIDR: Self-Deception	Main: Race (non-White more than White) Interaction: Race * Mode (non-White more self-deception on paper than computer, non-White more than White on paper)
BIDR: Impression management	No significant effects.

*SES-SFV refers to Sexual Experiences Survey Short Form Victimization, BIDR refers to Balanced Inventory of Desirable Reporting

Table 16
Summary of the Significant Findings Related to the Impact of Experimental Variables on Men's Self-Report

Measure	Statistically significant main effects or interactions
Sexual Behavior Measure	No significant effects.
Sexual Risk Survey	Main: Mode (paper less) Exp_cont (High contact more) Interaction: Race * Exp_cont (High contact more for non-White, Low contact more for White)
Sexual Strategies Scale	No significant effects.
Child Sexual Abuse Measure	Main: Race (non-White more)
SES-SFV	No significant effects.
Sexual Orientation Inventory	No significant effects.
Sexual Opinion Survey	No significant effects.
BIDR: Self-Deception	No significant effects.
BIDR: Impression management	Interaction: Race * Setting (non-White more in lab)

*SES-SFV refers to Sexual Experiences Survey Short Form Victimization, BIDR refers to Balanced Inventory of Desirable Reporting

misunderstanding of computers might lead them to feel motivated to provide more candid responses; a view that is not likely to be reflected by modern college age computer users. Interestingly, in spite of growing experience and comfort with technology, it appears that participants today are more comfortable disclosing sensitive sexual behavior on pencil-and-paper surveys than they are on computers. Given the growing popularity of digital survey research, researchers who are interested in particularly sensitive or uncommon behaviors may need to be mindful of this tendency when designing future studies.

Along with the main effects observed for mode, there were a number of race related interactions. For the SBM, race moderated each of the experimental variables. Non-White participants reported more behaviors than White participants in low contact, out of lab, pencil-and-paper based conditions but there were no racial group differences in the in high contact, in lab, or computer based conditions. Though I initially predicted that computer based surveys would yield more reported behaviors than paper-and-pencil amongst non-White participants, this reversal is consistent with the aforementioned reversal of the main effect for mode. In addition, a race by mode effect was also observed for the SRS, which suggested that non-White participants reported significantly fewer behaviors on computer based surveys than they did on those completed via pencil-and-paper. This is the first study that I am aware of that has considered race effects when examining mode of inquiry. Clearly, for sexual behavior, race is an important factor to consider when evaluating mode related effects, with much of the variability across conditions being accounted for by variability in the responding of non-White participants.

While the effect sizes observed for the significant mode related effects are considered to be small (η_p^2 = 0.02-0.03), I believe that these are still very meaningful

effects. In terms of the SBM, while the most common behaviors (e.g., kissing someone on the mouth) are likely to be captured regardless of methodology, these are not typically the behaviors of greatest interest to researchers. Rather, it is often the more uncommon behaviors (e.g., anilingus, anal sex) that have more clinical value in sex research, due to greater inherent risk of STIs. Similarly, with the SRS, small effects can be important when studying sexual behavior counts, particularly when those effects are more concentrated in specific demographic groups. For example, if researchers find a difference in sexual risk behaviors between racial groups, then they may wrongly assume that there are racial differences in the behavior when in fact, those differences might not be found (or might even be reversed) under different methodological conditions.

In terms of understanding the relationship between methodological factors and men's self-report of sexual behavior; it is much more difficult to draw any clear conclusions based on this study due to the very small sample of men. No significant main effects or interactions were observed for men on the SBM, and the main effect of mode on the SRS suggests that men reported more sexual behaviors on computer-based surveys, running counter to all of the other mode related effects observed in the study. It is possible that men are reporting more behavior in computer based conditions for the same reasons that women are reporting fewer behaviors. It is often considered more socially acceptable for men to engage in sexual behavior than women, and a desire to portray oneself in a more positive light could lead men and women to adjust their reporting in different directions, with men over-reporting and women under-reporting sexual behaviors (e.g., Crawford & Popp, 2003). However, main effects for mode were not observed for men on any of the other measures, and it is difficult to draw conclusions

based on the SRS alone. There has been some suggestion that men are less influenced by social pressures than women when it comes to reporting sensitive behaviors (e.g., Kays, et al., 2012), which would indicate that a much larger sample of men would be necessary in order to get a clear understanding of the methodological factors impacting men's reporting of sexual behavior.

Sexual Victimization and Perpetration

This is the first study of which I am aware that evaluates self-reported sexual victimization and perpetration in the context of the three methodological factors evaluated in this study. Given that acknowledgement of perpetration and victimization have important implications for prevention and intervention efforts; I see this as an area that is particularly deserving of attention. Reporting of victimization helps researchers to understand the scope of the problem and identify consequences and correlates of sexual victimization. Reporting of sexual perpetration aids researchers in better understanding the correlates of sexually aggressive behavior, which can be invaluable information for constructing and evaluating prevention programs. Further, because perpetration is particularly socially undesirably and, in some cases illegal, reporting may be heavily influenced by participants' perception of anonymity and confidentiality. Methodological factors are also particularly relevant for these experiences because the experiences are relatively rare and infrequent (although not as rare or infrequent as one would hope). As such, missing only a small number of victimization or perpetration experiences may amount to a substantial proportion of the experiences that have taken place.

I examined two measures of victimization in order to capture self-reported sexual victimization experiences in childhood and adulthood. In terms of childhood victimization, the CSAM demonstrated main effects for women on each of the methodological factors that I manipulated in this study. Consistent with the effect observed with other measures in this study, participants were more likely to acknowledge child sexual abuse (CSA) on pencil-and-paper surveys than they were when completing surveys on the computer. Further, they were more likely to report CSA in low contact conditions and out of lab conditions than in high contact and in lab conditions, respectively. These main effects are generally in line with the effects which I predicted, short of the aforementioned reversal in the direction of mode effects. There was also a race by setting effect for women, which suggested that non-White participants were more likely to report CSA in out of lab than in lab conditions; there was no significant effect of setting for White participants.

The continuous measure of child sexual abuse demonstrates the most complete support for my hypothesis in that mode, experimenter contact, and setting all proved to be significantly impactful factors on female participant's willingness to disclose CSA, with pencil-and-paper, low contact, and out of lab conditions being associated with more disclosure than computer, high contact, and in lab conditions, respectively. This may suggest that CSA is a particularly sensitive experience about which people are particularly guarded. This is not necessarily surprising, given the guilt and shame that are often associated with CSA (e.g., Dorahy & Clearwater, 2012). Notably, all but the main effect of mode and a race by mode interaction dropped out of significance when the CSAM was considered as a dichotomous measure of CSA experiences, independent of

severity. This suggests that less severe (and more common) experiences might be less impacted by methodological factors such as setting and experimenter contact. However, the mode related effects appear to be strong enough to reach significance regardless of the severity of the CSA experience.

Reporting of adult sexual victimization was also suggestive of a race by mode interaction. White female participants were significantly more likely to report a history of adult victimization when completing pencil-and-paper as compared to computer based surveys. This effect was similar to that which was observed with CSA.

Along with the two measures of sexual victimization, I also examined participants' self-report of sexual perpetration or sexually coercive behaviors using the SSS. For women, I found main effects for race and mode, suggesting that White participants and participants completing pencil-and-paper measures were more likely to acknowledge use of sexually coercive strategies than non-White participants and participants completing computer measures, respectively. The mode related effect is consistent with that which has been noted elsewhere in this study, and the race related moderation also reflects what appears to be a general tendency of non-White participants to be particularly conservative in reporting behaviors on computer based surveys as compared to paper-and-pencil surveys.

Interestingly it appears that the impact of methodological factors on self-report related to coercion perpetration parallel those which I observed in other domains as well (e.g., CSA, victimization, consensual sexual behavior). This suggests that people do not seem to actively employ a different strategy or approach for reporting sexual coercion

than they do when discussing other aspects of sexuality. Further, I did not see any significant effect for setting or experimenter contact, which suggests that these factors do not seem to impact acknowledgement of sexual perpetration in the same manner that they do for sexual victimization.

No effects for male participants were observed related to perpetration, and the only effect related to victimization was a main effect for race related to disclosure of CSA, with non-White participants reporting more CSA experiences than White participants. It is possible that the main effect for race reflects an increased risk for CSA amongst non-White participants, rather than a reporting bias. This seems plausible, given the tendency of non-White participants to report fewer behaviors or experiences than White participants on other sex related measures in this study. Though findings related to race as a potential risk factor for childhood maltreatment have been mixed, there are some indications that African American children experience higher rates of maltreatment than White American children, though observed differences may be confounded by socioeconomic differences (e.g., Lee et al., 2012).

Though no effects were observed for men on the CSAM related to my experimental variables, it is likely that this is a consequence of an insufficiently powered sample. A comparison of the CSAM mean scores for male and female participants as a function of condition suggests converging effects related to mode, experimenter contact, and setting (see table 12). Conceptually, it would make sense that male and female participants experience disclosure of CSA similarly, and such effects might be better captured for men with a larger sample. Clearly, more research is needed in order to understand the impact of methodological factors on men's disclosure of victimization and

perpetration.

Sexual Beliefs and Attitudes

This is also the first study that I am aware of that examines the impact of inquiry mode, experimenter contact, and setting on self-reported sexual beliefs and attitudes. This is particularly important given that beliefs and attitudes are inherently private and measurement depends completely on self-report. The SOI is the only sex-related measure included in this study that did not reveal any mode related interactions or main effects. Further, it is the only measure without any methodological main effects observed at all. While there is no obvious explanation for the absence of main effects, it is unique from other measures in that it includes behavioral counts ("With how many different partners have you had sex within the past year?"), behavioral forecasting ("How many different partners do you foresee yourself having sex with…), cognitive counts ("How often do you fantasize about having sex with someone other than.."), and scaled attitudinal statements (e.g., "sex without love is ok."). It is possible that the different types of items in the measure pulled for different types of methodological variance, resulting in effects "washing out" each other. Alternatively, it is possible that the different types of questions prevented participants from adopting a single mindset or approach to the measure.

The SOI did demonstrate a race by experimenter contact interaction for women, which supported my hypothesis related to the effect of experimenter contact. Non-White participants endorsed significantly more liberal attitudes (i.e., greater acceptance of casual sex) in low contact conditions than they did in high contact conditions. The effect was reversed for White participants, with more liberal attitudes being reported in high contact conditions. This suggests that the responding of both White and non-White

participants on the SOI was impacted by methodological factors although differentially so.

Main effects of mode were observed for women on the SOS, a measure of erotophilia versus erotophobia. Consistent with the trend observed across other measures, participants reported more positive views related to sex in pencil-and-paper conditions than they did on computer-based surveys. This suggests that participants completing pencil-and-paper surveys were willing to indicate stronger agreement with statements such as "masturbation can be an exciting experience" or disagreement with statements such as "I would not enjoy seeing an erotic movie" than participants completing computer surveys. This provides further support to the notion that pencil-and-paper surveys elicit more open or less socially-guarded responses from participants.

The SOS also demonstrated a race by experimenter contact interaction for women which was consistent with that observed in the SOI. Non-White participants endorsed significantly more positive views of sex and sexuality in low contact conditions than in high contact conditions. Interestingly, White participants reported significantly more positive views of sex and sexuality in high contact conditions than low contact conditions, accentuating the difference between the two groups.

The attitudinal measures present one of the few occasions when White participants showed evidence of being significantly impacted by the methodological variables manipulated in this study. Interpretation of this effect is challenging because the SOI and SOS are attitudinal measures; thus, what would be considered "desirable responding" is subjective. Undoubtedly, cultural factors influence attitudes towards sex,

and it is possible that the direction of influence differs for different racial groups. As such, it is possible that desirable responding efforts would lead some groups to bend towards conservative or erotophobic responses, and other groups to bend towards liberal or erotophillic responses. With this in mind, the significant differences observed across experimental contact conditions for both racial groups on the SOI and the corresponding trends observed with to SOS could potentially be explained by divergent efforts to provide culturally bound desirable responses in high contact conditions.

Yet again, the analyses that were run with men did not reveal any significant main effects or interactions related to sexual beliefs or attitudes. As has been previously mentioned, the lack of a significant effect may simply be the result of inadequate power. However, it is also possible that men are simply less variable or context dependent than women in their evaluation of attitudes toward sex. More research is clearly needed in order to draw any conclusions about the impact of methodological factors on men's attitudes towards sex.

Social Desirability

Previous studies that have sought to examine the impact of inquiry mode on socially desirable responding have yielded mixed results (Weigold et al., 2013). However, relatively few studies have actually employed formal measures of social desirability such as the BIDR. A recent study by Arne Weigold and colleagues (2013) is one of the few exceptions, in that they used the Marlowe-Crowne Social Desirability Scale (MCSDS) to directly assess for socially desirable responding. The authors also made efforts to manipulate experimenter contact, setting, and mode, while keeping other

factors constant. The authors found no differences across any condition and suggested that participants engaged in social desirability at equivalent rates, across all conditions.

Though the Weigold et al. study has a significant amount of overlap with this study in terms of evaluating social desirability, it also has a number of limitations. First, though the MCSDS is a well-known and established measure of social desirability, it is also somewhat dated, and does not capture separate factor scores for self-deception and impression management. Further, the Weigold study did not consider race, which I have consistently found to be an important factor in interaction with other measures evaluated in my study.

Indeed, in my study, with female participants, I found a significant main effect for race, which suggested that non-White participants engaged in more self-deception efforts than White participants. However, this significant main effect must be interpreted in light of an interaction between race and mode, which suggested that differences between racial groups on self-deception were largely occurring in pencil-and-paper conditions. These differences could not have been captured by Weigold and colleagues, as they did not consider self-deception as a separate outcome or race as a moderator. This highlights the challenges of capturing mode related effects, due to the complicated nature in which these effects seem to be manifest.

The race by mode interaction observed for non-White female participants on the self-deception scale is also notable in that it ran counter to the direction observed across the sex related dependent variables. Based on the higher rates of sexual behavior and more permissive attitudes observed on pencil-and-paper surveys, it would be intuitive to

anticipate that higher rates of self-deception would be observed on computer based surveys. These results suggest that for non-White female participants, pencil-and-paper surveys promoted higher levels of self-deception, more reported behaviors and endorsement of more permissive attitudes about sex. I ran correlational analyses between the total self-deception index score and each of the dependent sex related variables, and found no significant correlations for non-White female participants, suggesting that self-deception was not the direct cause of the differences observed across conditions. However, it remains unclear why this reversal was observed. It is possible that participants felt more comfortable in pencil-and-paper conditions, and in turn were more willing to explore (through self-report) their ideal level of sexual-behavior and more preferred sexual attitudes. Further research is clearly needed in order to better understand the relationship between self-deception and participants' reporting of sensitive behaviors and attitudes.

There were no significant differences for any of the methodological variables on women's use of impression management on the BIDR, suggesting that none of the conditions led to higher or lower levels of impression management. This is notable, given that effects were observed for at least one of the methodological variables across all but one of the other measures examined in this study, including the aforementioned measure of self-deception. Further, one would intuitively expect that motivation to present oneself in a more favorable light is at least partially responsible for the differences that I observed in sexual self-report across conditions.

One explanation for the lack of an effect for impression management on the BIDR is simply that participants were not engaging in any significant amount of impression management across any of the conditions and differences that were observed on the sexuality-related measures were the result of some other mechanism, such as recollection bias. However, this seems somewhat unlikely, given that there is no straightforward way in which to expect that recall would differentially impact ones self-report of sex related behaviors or attitudes across the various conditions of this study. Alternatively, it is possible that self-deception plays a more important role than impression management in influencing participants' responding. This would suggest that participants in certain conditions might be more prone to introspection, and in turn might be more likely to alter responses to preserve a favorable self-impression. Again however, it seems far from parsimonious to conclude that participants in high experimenter contact conditions are more likely to engage in self-deception than those in low contact conditions but no more likely to engage in impression management than their counterparts who have no interactions with other people while completing their surveys.

A final possible explanation for the lack of an effect for impression management is that there is something unique about questions related to sex, which provokes a different reaction than sensitive questions about other topics. Responding to BIDR statements such as "I am not always a safe driver when I exceed the speed limit" may be a qualitatively different experience than responding to questions about one's history of condom use or sexual aggression. Much of the research related to inquiry mode has treated sex as just another sensitive topic and thrown questions related to sex in with other "sensitive" questions (e.g., questions about drug use, cheating on tests, annual

income) in an effort to identify any differences across mode. As has been previously noted, these efforts have yielded inconsistent findings. By focusing specifically on sex, I have been able to more consistently demonstrate mode related effects, along with effects related to experimenter contact and setting. Even when participants are not managing their impressions related to other aspects of their life, they may be much more likely to guard themselves related to disclosure of sex related information under certain experimental conditions (e.g., computer based surveys, contact with experimenters, completion in lab).

Perceptions

Difficulties with participant recruitment and high rates of attrition between time 1 and time 2 resulted in a small time 2 sample size; thus, my analyses related to participant's perceptions of anonymity and confidentiality need to be interpreted very cautiously. Following the second administration of the otherwise identical survey, participants who returned to complete the time 2 surveys were asked to answer twelve additional questions related to perceived anonymity, perceived confidentiality, and perceived accuracy of their own responses. I was interested in capturing participant's perceptions because previous research has indicated that participants may misperceive the degree and direction by which methodological factors impact their self-report (Bates & Cox, 2008). This also served as something of a manipulation check, to ensure that participants experienced the high and low contact conditions as I intended them to.

My findings related to the impact of methodological variables on participant perceptions were largely consistent with the effects that were observed across the other measures included in the study. Participants who completed time 2 measures evidenced significant main effects for mode and level of experimenter contact on measures of perceived anonymity and confidentiality. Not surprisingly, participants in pencil-and-paper conditions and low experimenter contact conditions were significantly more likely to perceive the study as highly anonymous and highly confidential compared to participants in the computer conditions and the high experimenter contact conditions, respectively. This finding provides support for the assumption I have made that when participants' perceive research conditions as highly anonymous or confidential, they tend to report more sexual behaviors and more liberal beliefs and attitudes related to sex.

Though previous research has shown differences in participants' perceptions of anonymity and confidentiality across mode, these perceptions did not correspond to differential responding on self-report measures (Bates & Cox, 2008). My ability to capture differences both in participant perceptions and in self-reports of sexual behaviors may be partially explained by my consideration of race as a moderating factor. Many of the significant differences that I observed were moderated by race and would have gone unnoticed had race not been considered. Further, the content of measures included in my study differed from that of Bates and Cox (2008). As has been previously mentioned, there may be something about sex related research that makes it more likely that participants censure or edit their responses in situations that they perceive to be less anonymous and less confidential.

Interestingly, in the Bates and Cox study, participants reported a belief that their responses were more accurate in some situations than others, but these perceptions did not correspond with significant differences across conditions. In my study, participants did not indicate any significant differences in the accuracy of their responding across condition, but the results of the study suggest that participants did approach responding differently depending on the condition they were in. As such, one point of convergence between this study and the Bates and Cox study is that participants' perception of accuracy does not seem to correspond to the degree to which responses are edited. This provides some support for the possibility that the differences observed across condition may not be the product of conscious efforts to present oneself in a certain manner but rather unconscious attempts to conform to social expectations.

It is somewhat surprising that setting had no significant impact on participants' perceptions of anonymity or confidentiality. Given that setting had a significant effect for a number of the sexual measures that I examined; one might expect to see some of that effect captured in participants' perceptions. Given that individuals completing the study out of the lab might have completed it in a variety of locations (home alone, in a crowded computer lab, etc.), it is possible that the uncontrolled nature of out of lab settings led to substantial variability in perceptions of anonymity and confidentiality in that condition and thus obscured any differences between in lab and out of lab conditions. Regardless of the specific mechanism, this finding may suggest that the impact of setting on participants' self-report is influenced by something other than their perceptions of anonymity or confidentiality.

It is important to reiterate the need for cautious interpretation of these results given that the subset of my participants who returned to complete time 2 may not be random. Of the 337 participants who completed time 1, only 113 of them returned to complete a second survey at time 2 (34%). Such a high level of attrition certainly raises concerns about the representative nature of these 113 participants. Demand characteristics are also a concern in interpretation of the perception variables. Participants who completed time 2 were likely aware of the fact that they had completed the exact same survey 1-2 weeks earlier, with the only difference being the mode through which the survey was presented to them. It would not be a leap to expect that some of the participants had some awareness of the nature of the study by the time they got to the end of the second survey. However, given that these results related to participant perceptions are largely in keeping with the findings related to sexual self-report, I can cautiously consider them as further evidence that participants' perception of anonymity and confidentiality across condition may contribute to their overall level of disclosure related to sensitive sexual information.

Summary

Taken together, these results suggest that experimenter contact, setting, and inquiry mode all independently impact people's reporting of sexual behavior. Though I did not find consistent effects of each factor across each measure, I did observe a fair amount of consistency in the direction of these effects across the measures examined. Generally speaking, low experimenter contact, out of lab, pencil-and-paper based surveys were likely to yield higher rates of reported behavior and expression of more liberal

attitudes and beliefs about sexuality as compared to high contact, in lab, and computer based surveys.

The other consistent finding which emerged is that race serves as a fairly reliable moderating variable for the effects of mode, setting, and experimenter contact on participants' self-reports. It appears that generally, non-White participants are more sensitive to changes in these experimental variables and more likely than White participants to report fewer behaviors and more conservative attitudes when completing computer based surveys, in laboratory environments, or while in the presence of experimenters.

The vast majority of the effects which I observed were isolated to female participants. While there are a number of possible explanations for this, my limited sample of men limits the amount of speculation I can make about sex related differences in reporting across conditions. There were some indications that men and women may have experienced similar shifts in reporting based on condition (e.g., CSAM; table 12), but other measures showed considerably less convergence between genders across conditions (e.g., SBM; table 8).

Examination of participants' perceptions of anonymity and confidentiality across research conditions suggested that perceived anonymity and confidentiality may have contributed to higher levels of reporting in some conditions as compared to others. Though directionality cannot be established, this relationship supports the hypothesis that people are more likely to report more sexual behavior and more liberal sexual attitudes when they perceive as a situation as highly anonymous and confidential. However, the

lack of a relationship between setting and participant perceptions suggest that factors outside of the perceptions anonymity and confidentiality are also likely impacting participants' responding across conditions. Finally, the lack of observed difference related to participants' perceptions of accuracy suggest that observed differences in level of responding may not be the result of active misrepresentation or motivated editing on the part of participants. The notion that this may not be a conscious process is further supported by the lack of observed differences in impression management across research conditions. Based on BIDR literature, it would be expected that intentional efforts to portray oneself in a favorable light would be captured by the impression management scale (e.g., Paulhus, 1984). Self-deception (which did show a race moderated effect for mode) is a process that is more commonly associated with unconscious censuring of responses in order to maintain a certain view of oneself.

This study marks an important step in the direction of understanding inquiry mode in the context of other methodological variables. I have established mode as a uniquely impactful methodological factor and have unexpectedly shown evidence that in sex research, pencil-and-paper surveys may have inherent advantages over computer-based surveys. Further, I have expanded on prior research by examining a much wider range of self-report sexual behaviors in relation to mode effects. Along with examining a wider range of self-reported sexual behaviors than prior mode studies, I provided the first study of self-reported sexual perpetration and victimization, and the first comprehensive examination of sexual attitudes as function of research mode. Also, this is the first mode related study to directly consider the moderating effect of race on these three methodological variables. Given that race related interactions frequently explained much

of the effects observed in this study, race is likely to be an important factor in fully understanding the impact of methodological variables on sex related self-reports.

Limitations

This study had a number of important limitations which must be considered. Obviously one of the most glaring limitations relates to my inability to recruit a sufficiently large sample of male participants in order to fully examine the effect of gender on reporting across conditions. Though consideration of male and female participants separately using parallel analyses allowed for some exploration of possible gender related effects, my sample of men was still far too small for me to anticipate capturing differences across condition if they did exist. Further, I was not able to directly compare men and women within the same analyses, which is unfortunate given that there are theoretical reasons why differential effects might be observed between genders. Finally, by doubling my analyses through parallel analyses for men and women, I effectively doubled my error rate and increased the risk of identifying effects by chance. However, concerns that some of my findings reflect Type I error are tempered by the fact that the significant results generally follow a consistent and expected pattern. Nevertheless, it will be important for future studies to ensure that adequate samples of men and women can be recruited in order to replicate my findings and to determine the degree to which they apply to an adequately powered sample of men.

An additional concern related to recruitment is the relatively low percentage of potential participants who actually scheduled, attended, and completed the first survey for this study. Of the 676 potential participants who initially expressed interest in the study,

only 337 completed the initial survey (50.0%). Further, only 113 participants (16.7 %) completed the study in its entirety. Though I was able to adjust my analyses to account for the high rates of attrition between time 1 and time 2 using a between- instead of a within-subject design, I cannot fully account for the substantial group of potential participants who did not follow through with scheduling their initial appointments. However, it is important to consider this number in the context of the recruitment process. Indication of interest simply involved participants clicking on a link requesting information to what many of them assumed to be an "online study." Many of the potential participants expected that they could fulfill their participation requirement within minutes after expressing initial interest in the study. Based on emails which experimenters received from potential participants, it appears that some of them opted not to participate after discovering that participation was more demanding than a traditional online study.

Another related limitation is the differential rates at which potential participants scheduled initial appointments based on inquiry mode and setting. Logistic regression analyses revealed that participants in out of lab conditions were significantly more likely to schedule appointments in computer based conditions than they were when assigned to pencil-and-paper conditions. This suggests that potential participants assigned to out of lab conditions were more willing to participate in a study that simply required them to click on a link in an email than they were when asked to walk up to a laboratory door to pick up a pencil-and-paper survey. While this serves as a limitation in this study, it is also provides an important consideration for future methodological decisions. Even if disclosure rates are higher in some conditions (e.g., paper-and-pencil versus computer),

the potential benefit of these higher reporting rates must be weighed against peoples' willingness to participate in the study in the first place.

Concerns related to random assignment are somewhat tempered by the fact that there were minimal differences in participation across condition after potential participants had gone as far as scheduling an initial appointment. Given the relative ease by which potential participants could indicate interest in a study through the university subject pool web portal, it appears that individuals with very low motivation may have been dissuaded prior to scheduling an appointment. However, the possibility exists that the participants who were willing to take the additional step of scheduling initial appointments for out of lab and pencil-and-paper measures differed in some way from the potential participants who elected not to participate after being assigned to these conditions.

Another limitation in this study was that I did not measure participants' sexual identity (I did measure sexual behavior with same- and other-sex partners as part of the SBM). One possible explanation proposed for race as a moderator is that concerns related to racial prejudice might lead non-White participants to be more sensitive to methodological differences which increase the likelihood that they might be connected to their responses in some way. Given that sexual minorities often face discrimination directly as a result of some of the sexual behaviors and attitudes being assessed in this study, it conceptually makes sense that sexual identity might moderate methodological effects in similar ways as those observed for race. Future research is needed in order to determine the degree to which sexual identity impacts the effects of methodical variables on sex related self-reports.

A final limitation worth noting is my treatment of non-White participants. While the practical necessities of my sample size dictated my decision to combine all non-White participants together, it is likely that the racial groups captured in the "non-White" participant category were fairly heterogeneous in their cultural background, beliefs, and experiences. Though it is likely that minority status played a role in the effects which I observed (and minority status would be relevant to any non-White participants regardless of race or ethnicity), it would be helpful to have a better understanding of the degree to which other cultural factors, unique to different minority groups, may have impacted responding across conditions. Given that my sample of non-White participants was predominantly comprised of African Americans (71.3%), it is likely that the observed race related interactions are largely the result of differences in responding between White and Black participants. Future examination of methodological factors should consider a more nuanced examination of race related effects.

In spite of the aforementioned concerns, this study still serves as one of the most well controlled explorations of inquiry mode of which I am aware. The experimental design employed and the random assignment of participants to conditions, allows me to confidently interpret the observed results as being directly related to the experimental manipulation of methodological variables. It is unlikely that any of the observed effects could be better explained by any of the aforementioned difficulties with assignment. Further the scope of this study makes it one of the most widely applicable methodological studies of sexual self-report of which I am aware. By considering sexual behavior, attitudes, perpetration, and victimization in the same study, I am able to present a more complete understanding of the impact methodological decisions can have across multiple

domains of sex research. Additionally, though a more complete consideration of gender, race, and sexual identity would be undoubtedly valuable, this study marks the first consideration of any of these demographic variables as possible moderators of mode related effects. Hopefully future studies will build on my findings in order to better elucidate the interaction between participant characteristics and methodological variables.

Future Directions

This study has a number of clear implications for sex research moving forward. All else being equal, researchers may want to consider employing pencil-and-paper based surveys, low experimenter contact, and out of lab completion when studying sexual behavior and attitudes. This is particularly relevant to researchers who are interested in sex related research with non-White populations. Online research will almost certainly be an important part of sex research in the future, regardless of any mode dependent effects on self-report. It is a low cost mechanism to quickly reach a wide range of people, and it circumvents the time consuming and error prone data entry process required with traditional pencil-and-paper surveys. However, this study provides further evidence that equivalence cannot be assumed across modes of inquiry. As technology continues to rapidly evolve, providing additional modes from which to collect data, this reminder becomes increasingly important.

It is helpful to recognize that experimenter contact and setting of administration are relevant factors in their own right, but the mechanism through which mode related differences emerge is still elusive. Further research is needed to understand what is experientially different for participants between completing a survey on a computer

versus a printed pencil-and-paper form. Identification of experiential differences is likely going to have applications not only for the currently popular modes of differences (e.g., pencil-and-paper, computer), but also for the modes which are likely going to be dominating research over the next decade (e.g., mobile phone and tablet based surveys). Given that research will never be able to keep pace with technological innovation, it is essential that researchers isolate the underlying factors which contribute to differences.

Beyond applying the results of this study to future research, these findings also have implications for clinical practice. It is well known that people are often resistant to disclosing sensitive information in the context of face-to-face interviews. This resistance can serve as a barrier for access to care, for example, if victims of sexual assault are not willing to disclose their experiences. The results of this study suggest that providing clients with a survey that they can take home is likely to maximize their willingness to disclose CSA history. Further, completion of pencil-and-paper surveys within the clinic may increase client's willingness to report adult sexual victimization experiences. While it is likely premature and impractical to suggest that clinician's adopt a practice of prescribing multiple self-report measures to be completed in multiple different settings, I can suggest that flexibility be considered. Perhaps by sending clients home with a survey and asking them to review it in the waiting room prior to turning it in, clinicians could increase the likelihood that clients feel comfortable in disclosing multiple forms of sensitive information and in turn increase the likelihood that they have access to effective care.

Further research is also needed to replicate my findings and extend them to better consider gender related effects, effects related to sexual identity, and the differential impact of various racial groups. Expanding our understanding of these variables will allow researchers and clinicians to further tailor their approach to asking questions in order to maximize the likelihood that they are best able to capture the attitudes, behaviors and experiences of the specific groups with whom they are working.

While this study served mainly to consider the impact of methodological decisions on self-reports related to sex and sexuality, it is possible that my findings have broader implications as well. It is possible that methodological decisions may be more impactful for non-White participants than for White participants, regardless of the specific type of sensitive questions that are being asked. Future research is clearly needed in order to better understand race related moderation of methodological effects in other domains as well.

References

Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family Therapy, 67,* 1012-1028. doi: 10.1111/j.1741-3737.2005.00191.x

Alexander, M. G., & Fisher, T. D. (2003). Truth and consequences: Using the bogus pipeline to examine sex differences in self-reported sexuality. *Journal of Sex Research, 40*(1), 27-35. doi:10.1080/00224490309552164

Bates, S. C., & Cox, J. M. (2008). The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviors. *Computers in Human Behavior, 24,* 903-916. doi: 10.1016/j.chb.2007.02.021

Brown, J. L., & Vanable, P. A. (2009). The effects of assessment mode and privacy level on self-reports of risky sexual behaviors and substance use among young women. *Journal of Applied Social Psychology, 39*(11), 2756-2778. doi:10.1111/j.1559-1816.2009.00547.x

Browning, J. R., Hatfield, E., Kessler, D., and Levine, T. (2000). Sexual motives and interactions with gender, Archives of Sexual Behavior, 29, 139-152.

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*(3), 345-375. doi:10.1086/297758

Catania, J. A., Gibson, D. R., Chitwood, D. D., & 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*(3), 339-362. doi:10.1037//0033-2909.108.3.339

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in Web- or internet-based surveys. *Educational and Psychological Measurement, 60*(6), 821-836. doi:10.1177/00131640021970934

Cowart-Steckler, D., & Pollack, R. H. (1998). The Cowart–Pollack scale of sexual experience. In C. M. Davis, W. L. Yarber, R. Bauserman, G. Schreer, & S. L. Davis

(Eds.), Handbook of sexuality-related measures (pp. 104–105). Thousand Oaks, CA: Sage Publications

Crawford, M., & Popp, D. (2003). Sexual double standards: A review and methodological critique of two decades of research. *Journal of Sex Research, 40*(1), 13-26. doi:10.1080/00224490309552163

DiLillo, D., DeGue, S., Kras, A., Di Loreto-Colgan, A., & Nash, C. (2006). Participant responses to retrospective surveys of child maltreatment: Does mode of assessment matter? *Violence and Victims, 21*(4), 410-424. doi:10.1891/vivi.21.4.410

Dorahy, M. J., Clearwater, K. (2012). Shame and guilt in men exposed to childhood sexual abuse: A qualitative investigation. *Journal of Child Sexual abuse, 21,* 155-175. doi: 10.1080/10538712.2012.659803

Feigelson, M., & Dwight, S. (2000). Can asking questions by computer improve the candidness of responding?: A meta-analytic perspective. *Consulting Psychology Journal: Practice and Research, 52*(4), 248-255. doi: 10.1037/1061-4087.52.4.248

Finkelhor, D. (1981). *Sexually Victimized Children*. New York, NY: The Free Press.

Fisher, W.A. Byrne, D., White, L.A., & Kelley, K. (1988). Erotophobia-erotophillia as a dimension of personality. *The Journal of Sex Research, 25*, 123-151.

Fisher, T. D., Davis, C. M., Yarber, W. L., & Davis, S. L. (2010). *Handbook of Sexuality-Related Measures*. New York: Routledge.

Fox, S., & Schwartz, D. (2002). Social desirability and controllability in computerized and pencil-and-paper personality questionnaires. *Computers in Human Behavior, 18*(4), 389-410. doi:10.1016/S0747-5632(01)00057-7

Gillmore, M., Leigh, B., Hoppe, M., & Morrison, D. (2010). Comparison of Daily and Retrospective Reports of Vaginal Sex in Heterosexual Men and Women. *Journal of Sex Research, 47*(4), 279-284. doi:10.1016/S0191-8869(96)00137-7

Gosling, S., Vazire, S., Srivastava, S., & John, O. (2004). Should we trust web-based studies. *American Psychologist, 59*(2), 93-104. doi:10.1037/0003-066X.59.2.93

Hamilton, D., & Morris, M. (2010). Consistency of Self-Reported Sexual Behavior in Surveys. *Archives of Sexual Behavior*, 1-19. doi:10.1007/s10508-009-9505-7

Hardré, P. L., Crowson, H. M., &Xie, K. (2012). Examining contexts-of use for web-based and paper-based questionnaires. *Educational and Psychological Measurement*, 72, 1015-1038. doi: 10.1177/0013164412451977

Hines, D., Douglas, E., & Mahmood, S. (2010). The effects of survey administration on disclosure rates to sensitive items among men: A comparison of an internet panel sample with a RDD telephone sample. *Computers in Human Behavior*. doi:10.1016/j.chb.2010.04.006

Kays, K., Gathercoal, K., & Buhrom, W. (2012). Does survey format influence self-disclosure on sensitive question items? *Computers in Human Behavior, 28,* 251-256. doi:10.1016/j.chb.2011.09.007

Knapp, H., & Kirk, S. (2003). Using pencil and paper, Internet and touch-tone phones for self-administered surveys: does methodology matter? *Computers in Human Behavior, 19*(1), 117-134. doi:10.1016/S0747-5632(02)00008-0

Langhaug, L., Sherr, L., & Cowan, F. (2010). How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. *Tropical Medicine & International Health, 15*(3), 362-381. doi:10.1111/j.1365-3156.2009.02464.x

Lautenschlager, G., & Flaherty, V. (1990). Computer administration of questions: more desirable or more social desirability? *Journal of Applied Psychology, 75*(3), 310-314. doi: 10.1037/0021-9010.75.3.310

Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., Vehovar, V., & Berzelak, N. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research, 50*(1), 79.

McAuliffe, T., DiFranceisco, W., & Reed, B. (2007). Effects of question format and collection mode on the accuracy of retrospective surveys of health risk behavior: A comparison with daily sexual activity diaries. *Health Psychology, 26*(1), 60-67. doi:10.1037/0278-6133.26.1.60

Meston, C., Heiman, J., Trapnell, P., & Paulhus, D. (1998). Socially desirable responding and sexuality self-reports. *Journal of Sex Research, 35*(2), 148-157. doi:10.1080/00224499809551928

Morrison-Beedy, D., Carey, M., & Tu, X. (2006). Accuracy of audio computer-assisted self-interviewing (ACASI) and self-administered questionnaires for the assessment of sexual behavior. *AIDS and Behavior, 10*(5), 541-552. doi:10.1007/s10461-006-9081-y

Ollesch, H., Heineken, E., & Schulte, F.P. (2006). Physical or virtual presence of the experimenter: Psychological online-experiments in different settings.*International Journal of Internet Science,* 71-81.

Parks, K., Pardi, A., & Bradizza, C. (2006). Collecting data on alcohol use and alcohol-related victimization: A comparison of telephone and Web-based survey methods. *Journal of Studies on Alcohol and Drugs*, *67*(2), 318.

Paulhus, D. (1984). Two-component models of socially desirable responding. *Journal of personality and social psychology, 46*(3), 598-609. doi:10.1037//0022-3514.46.3.598

Paulhus, D.L. (1991). Balanced Inventory of Desirable Responding (BIDR). In Robinson, J.P., & Shaver, P.R. (Eds.), *Measures of personality and social psychological attitudes* (pp. 37-41). San Diego, CA: Academic Press.

Peterson, Z. D., McCallum, E., Janssen, E., Heiman, J. R., Goodrich, D., Thigpen, J., & Fortenberry, J. D. (2010, November). *Negative affect as a mediator in the relationship between child victimization and adult sexual perpetration: A study of men from urban STD clinics.* Paper presented at the annual meeting of the Association for Behavioral and Cognitive Therapies, San Francisco, CA.

Reddy, M., Fleming, M., Howells, N., Rabenhorst, M., Casselman, R., & Rosenbaum, A. (2006). Effects of method on participants and disclosure rates in research on sensitive topics. *Violence and Victims, 21*(4), 499-506. doi:10.1891/vivi.21.4.499

Richman, W., Weisband, S., Kiesler, S., & Drasgow, F. (1999). A meta-analytic study of social desirability response distortion in computer-administered and traditional questionnaires and interviews. *Journal of Applied Psychology, 84,* 754–775. doi: 10.1037/0021-9010.84.5.754

Rosenbaum, A., Rabenhorst, M., Reddy, M., Fleming, M., & Howells, N. (2006). A comparison of methods for collecting self-report data on sensitive topics. *Violence and Victims, 21*(4), 461-471. doi:10.1891/vivi.21.4.461

Ross, M., Daneback, K., Månsson, S., Tikkanen, R., & Cooper, A. (2002). Characteristics of men and women who complete or exit from an on-line Internet sexuality questionnaire: A study of instrument dropout biases. *Journal of Sex Research, 40*(4), 396-402. doi:10.1080/00224490209552205

Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research, 8,* 3-15. doi: 10.1177/096228029900800102

Schroder, K., Carey, M., & Vanable, P. (2003). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Annals of Behavioral Medicine, 26*(2), 104-123. doi:10.1207/S15324796ABM2602_03

Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: evidence for convergent and discriminant validity. Journal of Personality and Social Psychology, 60, 870–883.

Smith, T. (1992). A methodological analysis of the sexual behavior questions on the General Social Surveys. *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM, 8*, 309-309.

IBM SPSS Statistics for Windows 20.0 [Computer software]. (2011). Armonk, NY: IBM Corp.

Strang, E. T., Peterson, Z. D., Hill, Y. N., & Heiman, J. R. (in press). Discrepant responding across self-report measures of men's coercive and aggressive sexual strategies. *Journal of Sex Research*. doi: 10.1080/00224499.2011.646393

Struckman-Johnson, C., Struckman-Johnson, D., & Anderson, P. B. (2003). Tactics of sexual coercion: When men and women won't take no for an answer. *Journal of Sex Research, 40,* 76-86.

Testa, M., Livingston, J., & VanZile Tamsen, C. (2005). The impact of questionnaire administration mode on response rate and reporting of consensual and nonconsensual sexual behavior. *Psychology of Women Quarterly*, *29*(4), 345-352. doi:10.1111/j.1471-6402.2005.00234.x

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859. doi:10.1086/297751

Turchik, J.A., & Garske, j.P. (2009). Measurement of sexual risk taking among college students. *Archives of sexual behavior,* 38, 936-948.

Vannier, S. A., & O'Sullivan, L. F. (2008). The feasibility and acceptability of handheld computers in a prospective diary study of adolescent sexual behaviour. *Canadian Journal of Human Sexuality,* 17(4), 183-192.

Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based pencil-and-paper and internet data collection methods. *Psychological Methods, 18,* 53-70. doi: 10.1037/a0031607

Weinhardt, L., Forsyth, A., Carey, M., Jaworski, B., & Durant, L. (1998). 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*, *27*(2), 155-180. doi:10.1023/A:1018682530519

Wood, E., Nosko, A., Desmarais, S., Ross, C., & Irvine, C. (2006). Online and traditional pencil-and-paper survey administration: Examining experimenter presence, sensitive material and long surveys. *Canadian Journal of Human Sexuality*, 15(3/4), 147-155.