

Violence Against Women

<http://vaw.sagepub.com/>

The Effects of Survey Question Wording on Rape Estimates: Evidence From a Quasi-Experimental Design

Bonnie S. Fisher

Violence Against Women 2009 15: 133

DOI: 10.1177/1077801208329391

The online version of this article can be found at:

<http://vaw.sagepub.com/content/15/2/133>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Violence Against Women* can be found at:

Email Alerts: <http://vaw.sagepub.com/cgi/alerts>

Subscriptions: <http://vaw.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://vaw.sagepub.com/content/15/2/133.refs.html>

>> [Version of Record](#) - Jan 6, 2009

[What is This?](#)

The Effects of Survey Question Wording on Rape Estimates

Evidence From a Quasi-Experimental Design

Bonnie S. Fisher

University of Cincinnati

The measurement of rape is among the leading methodological issues in the violence against women field. Methodological discussion continues to focus on decreasing measurement errors and improving the accuracy of rape estimates. The current study used a quasi-experimental design to examine the effect of survey question wording on estimates of completed and attempted rape and verbal threats of rape. Specifically, the study statistically compares self-reported rape estimates from two nationally representative studies of college women's sexual victimization experiences, the National College Women Sexual Victimization study and the National Violence Against College Women study. Results show significant differences between the two sets of rape estimates, with National Violence Against College Women study rape estimates ranging from 4.4% to 10.4% lower than the National College Women Sexual Victimization study rape estimates. Implications for future methodological research are discussed.

Keywords: *quasi-experiment; rape; rape estimates; survey question*

The measurement of rape is one of the key methodological issues in the study of violence against women (Fisher & Cullen, 2000). Obtaining accurate rape estimates has garnered worldwide interest by scholars and policy makers from a range of disciplines (Hamby & Koss, 2003; Jaquier, Fisher, & Killias, 2006; Johnson, Ollus, & Nevala, 2008; Kilpatrick, 2004; Koss et al., 2007; Koss, Gidycz, & Wisniewski, 1987; Lynch, 1996b; Saltzman, 2000a, 2000b, 2004; Walby & Myhill, 2001).

Author's Note: This project was supported under Award 95-WT-NX-0001 for the National Institute of Justice and Award 97-MU-MU-0011 from the Bureau of Justice Statistics, U.S. Department of Justice. Points of view in this article are those of the author and do not necessarily represent the official position of the U.S. Department of Justice. The author thanks Professors Mary Koss and Dean Kilpatrick for their continual interest in rape measurement issues, especially Dean for his support by continuously referencing the results reported in this article. The author also is thankful to Professor John Sloan for his comments on earlier drafts. The author is responsible for any mistakes.

Issues surrounding the measurement of rape were initially presented by feminists as they brought rape to the forefront of legal reforms and helped to change its definition (see Bachman & Paternoster, 1993). This impetus resulted in several interwoven arguments that ultimately changed how rape was measured, especially in self-report surveys (Fisher & Cullen, 2000; Koss et al., 2007). During the 1980s, Koss and her colleagues (Koss et al., 1987; Koss & Gidycz, 1985; Koss & Oros, 1982) developed the Sexual Experiences Survey and administered it to a nationally representative sample of college women. Their results, especially concerning the accuracy of rape estimates, ignited what has become a long-standing debate between feminist scholars and their critics around the validity of the Sexual Experiences Survey in measuring rape and raised serious questions about whether the extent of women's rape is a "true" social and public health problem or a misguided "social construction of reality" (Gilbert, 1992, 1994, 1995, 1997; Koss, 1992, 1996).

The results of Koss et al.'s (1987) study also fueled scholarly discussion of whether the two primary sources of rape estimates at the time, "official" statistics from the Uniform Crime Reports and those from the nationwide National Crime Survey (NCS), substantially underestimated the true incidence of rape. Critics persuasively contended that the NCS was poorly designed to elicit reports of rape (and other forms of sexual violence) from respondents who, in fact, had experienced these victimizations (Fisher & Cullen, 2000). The crux of their criticism was simple: The NCS did not include specific questions directly asking about rape experiences (see Bachman & Taylor, 1994; Eigenberg, 1990; Koss, 1993a, 1993b; Lynch 1996a, 1996b). Lacking, however, in their well-formulated and convincing criticisms was empirical evidence generated from a well-designed comparative study of how rape is measured and estimated.

Debates surrounding the validity of how rape is measured and its estimates—or perhaps because of them—resulted in several methodological advancements. First, studies abound that further elaborate on the discussion of the effects that different research designs, operationalizations of rape, and wording of survey questions have on rape estimates. Each study provides plausible methodological explanations that are critical for better understanding why such widely diverging rape estimates exist (see Bachman, 2000; Cantor & Lynch, 2005; Hamby & Koss, 2003; Lynch, 1996a, 1996b; Jaquier et al., 2006; Schwartz, 2000). Second, substantial revisions in both the NCS and the Sexual Experiences Survey have occurred. The redesigned NCS—now called the National Crime Victimization Survey (NCVS)—was first administered in 1992. The NCVS addressed a considerable number of the methodological shortcomings inherent in its previous form, including a broader definition of rape and the use of "short cue" and additional victimization screen questions to improve self-reporting of rape and sexual assault incidents (Bachman & Saltzman, 1995; Bachman & Taylor, 1994; Cantor & Lynch, 2005; Lauritsen, 2005). Using methodological lessons learned from previous studies and from the NCVS, the Sexual Experiences Survey was also recently redesigned to address its widely known

weaknesses, including the vocabulary used and ambiguous assessment of consent, and to build on its original strengths, such as behavioral specification of acts and tactics (Koss et al., 2007).

Although methodological innovations have occurred, very few published, national-level empirical studies completed to date have advanced our understanding of how diverse methods affect rape estimates. Two studies, Bachman (2000) and Jaquier et al. (2006), used post hoc comparisons between different national surveys that measured rape. Each of the researchers performed considerable post hoc manipulation of the survey designs to address differences in their attempts to equate the respective survey designs to make meaningful comparisons.

Bachman (2000) statistically compared annual rape estimates from two national-level studies, the NCVS and the National Violence Against Women Survey (NVAWS). For her comparison, she constructed the NCVS “as comparable as possible” to the NVAWS (Bachman, 2000, p. 839). These two studies were not originally designed to be compared, so Bachman could not make comparable several methodological attributes that other researchers have argued are reasons for diverging rape estimates (see Fisher & Cullen, 2000). For example, the NCVS uses a two-stage measurement process with victimization screen questions and incident reports (see Bachman & Saltzman, 1995; Bachman & Taylor, 1994). Responses to questions in the incident report are used to classify which category of crime, if any, the respondent experienced. The NVAWS used a one-stage measurement process that incorporated behaviorally specific questions to determine if victimization occurred (see Tjaden & Thoennes, 1998). Despite the previously noted changes in the NCVS, Bachman (2000) concluded that “the NVAWS has a greater likelihood of capturing incidents of intimate-perpetrated rape . . . compared to the NCVS” (p. 860). Her conclusion concerning the NCVS supports critics who argued the NCS, the precursor to the NCVS, underestimated the extent of rape.

Jaquier and her colleagues (2006) compared rape estimates from the NVAWS and the International Violence Against Women Survey administered in Switzerland after they equated, post hoc, the surveys’ differences to make their estimates of completed and attempted rape as close as possible for comparison purposes. The results illuminated the effects of not addressing survey design differences, in particular how rape is operationalized, when doing comparative rape estimate analyses that involve two different surveys.

Bachman’s (2000) and Jaquier et al.’s (2006) research is innovative, advances understanding of the importance of performing comparative rape estimate research, and shows differences in rape estimates occurring because of methodological differences. However, the studies are characterized by similar weaknesses: Each study relies on post hoc manipulation to equate its data sets, and each was not originally designed to be a comparative study of rape estimates (or other types of sexual victimization).

The current study addresses the gaps noted in the discussion above. First, unlike previous methodological studies that identify measurement shortcomings that might

influence rape estimates, the current study empirically examines methodological issues that affect rape estimates. Second, contrasted to the post hoc research designs used by Bachman (2000) and Jaquier and her colleagues (2006) to compare rape estimates generated from different studies, the current study was explicitly designed to test if survey design differences, specifically survey question wording used to operationalize rape, had any effect on rape estimates. Thus, the current study utilizes a quasi-experimental research design to compare self-reported rape estimates from two nationally representative samples of college women, the National College Women Sexual Victimization (NCWSV) study and the National Violence Against College Women (NVACW) study. Estimates of completed and attempted rape and verbal threats of rape based on different question wording found in the two studies are presented for the purpose of informing ongoing methodological discussions surrounding the measurement of rape (and other forms of sexual violence). Implications of the results for future research are then discussed.

Research Design Attributes of the NCWSV and NVACW Studies: The Development of a Quasi-Experimental Research Design

The NCWSV and NVACW studies jointly provided a unique opportunity to empirically compare rape estimates generated from a quasi-experimental research design. The comparison also allowed the taking into account of several methodological issues (e.g., sampling design, reference period) that scholars had speculated influenced diverging estimates of rape while simultaneously examining the effects operationalization may have on rape estimates.

Table 1 outlines the research design attributes for the NCWSV and the NVACW studies. Most of the attributes of the research design were identical across the two studies, whereas a few were intentionally designed to differ across the two studies (see Fisher & Cullen, 1998, 1999, 2000).

Sampling Design

The sampling design employed in the NCWSV and the NVACW studies were identical (see Table 1, rows 2-5). The population included all 4-year and 2-year institutions of higher education that had a total enrollment of at least 1,000 students. The college student sampling frame for both studies was provided by American Student List Company, LLC, for the same academic year enrollment period.

The studies employed an identical two-stage probability sampling design. First, a total of 233 institutions of higher education were selected from 12 strata (three categories of location: urban, suburban, and small town/urban; and four categories of total student enrollment: 1,000 to 4,999, 5,000 to 9,999, 10,000 to 19,999, and 20,000 or

Table 1
Overview Comparison of the National College Women Sexual Victimization Study and National Violence Against College Women Study

Research Design Attribute	National College Women Sexual Victimization Study	National Violence Against College Women Study
Sampling design		
Sampling frame	<i>Four-year and 2-year institutions of higher education in the United States that had a total student enrollment of at least 1,000 students</i>	<i>Four-year and 2-year institutions of higher education in the United States that had a total student enrollment of at least 1,000 students</i>
Sampling design	<i>Two stages: (a) stratified institutions of higher education by total student enrollment and location of school and (b) randomly selected women enrolled in selected institutions of higher education</i>	<i>Two stages: (a) stratified institutions of higher education by total student enrollment and location of school and (b) randomly selected women enrolled in selected institutions of higher education</i>
Sample size: schools	Two hundred thirty-three institutions of higher education total: 194 4-year institutions of higher education and 39 2-year institutions of higher education	Two hundred thirty-three institutions of higher education total: 191 4-year institutions of higher education and 42 2-year institutions of higher education
Sample size: female college students	4,446	4,432
Context of study in the cover letter		
Title of survey	The Extent and Nature of Sexual Victimization of College Women	Victimization Among College Women
Description of study context	Unwanted sexual experiences that women may experience during college ^a	Criminal victimization that women may experience during college
Interviewing		
Survey firm	<i>Schulman, Ronca, and Bucuvalas, Inc.</i>	<i>Schulman, Ronca, and Bucuvalas, Inc.</i>
Interviewers	<i>Professionally trained women</i>	<i>Professionally trained women</i>
CATI	<i>Yes</i>	<i>Yes</i>
Average interview time	25.9 minutes	12.7 minutes
Field period	February 21, 1997, to May 5, 1997	March 27, 1997, to May 14, 1997
Response rate (%) ^b	85.6	91.6
Introduction to survey		
Wording used in introduction to telephone interview	<i>As you may recall, the purpose of the study is to better understand the extent and nature of criminal victimization among college women. Regardless of whether or not you have ever personally been victimized, your answers will help us to understand and deal with the problem of victimizations at your campus and nationally.</i>	<i>As you may recall, the purpose of the study is to better understand the extent and nature of criminal victimization among college women. Regardless of whether or not you have ever personally been victimized, your answers will help us to understand and deal with the problem of victimizations at your campus and nationally.</i>

(continued)

Table 1 (continued)

Research Design Attribute	National College Women Sexual Victimization Study	National Violence Against College Women Study
Definition of rape		
Completed rape	Unwanted completed penetration by physical force or the threat of physical force. Penetration includes penile-vaginal, mouth on your genitals, mouth on someone else's genitals, penile-anal, digital-vaginal, digital-anal, object-vaginal, and object-anal. ^c	Forced sexual intercourse, including both psychological coercion as well as physical force. Forced sexual intercourse means vaginal, anal, or oral penetration by the offender(s). This category also includes incidents where the penetration is from a foreign object such as a bottle. ^{d,e}
Attempted rape	Unwanted attempted penetration by force or the threat of force.	Attempted forced sexual intercourse, including the use of both psychological coercion as well as physical force.
Threat of rape	Threat of unwanted penetration with force and threat of force.	Threatened forced sexual intercourse, including both psychological coercion as well as physical force.
Operationalizing rape		
Measurement approach	<i>Two stages: (a) screen questions and (b) detailed incident report</i>	<i>Two stages: (a) screen questions and (b) detailed incident report</i>
Screen questions cueing strategy	Behaviorally specific	Short cue, direct, broad net
Incident questions	Multiple questions concerning (a) type of completed, attempted, and threatened penetration and (2) physical force used or threatened with physical force	Multiple questions concerning (a) what actually happened, how victim was attacked, how offender tried to attack, and how offender threatened and (b) clarification of rape, attempted rape, or unwanted sexual contact with force
Reference period	<i>Since school began in fall 1996</i>	<i>Since school began in fall 1996</i>
Victimization categorization criterion	<i>Hierarchical scoring procedure</i>	<i>Hierarchical scoring procedure</i>

Note: Italics font denotes research attributes that were identical. CATI = computer-aided telephone interviewing.

a. Examples, such as stalking, sexual assault, and sexual harassment were provided.

b. For both samples, we summed the total number of respondents completing the survey and the total number of respondents who were screened out and divided this figure by the total number of potential respondents contacted by Schulman, Ronca, and Bucuvalas, Inc.

c. This definition for penetration is used by the National College Women Sexual Victimization study for attempted and threatened rape.

d. This is the definition used in the National Crime Victimization Survey (see Bureau of Justice Statistics, 2000, p. 175).

e. This definition for forced sexual intercourse is used by the National Violence Against College Women study for attempted rape and threat of rape.

more). Institutions within each stratum were then selected using a probability-proportionate-to-size of female enrollment. Second, within each selected institution, female students were then randomly selected. For each stratum, the sample size for institutions of higher education and students was determined based on a standard acceptable margin of error. The total sample sizes for the NCWSV study and the NVACW study were large: 4,446 and 4,432 college women, respectively.

Context of the Study in a Cover Letter

Sample members were informed about the context of the study in a cover letter sent via the U.S. Postal Service. Each sample member was sent this letter at her current school address approximately 2 weeks prior to being called on the telephone by an interviewer (see Table 1, rows 7-8). In the cover letter, the title of the survey and the wording used to describe the context of the study were somewhat different between the two studies. The NCWSV study referred to an “unwanted sexual experience,” whereas the NVACW study referred to “criminal victimizations” in the introductory paragraph to describe the context of the study. Other than these two wording differences, the content of the cover letters was identical. Each letter provided such information as identifying the sponsor of the study, indicating that participation was voluntary, and providing information about whom to contact if the sample member had questions regarding the legitimacy of the study and/or wanted a copy of the results (e.g., an 800 number and an e-mail address were provided).

Interviewing

Female interviewers, who were hired and professionally trained by Schulman, Ronca, and Bucuvalas, Inc., administered the respective survey using a computer-aided telephone interviewing (CATI) system (see Table 1, rows 10-15). The two field periods were not identical; there was, however, considerable overlap between them. The NCWSV study field period began February 21, 1997, and ended May 5, 1997. The NVACW study field period started approximately 1 month later, on March 27, 1997, and lasted 9 days longer, ending on May 14, 1997. Administration of the NCWSV survey lasted twice as long as administration of the NVACW survey: 26 minutes compared to 13 minutes. It is instructive to note that the surveys used identical wording in the introduction for the telephone interview. Interviewers read this introduction to all the respondents, including those who had and those who had not recalled receiving the cover letter. Due to the telephone administration of the survey, respondents provided informed consent by verbally agreeing to participate in the survey. At any time, they could terminate the interview, reschedule to complete the survey, or refuse to answer any question. Participation in the study was voluntary. Participants' responses were confidential, and their anonymity was guaranteed. The author's institutional review board approved each study's protocol.

Introduction to the Survey

In addition to the cover letter, the introduction read to the respondents set the context in which information about victimizations was solicited (see Table 1, row 16). After assessing if the respondent had received the cover letter, if the respondent agreed to participate in the study, and if the respondent was eligible to participate, the interviewers then read the identical introduction to respondents in the NCWSV study and the NVACW study.

Definition of Rape

Each study measured completed, attempted, and threatened rape (see Table 1, rows 18-20). Both studies included in their definitions of rape forced vaginal, anal, or oral penetration by the perpetrator(s), which can also include experiences where the penetration is from a foreign object. The NCWSV study definition of rape explicitly referred to physical force and the threat of physical force. The NVACW study definition of rape also referred to physical force but also incorporated “psychological coercion.” Koss (1996) noted that this term “is probably meant to refer to verbal threats of bodily harm or rape, which are crimes” (p. 60). She further noted that it might also suggest situations involving verbal strategies to coerce sexual intercourse (e.g., continual nagging), which are undesirable but are not crimes.

Operationalizing Rape

There were similarities and differences in how rape was operationalized in the two studies (see Table 1, rows 22-26). Identical to the NCVS measurement strategy, both studies employed a two-stage measurement process: (a) victimization screen questions and (b) incident reports. Both studies asked a series of “screen questions” to determine if a respondent experienced an act “since school began in the Fall of 1996” that may be defined as a victimization. If the respondent answered “yes,” then for each number of times that experience happened, the respondent is asked by the interviewer to complete an “incident report.” The report contains detailed questions about the nature of the events that occurred in the incident. The incident report was used to classify the type of victimization that took place; that is, responses to questions in the incident report, not the screen questions, were used to categorize the type of victimization, if any, that occurred.

There were two differences in how rape was operationalized by the NCWSV and the NVACW: (a) the number and wording of the screen questions and (b) the wording of the incident-level questions used to determine the type of incident. The NCWSV substantially modified the NCVS format, most notably to include a range of 12 behaviorally specific sexual victimization screen questions (including 1 for stalking; see Fisher & Cullen, 2000). A behaviorally specific question is one that does not ask simply if a respondent “had been raped,” but rather describes an incident in graphic

language that covers the elements of a criminal offense (e.g., someone “made you have sexual intercourse by using force or threatening to harm you . . . by intercourse I mean putting a penis in your vagina”) (see Fisher et al., 2000, Table 1). Each rape screen question asked the respondent about a different form of penetration wherein force or the threat of harm was used; a statement then followed that defined the type of penetration (e.g., anal sex is defined as “putting a penis in your anus or rectum”). The other screen questions provided examples of the types of behavior about which we were asking respondents. Noteworthy is that the rape measurement work of Koss et al. (1987); Kilpatrick, Edmunds, and Seymour (1992); and Tjaden and Thoennes (1998) was influential in the development of these screen questions.

In contrast, the NVACW study used a format that was as closely aligned as possible with that of the NCVS. All seven of the individual-level screen questions used in the NVACW came directly from the NCVS, as did the incident-level questions used to determine what type of violent victimization the respondent experienced. In the NVACW, the NCVS screen question specifically asked whether a respondent “has been forced or coerced to engage in unwanted sexual activity,” as were questions about having something stolen, having something attempted to be stolen, being attacked, and being threatened (U.S. Department of Justice, 2000, p. 129).

For each study, within an incident, the same categorization criterion was used: a hierarchical scoring procedure. The incident was categorized using the most serious type of victimization that the respondent reported had occurred. For example, if in one incident, two victimizations took place—say a completed rape and a simple assault or sexual coercion—the incident was categorized as a completed rape.

The two studies also differed in how rape was operationalized within an incident report. The NCWSV specifically asked about what acts were completed, attempted, and/or threatened. For each of these three forms of behavior, respondents were asked multiple response questions to identify which type(s) of penetration they had experienced. After these questions, we asked two questions about whether actual or threatened physical force was used. In contrast, in the NVACW, if a respondent indicated in any of the “what happened” questions (e.g., what actually happened, how did the offender try to attack you, or how were you threatened) that an unwanted sexual contact with force occurred, she was then asked if she meant forced or coerced sexual intercourse, including attempts. If she said “yes” to this question, the incident was categorized as a rape. Also, if the respondent indicated the offender hit her, knocked her down, or attacked her and that among her injuries was a rape or an attempted rape, she was then asked if she meant forced or coerced sexual intercourse, including attempts. The incident was then categorized according to one of three types of rape (see Fisher & Cullen, 2000).

Every effort was made to ensure that aside from using different screen and incident report questions, the methodology used in the NCWSV study and the NVACW study was identical. To date, this is the strongest research design employed to examine how these two differences affected rape estimates.

Table 2
Estimates From the National College Women Sexual Victimization Study and the National Violence Against College Women Survey

Type of Victimization	National College Women Sexual Victimization Study		National Violence Against College Women Study		<i>p</i> Value ^a
	Percentage of Victims, 95% Confidence Interval (<i>n</i>)	Rate of Victimization per 1,000 Female Students (<i>n</i>)	Percentage of Victims, 95% Confidence Interval (<i>n</i>)	Rate of Victimization per 1,000 Female Students (<i>n</i>)	
Rape					
Completed rape	1.66, 1.29–2.04 (74)	19.34 (86)	.16, .04–.27 (7)	2.0 (9)	.0000
Attempted rape	1.10, .80–1.41 (49)	15.97 (71)	.18, .06–.30 (8)	1.8 (8)	.0000
Verbal threat of rape	.31, .15–.48 (14)	9.45 (42)	.07, .01–.14 (3)	.7 (3)	.0095

a. The estimated percentage of victims from the two studies was statistically compared using a two-sample test of proportions calculated in STATA 8.2.

Estimates of Rape From the NCWSV Study and the NVACW Study

Tjaden and Thoennes (1998) and Bachman (2000) reported that studies using behaviorally specific questions generally find higher levels of sexual victimization than those reported by the NCVS. Examining Table 2, it is clear that the estimates for completed rape, attempted rape, and threats of rape from the NVACW study are statistically significantly lower than the estimates from the NCWSV.

The percentage of the sample that reported experiencing a completed, attempted, or threatened rape in the NVACW was significantly smaller than in the NCWSV. For completed rape, 0.16% of the NVACW sample reported having experienced a completed rape, compared to 1.66% of the women in the NCWSV. For attempted rape, 0.18% of the NVACW sample reported they had experienced such behavior, compared to 1.10% of the NCWSV sample. A similar pattern was found for threatened rape: 0.07% of the NVACW sample reported a threatened rape, compared to 0.31% of the sample in the NCWSV. Noteworthy also is the magnitude of the differences between the rape estimates from the NVACW study and NCWSV study. The NVACW rape estimates are significantly smaller than those from the NCWSV study: 10.4 times smaller for completed rape, 6.1 times smaller for attempted rape, and 4.4 times smaller for threatened rape.

What accounts for these differences? Given the similarities between the two studies, it would appear that the most likely difference is that the NCWSV used a range of behaviorally specific screen questions. Compared to the NCVS screen questions

employed by the NVACW, the use of graphically worded screen questions in the NCWSV likely prompted more women who had experienced a sexual victimization to report this fact to the interviewer. Although not all of those answering “yes” to a rape screen question were subsequently classified as rape victims based on their responses in the incident report (see Fisher & Cullen, 2000), it appears that behaviorally specific screen questions are more successful in prompting women who have in fact been sexually victimized to answer in such a way that they are then “skipped into” the incident report by interviewers. These results support those reported by Tjaden and Thoennes (1998) and Bachman (2000). When these results, including the current study’s empirical evidence generated from a quasi-experimental research design, are examined and coupled with the arguments articulated by Koss and her colleagues (1987; Koss, 1992, 1993a, 1993b, 1996; Koss et al., 2007) and Kilpatrick and his colleagues (1992; Kilpatrick, 2004), it seems likely the NCVS underestimates the “true” incidence of rape in the United States.

Conclusion and Implications for Future Methodological Research

Measuring rape (as well as other forms of sexual victimization) is a complicated and, to a degree, imperfect enterprise (Fisher & Cullen, 2000). According to Smith (1987, p. 185), it is the “biggest methodological challenge in survey research.” The challenges are especially daunting when one is attempting to discern when, in an intimate encounter, a sexual advance crosses the line from imprudence to criminal behavior. But the salience of the methodology of measuring rape is intensified even further because the “findings” are integral to the ongoing debate between feminist and conservative scholars about whether women’s rape is a true social problem or a misguided social construction of reality, and about policy responses. No single study, including the comparison between the NCWSV and NVACW, can fully resolve this debate; this study, however, can inform current discussion and provide guidance for improving the measurement of rape (and by extension, a full range of violence against women).

The results have at least four important implications for the measurement of rape. First, the use of behaviorally specific questions cannot be overemphasized, not necessarily because they produce larger estimates of rape but because they use words and phrases that describe to the respondent exactly what behavior is being measured. Using behaviorally specific screen questions appears to cue more women to recall what they experienced. The use of behaviorally specific questions is not a panacea for addressing measurement error associated with estimating rape (and other forms of victimization), but it is a step forward in understanding how question wording affects self-report survey responses (see Fisher & Cullen, 2000; Walby & Myhill, 2001).

Second, drawing on the strength of the NCVS, the two-stage measurement process—screen questions and incident reports—appears to be a promising way to address the measurement error typically associated with a single-stage measurement process, although it still needs further rigorous testing (Fisher & Cullen, 2000). There is, however, some evidence to suggest that single measurement may introduce measurement error. For example, in the NCWSV, of the 325 incidents that screened in on the rape screen questions, 21 of them could not ultimately be classified because the respondent could not recall enough detail in the incident report; 59 were then classified as “undetermined” because the respondent refused to give answer questions or answered “don’t know” to one or more questions in the incident report that would have allowed the incident to be categorized as a rape; 155 were classified as a type of sexual victimization other than rape; and 90 were classified as rape (completed, attempted, or threatened). The other 109 rape incidents screened in from the other sexual victimization screen questions (see Fisher & Cullen, 2000). These results provide us with some understanding of how using only behaviorally specific questions would fail to count women whose recall is prompted by other types of screen questions. To date, we have only a preliminary understanding of what sources of measurement error the use of incident reports might introduce; further research is needed to examine this. One avenue of research might consider how the use of structured qualitative questions that allow the respondent to “tell her own story” help the researcher to identify and understand the sources of measurement error in both the behaviorally specific questions and the incident report (Hamby & Koss, 2003; Walby & Myhill, 2001).

Third, there is one other factor that might have contributed to significant differences between the NCWSV study and NVACW study estimates: the “context” of the two surveys. Recall from Table 1 (see rows 7-8) that it is plausible that NCWSV respondents were sensitized to report a broad range of sexual victimization incidents they experienced, whereas NVACW respondents limited their reports to incidents they defined as criminal. If so, the contextual difference would mean the NVACW study was measuring a much narrower domain of sexual victimization than was the NCWSV study (see Saltzman, 2004). One caution to this line of reasoning is that nearly half of the completed rape victims, when asked if they considered the incident a “rape,” said “yes.” Even when the count of completed rape is limited to this group, the incidence of rape victims is still several times greater in the NCWSV than in the NVACW study. The impact of the survey question context on respondents’ responses to sexual victimization questions remains an area that warrants further methodological examination.

Fourth, to further the understanding of rape and a full range of violence, comparative research employing experimental and quasi-experimental designs should not be overlooked. The strength of these types of research designs will allow researchers to manipulate sources of possible measurement error and estimate their effects on estimates of different types of victimization. At present, this research remains in its beginning stages, but the current study shows that this type of research can be done.

The results from the current study are one step toward addressing issues surrounding the measurement of rape. Future researchers are encouraged to build from the current study and take steps to more rigorously test which other aspects of measurement, if any, affect estimates of rape and other forms of violence against women. Their work can be used to inform the production of new knowledge and development of theories and social and health policies with a better understanding of measurement of these estimates and, it is hoped, more accurate estimates of the extent of rape and other forms of violence against women.

References

- Bachman, R. (2000). A comparison of annual incidence rates and contextual characteristics of intimate partner violence against women from the National Crime Victimization Survey (NCVS) and the National Violence Against Women Survey (NVAWS). *Violence Against Women, 6*, 839-867.
- Bachman, R., & Paternoster, R. (1993). Changes in the adjudication of rape: 1970-1991. A contemporary look at the effects of rape law reform: How far have we really come? *Journal of Criminal Law and Criminology, 84*, 554-574.
- Bachman, R., & Saltzman, L. E. (1995). *Violence against women: Estimates from the redesigned survey*. Washington DC: U.S. Department of Justice.
- Bachman, R., & Taylor, B. M. (1994). The measurement of family violence and rape by the redesigned National Crime Victimization Survey. *Justice Quarterly, 11*, 499-512.
- Bureau of Justice Statistics. (2000). *Criminal victimization in the United States, 1995: A National Crime Victimization Survey report*. Washington, DC: U.S. Department of Justice.
- Cantor, D., & Lynch, J. P. (2005). Exploring the effects of changes in design on the analytical uses of the NCVS data. *Journal of Quantitative Criminology, 21*, 293-319.
- Eigenberg, H. M. (1990). The National Crime Survey and rape: The case of the missing question. *Criminology, 7*, 655-671.
- Fisher, B. S., & Cullen, F. T. (1998, August). *The sexual victimization of college women: Findings from a national-level study*. Draft of final report submitted to the National Institute of Justice.
- Fisher, B. S., & Cullen, F. T. (1999, March). *The victimization of college women: Findings from a national-level study*. Draft of final report submitted to the Bureau of Justice Statistics.
- Fisher, B. S., & Cullen, F. T. (2000). Measuring the sexual victimization of women: Evolution, current controversies, and future research. In D. Duffee (Ed.), *Measurement and analysis of crime and justice: Criminal justice 2000* (Vol. 4, pp. 317-390). Washington DC: U.S. Department of Justice.
- Gilbert, N. (1992, May/June). Realities and mythologies of rape. *Society, pp.* 4-10.
- Gilbert, N. (1994, March/April). Miscounting social ills. *Society, pp.* 18-26.
- Gilbert, N. (1995). *Was it rape? An examination of sexual assault statistics*. Menlo, CA: Henry J. Kaiser Family Foundation.
- Gilbert, N. (1997). Advocacy research and social policy. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 22, pp. 101-148). Chicago: University of Chicago Press.
- Hamby, S. L., & Koss, M. P. (2003). Shades of gray: A qualitative study of terms used in the measurement of sexual victimization. *Psychology of Women Quarterly, 27*, 243-255.
- Jaquier, V., Fisher, B. S., & Killias, M. (2006). Cross-national survey designs: Equating the National Violence Against Women Survey and the Swiss International Violence Against Women Survey. *Journal of Contemporary Criminal Justice, 22*, 90-112.
- Johnson, H., Ollus, N., & Nevala, S. (2008). *Violence against women: An international perspective*. New York: Springer.

- Kilpatrick, D. G. (2004). What is violence against women? Defining and measuring the problem. *Journal of Interpersonal Violence, 19*, 1209-1234.
- Kilpatrick, D. G., Edmunds, C. N., & Seymour, A. K. (1992). *Rape in America: A report to the nation*. Arlington, VA: National Victim Center.
- Koss, M. P. (1992). The underdetection of rape: Methodological choices influence incidence estimates. *Journal of Social Issues, 48*, 61-75.
- Koss, M. P. (1993a). Detecting the scope of rape: A review of prevalence research methods. *Journal of Interpersonal Violence, 8*, 198-222.
- Koss, M. P. (1993b). Rape: Scope, impact, interventions, and public policy responses. *American Psychologist, 48*, 1062-1069.
- Koss, M. P. (1996). The measurement of rape victimization in crime surveys. *Criminal Justice and Behavior, 23*, 55-69.
- Koss, M. P., Abbey, A., Campbell, R., Cook, S., Norris, J., Testa, M., Ullman, S., West, C., & White, J. (2007). Revising the SES: A collaborative process to improve assessment of sexual aggression and victimization. *Psychology of Women Quarterly, 31*, 357-370.
- Koss, M. P., & Gidycz, C. A. (1985). The Sexual Experiences Survey: Reliability and validity. *Journal of Consulting and Clinical Psychology, 53*, 442-443.
- Koss, M. P., Gidycz, C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Counseling and Clinical Psychology, 55*, 162170.
- Koss, M. P., & Oros, C. (1982). Sexual Experiences Survey: A research instrument investigating sexual aggression and victimization. *Journal of Consulting and Clinical Psychology, 50*, 455-457.
- Lauritsen, J. L. (2005). Social and scientific influences on the measurement of criminal victimization. *Journal of Quantitative Criminology, 21*, 245-266.
- Lynch, J. P. (1996a). Clarifying divergent estimates of rape from two national surveys. *Public Opinion Quarterly, 60*, 410-430.
- Lynch, J. P. (1996b). Understanding differences in the estimates of rape from self-report surveys. In R. J. Simon (Ed.), *From data to public policy: Affirmative action, sexual harassment, domestic violence and social welfare* (pp. 121-142). Lanham, MD: University Press of America.
- Saltzman, L. E. (2000a). Building data systems for monitoring and responding to violence against women, part 1. *Violence Against Women, 6*, 699-704.
- Saltzman, L. E. (2000b). Building data systems for monitoring and responding to violence against women, part 2. *Violence Against Women, 6*, 811-814.
- Saltzman, L. E. (2004). Issues related to defining and measuring violence against women: Response to Kilpatrick. *Journal of Interpersonal Violence, 19*, 1235-1243.
- Schwartz, M. D. (2000). Methodological issues in the use of survey data for measuring and characterizing violence against women. *Violence Against Women, 6*, 815-838.
- Smith, M. D. (1987). The incidence and prevalence of woman abuse in Toronto. *Violence and Victims, 2*, 173-187.
- Tjaden, P., & Thoennes, N. (1998). *Prevalence, incidence, and consequences of violence against women: Findings from the National Violence Against Women Survey*. Washington, DC: National Institute of Justice and Centers for Disease Control and Prevention.
- U.S. Department of Justice, Bureau of Justice Statistics. (2000). *Criminal victimization in the United States, 1995*. Washington, DC: Author.
- Walby, S., & Myhill, A. (2001). New survey methodologies in researching violence against women. *British Journal of Criminology, 41*, 502-522.

Bonnie S. Fisher is a professor in the Division of Criminal Justice at the University of Cincinnati. She was the principal investigator for four federally funded research projects involving the victimization of college students, sexual victimization of and violence against college women, and campus-level responses to reports of sexual assault. Her research interests include repeat victimization, the self-protection–sexual victimization nexus, fear of crime, and the measurement of victimization. She is currently a co–principal investigator on a National Institutes of Health grant to examine the quality of medico-legal evidence—in particular, the detection of vaginal-anal injury—in sexual assault cases.