Psychology of Women Quarterly, **28** (2004), 256–265. Blackwell Publishing. Printed in the USA.

ASSESSING WOMEN'S EXPERIENCES OF SEXUAL AGGRESSION USING THE SEXUAL EXPERIENCES SURVEY: EVIDENCE FOR VALIDITY AND IMPLICATIONS FOR RESEARCH

Maria Testa, Carol VanZile-Tamsen, and Jennifer A. Livingston
University at Buffalo

Mary P. Koss
University of Arizona

In this study we examined the ability of a modified Sexual Experiences Survey (SES; Koss, Gidycz, & Wisniewski, 1987) to assess sexual victimization among a local community sample of women (n = 1,014). Women who reported sexual victimization were interviewed regarding the most recent incident. Those who responded negatively to all SES items were asked whether they had ever feared they would be sexually assaulted but were not, and to describe that incident. Independent coders read a subset of transcripts (n = 137) and classified each incident as reflecting: one of the SES items, a form of unwanted sex not included on the SES, or not unwanted sex. Coders viewed nearly all incidents elicited by the SES as reflecting some type of unwanted sex. Respondent-coder agreement for rape and coercion incidents was high, but low for contact and attempted rape incidents. The SES scoring continuum, reflecting objective severity of acts, was only modestly associated with subjective trauma associated with rape, attempted rape, coercion, and contact.

Self-report is often the only way of obtaining information on many private and sensitive experiences because they are not officially reported, witnessed, or even disclosed to others. This is particularly true in the case of women's sexual aggression experiences, which are believed to be stigmatizing and prone to underreporting. Crime surveys, which ask about rape and sexual assault in the context of other violent crimes, typically obtain very low estimates of prevalence (see Koss, 1993, 1996, for reviews). Similarly, use of a single global screening item, followed by additional questions only for those who respond positively, also appears to result in underestimation of the rate of rape (Koss, 1993). Questions that use terms such as "rape" are also likely to yield underestimates because the majority of women do not label their experiences as rape, even when they meet the legal definition (Fisher, Cullen, & Turner, 2000; Koss, 1988; Layman, Gidycz, & Lynn, 1996).

The use of multiple, behaviorally specific questions describing concrete events has been advocated as a way of assessing sexual assault experiences (Koss, 1993). Use of such items has the advantage of providing memory cues, thereby increasing the number of events that are able to be recalled from memory (Converse & Presser, 1986). Moreover, with such an approach, a respondent is not required to apply a potentially stigmatizing label to her experiences. Supporting the use of behaviorally specific questions, Fisher et al. (2000) found that estimates of sexual assault prevalence obtained using behaviorally specific questions were 4 to 11 times higher than estimates obtained using nonspecific questions regarding criminal victimization, using otherwise identical methodology. Although use of multiple, behaviorally specific questions has its advantages, it is not without disadvantages. The effectiveness of a question as a memory cue depends on the specificity of the cue and its relationship to representations of events stored in memory (Schwartz, 2003). The meaning that a researcher intends to

Maria Testa, Carol VanZile-Tamsen, and Jennifer A. Livingston, Research Institute on Addictions, University of Buffalo; Mary P. Koss, Department of Psychology, University of Arizona.

This research was supported by grants R01 AA12013 from the National Institute on Alcohol Abuse and Alcoholism and NIH Director's Office of Research on Women's Health and K02 AA00284 from the National Institute on Alcohol Abuse and Alcoholism. We thank Gina Bellavia, Ronda Dearing, and Julie Schumacher for serving as coders; Quince Hopkins and Jill Lowell for their legal research; and Kimberly Bauerlein and Cassandra Hoebbel for their assistance in preparing this manuscript. Portions of this research were presented at the annual meetings of the American Psychological Association in Chicago, Illinois, August 2002 and Toronto, Ontario, Canada, August 2003.

Address correspondence and reprint requests to: Maria Testa, Research Institute on Addictions/University at Buffalo, 1021 Main Street, Buffalo, NY 14203. E-mail: testa@ria.buffalo.edu

convey with a given question may not be the meaning that a respondent extracts from that question. Further, when a respondent finds a question difficult to comprehend or answer, he or she is likely to modify the question in a way that makes it easier to answer (Converse & Presser, 1986). It is also possible that behaviorally specific questions may be so specific as to fail to cue memories of related and relevant events that do not exactly match the description.

The Sexual Experiences Survey (SES; Koss, Gidycz, & Wisniewski, 1987) has been used extensively over the past two decades as a way of assessing women's experiences of sexual assault. This 10-item measure provides multiple concrete examples of sexually aggressive experiences, intended to facilitate recall of relevant events while avoiding use of potentially stigmatizing labels such as rape and sexual assault. The measure is commonly used not only to identify victims of sexual aggression but also to determine the nature of women's sexually aggressive experiences and to classify them as rape, attempted rape, verbally coerced intercourse, or sexual contact. In contrast to the low rates of sexual assault derived from crime surveys, studies that have used the SES typically reveal that approximately half of college students and young adult women report some type of sexual aggression since age 14, with 15–20% of young women reporting rape (Abbey, Ross, McDuffie, & McAuslan, 1996; Humphrey & White, 2000; Koss et al., 1987; Testa, Livingston, VanZile-Tamsen, & Frone, 2003).

Despite its widespread usage and acceptance as the best available measure of sexual aggression, psychometric data on the SES are sparse. One study examined how well the SES items conform to legal definitions of rape and other sex offenses, using prosecuting attorneys to evaluate the items (Gylys & McNamara, 1996). Two of three rape items were rated as accurate reflections of relevant statutes; however, only a minority of attorneys viewed items describing alcohol-related rape and attempted rape, and verbally coerced contact as reflective of existing laws. A few studies have examined whether women's reports of sexual aggression are consistent across administrations of the measure. Koss and Gidycz (1985) report test-retest reliability of .93 over 1 week using the Koss and Oros (1982) version of the Sexual Experiences Survey, rather than the more commonly used Koss et al. (1987) version. Similarly, using a German translation of the SES that included items from the Koss and Oros and the Koss et al. versions, Krahe, Reimer, Scheinberger-Olwig, and Fritsche (1999) found that overall, there was high agreement between responses on two administrations (95%). However, consistency was more modest when an exact match between items between two administrations of the test was required (69%). That is, although women were consistent in reporting whether or not they had experienced sexual aggression, they did not always endorse the same items on both administrations.

It has been common to score the SES to reflect an objective severity continuum. Using the Wolfgang Crime Severity Index (Wolfgang, Figlio, Tracey, & Sincer, 1985) as a guide, Koss et al. (1987) developed a scoring method in which a woman reporting rape, the most severe experience, is assigned a score of 4. Women whose most serious experience is attempted rape are assigned a score of 3, followed by coercion (2), and contact (1). Women who report no victimization may be assigned a 0. This continuum score, or a variant of it, has been used as both a predictor of subsequent outcomes (e.g., Gidycz, Coble, Latham, & Layman, 1993) and as an outcome measure in studies examining predictors of sexual assault (e.g., Abbey, Clinton-Sherrod, McAuslan, Zawacki, & Buck, 2003; Abbey et al., 1996; Koss & Gaines, 1993). Although the continuum of increasing severity has face valid appeal, psychometric evidence in support of this scoring method is limited. Using the Koss and Oros (1982) version of the SES, which yields a somewhat different scoring continuum (no victimization, sexual coercion, sexual abuse, and sexual assault), Koss and Gidycz (1985) report that the correlation between level of victimization reported on the paper and pencil measure, and level of victimization reported to an interviewer was .73. However, this correlation may be artificially high because it includes women who consistently reported no victimization on both occasions (e.g., Scott & Aneshensel, 1997). Nearly one quarter of women who reported some type of sexual aggression (23.5%) changed their responses from the survey to the interview, resulting in reclassification. Twice as many of these reclassifications resulted in a lower as opposed to a higher category of victimization, although only 3% of rape victims were subsequently reclassified.

The current study was undertaken to examine the ability of the SES to assess experiences of sexual aggression among a representative local community sample of young women. Traditional approaches to validity are not feasible due to the hidden, private nature of most sexual aggression experiences and hence the absence of a "gold standard" (Sudman & Bradburn, 1982). Rather, we examined whether independent observers would classify events underlying women's positive responses to the SES in a manner comparable to the respondent's endorsement. That is, after reading a description of an experience that caused the respondent to endorse, for example, a rape item, would an independent observer also endorse rape as the best description of the incident? The association between SES items endorsed by the respondent and SES items endorsed by independent coders provides an estimate of the sensitivity of the measure in distinguishing between different types of sexual aggression. Moreover, by providing coders with the opportunity to indicate that an incident was not reflective of unwanted sex, we are able to estimate the specificity of the SES or the extent to which it appears to yield false positives.

In establishing the validity of the SES, it is important to consider whether the measure fails to detect certain experiences that might reasonably be considered sexually aggressive. In our study, women who responded negatively to all SES items were asked to describe an incident in which they feared that sexual aggression might occur but did not. Descriptions of these events were also provided to independent coders who were instructed to choose the SES item

that best described the event, to indicate that the event was unwanted sex but not adequately described by any of the items, or to indicate that the incident was not an incident of unwanted sex. In this manner, we examined the extent to which the SES may fail to identify women with other types of unwanted sex experiences (false negatives) and considered the nature of these other incidents.

Finally, given that the SES is frequently scored as a continuum thought to represent the objective severity of sexually aggressive acts perpetrated upon the woman, we examined how well this continuum predicted subjective severity associated with the incident. This may be viewed as one way of establishing the external validity of the continuum approach. Previous research has found a positive correlation between objective severity, represented by the SES continuum, and retrospective reports of trauma immediately following the incident (Testa & Livingston, 1999). Similarly, severity of previous victimization was associated with current depression and anxiety levels (Gidycz & Koss, 1989). As a way of approximating subjective severity, we examined women's self-reported trauma associated with sexual assault incidents, both at the time the incident occurred and at present. We hypothesized, consistent with the continuum approach, that rape incidents would be considered the most traumatic, followed by attempted rape, coercion, and contact incidents.

METHOD

Sample

The current sample consisted of 1,014 women participating in the first wave of a three-wave prospective study of alcohol and sexual behavior. Women 18–30 years of age living in Buffalo, New York and its immediate suburbs in Erie County were identified using random-digit dialing between May 2000 and April 2002. In-person interviews were completed with 61% of eligible women identified, a rate that is comparable or superior to completion rates for surveys that were conducted solely by telephone (Greenfield, Graves, & Kaskutas, 1999; Welte, Barnes, Wieczorek, Tidwell, & Parker, 2001). The sample matched closely the characteristics of the local population. For example, 75% of the sample was White and 17% were African American, compared to 72% and 21%, respectively, for the geographic area from which the sample was drawn. Also consistent with local demographics, median household income for the sample was between $30,000 and $40,000, and 95% were high school graduates (compared to 89% of 18–34-year-old women in Erie County). The man age was 23.76 ($SD = 3.71$) and most were currently unmarried (76% never married, 3% divorced or legally separated).

Procedure

Eligible women were recruited to participate individually in an initial data collection session at the Research Institute on Addictions, University at Buffalo (RIA). They were told that they would be asked about aspects of their personality, alcohol and drug use, and sexual experiences using computer-administered questionnaires and a face-to-face confidential interview. Upon arrival at RIA, study procedures were explained and informed consent obtained. Participants were told that some questions might make them uncomfortable and informed that they had the right to refuse to answer any question or to discontinue at any time. Refusal options were provided for all computer-administered questions. Permission to tape-record the interview was also obtained; however, women could decline to be taped and still be interviewed. Women were paid $50 for participation.

We opted to use computer-assisted self-interview (CASI) for most questionnaires, including the SES, because studies suggest that participants are more willing to report sensitive behaviors using this mode of administration (Gribble, Miller, Rogers, & Turner, 1999). However, we conducted semi-structured face-to-face interviews, administered by trained female interviewers, to gather in-depth information about incidents of sexual assault.

Measures

Sexual aggression experiences. The SES (Koss et al., 1987) was used to assess sexual aggression experiences occurring since age 14. Women indicated whether or not each experience had occurred, and for each item endorsed, indicated how old they were the most recent time it occurred. We made modifications to the wording of several items for clarity's sake (see Table 1 for wording of items). Pilot testing suggested that women found the term *sex play* off-putting and confusing. Hence, the sexual contact items were rephrased, inquiring whether the woman had ever been "fondled, kissed, or touched sexually." The original phrasing of the attempted rape items was also judged to be confusing and hence reworded to simplify grammar. The original Koss et al. (1987) items describing completed or attempted sexual intercourse "because a man gave you alcohol or drugs" were modified to specify that substances were given "without your knowledge or consent" to make them conform more closely with legal statues (Gylys & McNamara, 1996). Finally, we added an item, adapted from Rodzinka et al. (1999), to reflect rape that occurs because a woman is too incapacitated by alcohol or drugs to resist.

Event-level interview. After the respondent completed the CASI, the interviewer retrieved a summary of her SES responses from the computer. Women who endorsed one or more items on the SES were subsequently interviewed regarding the most recent occurrence of sexual aggression. We opted to focus on the most recent incident, as opposed to the most severe or best remembered, to obtain a sampling of the full range of sexual aggression incidents. For participants who reported some sexual aggression, the

Table 1

Prevalence of Sexual Experiences Since Age 14: Modified Sexual Experiences Survey

	Number	*Percent*
1. Have you ever been fondled, kissed, or touched sexually when you didn't want to because you were overwhelmed by a man's continual arguments and pressure?	220	21.7
2. Have you ever been fondled, kissed, or touched sexually when you didn't want to because a man used his position of authority (boss, teacher, camp counselor, supervisor) to make you?	30	3.0
3. Have you ever been fondled, kissed, or touched sexually when you didn't want to because a man threatened or used some degree of physical force (twisting your arm, holding you down, etc.) to make you?	129	12.7
Total Reporting Sexual Contact	**271**	**26.7**
4. Have you given in to sexual intercourse when you didn't want to because you were overwhelmed by a man's continual arguments and pressure?	185	18.2
5. Have you had sexual intercourse when you didn't want to because a man used his position of authority (boss, teacher, camp counselor, supervisor) to make you?	10	1.0
Total Reporting Sexual Coercion	**191**	**18.8**
6. Have you had a man attempt to insert his penis (but intercourse did not occur) when you didn't want him to by threatening or using some degree of force (twisting your arm, holding you down, etc.)?	82	8.1
7. Have you ever had a man attempt to insert his penis (but intercourse did not occur) when you didn't want him to by getting you intoxicated on alcohol or drugs without your knowledge or consent?	51	5.0
Total Reporting Attempted Rape	**121**	**11.9**
8. Have you had sexual intercourse when you didn't want to because a man made you intoxicated by giving you alcohol or drugs without your knowledge or consent ?	29	2.9
9. Have you been in a situation in which you were incapacitated due to alcohol or drugs (that is, passed out or unaware of what was happening) and were not able to prevent unwanted sexual intercourse from taking place?	86	8.5
10. Have you had sexual intercourse when you didn't want to because a man threatened or used some degree of physical force (twisting your arm, holding you down, etc.) to make you?	100	9.9
11. Have you had sex acts (anal or oral intercourse or penetration by objects other than the penis) when you didn't want to because a man threatened or used some degree of physical force (twisting your arm, holding you down, etc.) to make you?	32	3.2
Total Reporting Rape	**174**	**17.2**

interviewer began the interview by reminding the woman of the item or items she had endorsed as occurring at the most recent age. If more than one item was reported at the most recent age, the respondent was asked whether the experiences had occurred separately or as part of a single incident, the goal being to identify the most recent event. Once the most recent incident of sexual assault was identified, the woman was asked to describe the event in her own words. Respondents also were asked specific questions about the incident, including when it occurred, her relationship to the perpetrator, the activities that preceded the incident, whether penetration had occurred, whether physical force or injury was involved, her responses or resistance to his actions, and the alcohol and drug use of the woman and the perpetrator at the time of the assault.

Women who endorsed no items on the SES were subsequently asked, "Have you ever had an experience in which you were concerned that a man might be sexually aggressive toward you, or attempt to go further sexually than you were prepared to go, but for whatever reason no aggression occurred?" This question was intended to provide us with an understanding of incidents that women perceive as sexually threatening but do not result in endorsement of any items on the SES, either because they don't fit the wording provided, or because the woman was able to escape or diffuse the situation before it evolved into something more serious. We worded the question to focus on events that did not result in completed intercourse, because we believed that these were the experiences most likely to be missed by the SES (e.g., attempted coercion), particularly with the addition of the incapacitated rape item. Women who responded positively to this close-call item were subsequently asked a series of questions, similar to those asked regarding incidents of sexual aggression reported on the SES.

Trauma

At the conclusion of the interview regarding the most recent sexually aggressive event or the close-call incident, women were asked, "At the time that these experiences happened, how upsetting or traumatic were they for you?" Women responded using a 6-point Likert scale, ranging from 1 (*not at all traumatic*) to 6 (*most traumatic thing possible*). They used the same scale to answer the question "As you look back at these experiences, how traumatic or upsetting are these experiences for you now?"

Coder Ratings

Participants' descriptions of their most recent incident of sexual assault, as well as incidents in which they feared sexual assault (close calls), were transcribed from audiotaped interviews. A subset of these transcripts was provided to three independent coders. All three were female scientists at RIA who had doctorates in psychology but no prior experience in sexual aggression research, no familiarity with the SES, and no connection to the study. We chose to use postdoctoral fellows as coders both because of the cognitive and emotional difficulty level of the task, as well as the fact that we could not violate respondent confidentiality by sharing transcripts of actual incidents with persons outside of RIA.[1] Coders were provided with a list of SES items identical to those given to study participants and instructed to choose the SES item that best fit the description provided in the transcript. Coders were not given explicit definitions, but rather were asked to use their best lay judgment in selecting the item that best fit each incident description. We did this deliberately so as to approximate the decision that a respondent makes in deciding whether to endorse a particular item based on her experiences and so as to avoid transmitting any of our own biases. Coders were also given the option of coding the incident as: (a) not an incident of unwanted sex, (b) an incident of unwanted sex but none of the items are adequate to describe it, or (c) insufficient information to code. We used the term *unwanted sex* because pilot research suggested that many women interpret the phrase *sexually aggressive* as indicating use of physical force. Coders were reminded not to over-think their codes and not to discuss codes with each other.

RESULTS

Prevalence of Sexual Assault Experiences

Of 1,014 respondents, 383 (37.8%) responded positively to one or more items on the SES, indicating some sexual victimization since age 14. Of the 383 women who reported sexual aggression, 243 responded positively to two or more items ($M = 3.35$, $SD = 1.50$). By asking the age at which the most recent incident of each type had occurred, we were able to determine that for at least 138 respondents these represented multiple incidents. However, for 105 women, multiple items were reported as occurring at the same age, suggesting that experiences may have occurred within a single incident.[2] Internal consistency among the items, calculated using Cronbach's alpha, was .73 and nearly identical to that reported by other researchers (e.g., Abbey et al., 1996; Koss et al., 1987). However, this inter-item consistency reflects both the occurrence of multiple incidents of sexual assault and of multiple experiences occurring within the same incident.

The frequencies with which women endorsed each SES item are presented in Table 1. In addition, the percent reporting each type of sexual assault—contact (items 1, 2, 3), coercion (items 4 and 5), attempted rape (items 6 and 7), and rape (items 8–11)—are presented. The continuum method of scoring the SES may be used to assign a score based on the most serious level of aggression experienced. For example, if a woman experienced both coercion and rape, she would be assigned the highest score, reflecting her rape experience. Given that multiple acts of sexual aggression may occur as part of the same incident, this scoring method avoids the potential overestimation of prevalence that could occur from reporting on the independent frequency of various acts. Using the continuum scoring method, 174 (17.5%) reported rape as the most serious experience, 47 (4.7%) reported attempted rape, 97 (9.8%) reported sexual coercion, 65 (6.6%) reported contact, and 609 (61.4%) reported no aggression.

For subsequent analyses, we focused on the most recent incident of sexual assault, recognizing that this may not coincide with the lifetime continuum score (e.g., a woman who was raped last year but coerced last month would report on the coercion incident). Event-based interviews regarding the most recent incident of sexual aggression were completed for 361 of the 383 women who reported some sexual aggression using the CASI-administered SES. These incidents included 123 rapes, 38 attempted rapes, 126 coercions, and 74 contact incidents. Similar interviews were completed with the 71 women (7.0% of the sample) who reported no SES experiences but indicated that they had experienced a close call.[3] On average, incidents had occurred 4.53 years ago ($SD = 3.80$).

The categorization of sexual assault experiences according to the SES presumes that rape and coercion involve penetration, whereas contact and attempted rape do not. Women were asked, during the event-level interview, whether penetration had occurred, providing us an opportunity to examine the correspondence between SES responses and a direct question regarding penetration. As expected, nearly all incidents categorized as penetrative according to the SES (i.e., rape or coercion) were reported to have involved vaginal or anal penetration (96.4%). However, 18.9% of contact and attempted rape incidents also were reported in the interview to have involved vaginal or anal penetration. This substantial proportion of false negatives suggests that the SES may underreport the severity of women's experiences.

Because the four types of sexual aggression incidents were not equally prevalent, transcripts were randomly selected for coding to ensure that relatively equal numbers of the four types were included in the sample. Hence, coders were provided with interview transcripts representing 16 rapes, 19 attempted rapes, 21 coercions, and 20 contact incidents. In addition, they were provided with all usable transcripts describing close-call incidents ($n = 61$), yielding a total of 137 transcripts for each coder.

Table 2

Frequency of Coding Matches (Number of Possible Coding Matches per Category in Parentheses)

Categories Based on Participant SES Responses				
Assigned Codes	*Contact (n = 60)*	*Coercion (n = 63)*	*Attempted Rape (n = 57)*	*Rape (n = 48)*[1]
Contact	34	0	23	0
Coercion	0	56	0	2
Attempted Rape	4	0	23	0
Rape	4	6	1	42
Not Enough Information	0	0	1	0
Not Unwanted Sex	2	0	0	4
Unwanted Sex Not on the SES	16	1	9	0

[1]Sample size for each category is the number of transcripts per category multiplied by the number of coders (3).

Agreement Between Respondent SES Categories and Coder Ratings: Sexual Aggression Incidents

Although the SES contains 10 (or in this case, 11) individual items, prevalence data are most frequently presented according to the four categories of sexual aggression: rape, attempted rape, coercion, and contact. Hence, we converted individual item ratings provided by respondents and by coders to category ratings and computed agreement statistics based upon category, not individual item, agreement. We used as our gold standard the classification of an incident according to the participant's response and examined the extent to which the coders' classifications matched those of the respondent. Table 2 depicts agreement between respondents and the three coders. For each transcript there are three possible points of agreement. For example, for the 16 rape transcripts that were coded by the three coders, perfect agreement would involve 48 (16 × 3) matches between respondent and coders. For the four types of sexual aggression—contact, coercion, attempted rape, and rape—the numbers on the diagonal show the extent to which the three coders classified the incidents as the respondents did. The numbers on the off-diagonals reveal systematic biases in the way incidents were viewed. For example, it was not uncommon for incidents classified as contact by the respondent to be labeled as examples of unwanted sex not described by any SES item. These incidents included attempted coercion, touching without force, or consensual kissing or petting that began to go too far. Attempted rape incidents were frequently classified as contact by the coders. Many of these incidents involved a man getting on top of the woman and touching her or removing her clothing; however, the incident ended before he attempted to insert his penis.

To determine whether there are individual differences in how sexual aggression incidents are classified, we computed percent agreement between respondent and each individual coder. As another way of expressing agreement, we also computed separate Kappa statistics. Cohen's Kappa is a measure of agreement between an independent rater and a participant's response on an item, in this case the type of sexual victimization experienced. Kappa is preferred over percent agreement because it is chance-corrected and indicates the magnitude of agreement over and above what would be expected due to chance alone, thus it provides a measure of inter-rater reliability for a measure or classification system. The closer the magnitude of the kappa coefficient to 1, the greater the level of agreement and the greater the reliability. Agreement between individual coders and respondent was consistently high for incidents classified as rape (81–94% agreement, Kappa .76–.81) or coercion (86–95% agreement, Kappa .86–.93) according to respondent codes. However, agreement between participant and coder ratings was much more variable, and lower overall, for incidents classified as contact (40–85% agreement, Kappa .32–.53) and attempted rape (16%–68%, Kappa .22–.64).

To better understand the nature of disagreements for the contact and attempted rape incidents, we examined the pattern of agreement versus disagreement for the three coders separately. This revealed distinct individual differences in how the three coders viewed the incidents. For example, Coder 2 coded 16 of 20 contact incidents "correctly"; however, she did quite poorly in coding attempted rape incidents because of an apparent bias toward coding incidents without penetration as contact rather than as attempted rape. In contrast, Coder 1 did fairly well recognizing attempted rape incidents as such but poorly in identifying contact. This coder was the only one of the three who classified some contact incidents as attempted rape. Coder 3 appeared to interpret the SES items literally, applying the code "not assessed by SES" more liberally for contact incidents that involved elements of attempted verbal coercion or that did not explicitly involve physical force, authority, or verbal pressure (e.g., he just grabbed her).

Despite difficulties in distinguishing among the types of incidents, particularly contact and attempted rape, it is noteworthy that coders overwhelmingly indicated that the incidents were reflective of some type of unwanted sex,

either one of the SES items, or some other type of unwanted sex not assessed by the SES. Of the 57 transcripts representing incidents of sexual aggression, there were only six codes of "not unwanted sex" assigned out of the 228 codes assigned by the three coders.[4]

Close Calls

We also considered the extent to which the SES may fail to capture some sexually aggressive incidents. Women who responded negatively to all SES items but reported an incident in which they feared they might experience sexual aggression were interviewed regarding these close-call incidents. As described above, transcripts of these event-based interviews ($n = 61$) were provided to the coders interspersed with transcripts representing SES-elicited events and coded in the same manner. Coders most frequently described these incidents as unwanted sex not on the SES (79/183) or as contact (61/183). Many of these close-call experiences involve attempts by a man to go farther than the woman wanted to go, using verbal coercion and sometimes physical force. What stands out in these descriptions is that the man's advances are thwarted, through what the woman perceives as her own efforts (telling him no, threatening him) or through someone else's intervention, as illustrated in the following example:

> I was at a party at my friend's house and I was just sitting around talking with this kid and he seemed really nice. He said, "Let me go introduce you to my friend" and I said okay. It was his house and he took me in the back of the house and brought me into the bedroom and shut the door and locked it and he was like trying to take my shirt off. I kept telling him I had a boyfriend. He pushed me down on the bed and started to kiss me and take my shirt off. The only reason I honestly think it didn't happen is because someone came and knocked on the door and I jumped up and ran out the door.

Other close-call situations included a stranger exposing himself, a man grabbing the woman in a bar, or a situation in which the woman felt threatened due to a man's proximity, suggestive language, or isolated situation but sexual contact did not occur.

Trauma Following Sexual Assault

The continuum method of scoring the SES is thought to represent increasing objective severity of sexual assault experiences, ranging from contact to rape. Consistent with the ordering thought to underlie this continuum, we hypothesized that subjective trauma following these experiences would follow this pattern, with more trauma associated with increasingly severe types of experiences. After describing the incident of sexual assault or the close call, women were asked: (a) how traumatic or upsetting the incident was at the time, and (b) how traumatic or upsetting the incident is now. To maximize power for these analyses, we used all sexual aggression incidents ($n = 361$) and close-call incidents ($n = 65$) for which we had quantitative data. Level of victimization was positively correlated with ratings of immediate trauma ($r = .18$, $p < .001$) and later trauma ($r = .24$, $p < .001$). Next, we performed a repeated measures ANOVA followed by simple effects Tukey b post hoc tests, to determine whether ratings of immediate trauma and current trauma were ordered according to the severity continuum thought to underlie the SES (with close calls considered as less severe than contact). As shown in Table 3, there was a significant main effect for time, such that subjective trauma was higher immediately after the incident ($M = 4.09$, $SD = 1.44$) compared to currently ($M = 2.84$, $SD = 1.58$), $F(1, 416) = 241.76$, $p < .001$. For retrospective ratings of trauma immediately following the incident, respondents rated coercion incidents as significantly less traumatic than all other incidents, including close calls. Rape incidents were rated the most traumatic but did not differ from contact or attempted rape. For ratings of current trauma, rape incidents were associated with significantly more current trauma than all other types of incidents, which did not differ from each other.

DISCUSSION

Findings suggest that the SES functions well as a surveillance measure of sexual aggression. Virtually all of the incidents underlying women's positive response to the SES were rated by independent coders as reflecting one of the SES items or, more rarely, as reflecting some other type of unwanted sex not assessed by the measure. Hence, there is

Table 3

Mean Trauma Resulting from Different Types of Sexual Aggression Experiences

	Type of Incident				
	Close Call ($n = 65$)	*Contact* ($n = 72$)	*Coercion* ($n = 126$)	*Attempted Rape* ($n = 37$)	*Rape* ($n = 123$)
Trauma Then	3.90^{a}	4.32^{ac}	3.44^{b}	4.22^{ac}	4.67^{c}
SD	(1.47)	(1.31)	(1.47)	(1.13)	(1.30)
Trauma Now	2.29^{a}	2.74^{a}	2.72^{a}	2.19^{a}	3.53^{b}
SD	(1.39)	(1.74)	(1.51)	(1.45)	(1.47)

Note. Within a row, means with different superscripts differ at the $p < .05$ level.

little to suggest that the SES is overly inclusive in classifying women as victims of sexual assault.

On the other hand, women's responses to our question regarding incidents in which she feared sexual aggression suggest that the measure may fail to capture some incidents that might reasonably be considered sexually aggressive, resulting in some false negatives. Of 1,014 women, there were 71 (7%) who responded negatively to all SES items but indicated that they had experienced an incident in which they feared they would be sexually assaulted but were not. A majority of these incidents were coded as representing either contact or some other type of unwanted sex not included on the SES. Although many of these incidents involved primarily a perceived threat of sexual assault with little or no physical contact (e.g., sexual harassment), others appear similar to contact or attempted rape incidents elicited by the SES in terms of the perpetrator's behavior (e.g., verbal and physical pressure to have sex). They differ in their emphasis on how the woman's resistance or outside intervention halted the episode. This focus on "what could have happened but didn't" likely reflects the particular wording of question we used to elicit these experiences ("but for whatever reason, sexual assault did not occur"). Nonetheless, the close-call data suggest, importantly, that asking about potential sexual assault experiences in a different way may elicit recall of experiences not elicited by the SES items.

Regarding the precision of the SES, rape and coercion incidents were identified by coders with a good deal of accuracy, suggesting that these questions do a good job of cuing recall of incidents that match definitions of these constructs. However, coders had difficulty distinguishing contact and attempted rape incidents. Moreover, many close-call incidents appeared similar to incidents elicited by the SES in that they involved a man using a variety of sexually aggressive strategies, verbal and physical, in an apparent attempt to have sexual intercourse. There are several possible reasons for this difficulty in distinguishing among incidents involving noncompleted intercourse.

First, it is possible that poor coder agreement for contact and attempted rape reflects weaknesses in our method. We altered the contact and attempted rape items to simplify wording and grammar following comments from respondents and interviewers that the original questions were awkwardly phrased. Modification of the SES is common in sexual assault research (e.g., Abbey et al., 1996; Humphrey & White, 2000). In fact, few have used the SES in its original form. However, subtle changes in wording may alter the way questions are interpreted, at least by some respondents, and skew the types of events that are recalled (or not recalled). For example, our rate of verbally coerced contact was half the 44% rate reported by Koss et al. (1987) in her national college sample (using the original sex play wording). The majority of contact incidents described by respondents in this study appeared fairly serious, involving at least some physical pressure or force. It is possible that the changes in wording failed to elicit milder incidents, although it is also possible that the difference in prevalence resulted from differences in sample (college vs. community), age (older women forget milder incidents), or time period. The rate of physically forced contact was nearly identical in this sample (12.9%) as in Koss et al.'s (13%) despite the change in wording.

It is also possible that difficulty in distinguishing contact and attempted rape incidents reflects the quality of the narratives that we obtained, due to self-censoring by respondents or inadequate questioning by interviewers. Rape and coercion are relatively easy to identify because we explicitly asked whether intercourse took place, and there are readily apparent differences among incidents completed due to verbal coercion, physical force, or incapacitation (Abbey et al., 1996; Livingston, Buddie, Testa, & VanZile-Tamsen (in press); Testa et al., 2003). However, distinguishing between unwanted contact and attempted rape may require more information than we typically obtained. We did not, for example, ask whether the man attempted to insert his penis and women rarely volunteered such information, typically describing the incident in more general terms without reference to specific sexual acts.

We believe, however, that there is a true blurring of the line between contact and attempted rape, and that individuals may draw that line in different places. Most incidents that involve noncompleted intercourse include a combination of physical touching, some force (e.g., getting on top of her, pushing her against the wall), and verbal indications that intercourse is desired. However, the SES items are worded narrowly, referring to contact "due to continual arguments and pressure" or "due to physical force," whereas the attempted rape item includes the phrase "attempt to insert penis." Whether or not a woman responds positively to the attempted rape item may depend upon her perception of the perpetrator's intent, rather than the actual severity of sex acts. For example, the following incident resulted in endorsement of forcible attempted rape, although it ends before intercourse is actually attempted:

> I was babysitting. Taking care of his two girls. His wife and him had went out. He returned early. We were sleeping. I had put the girls to sleep and I was sleeping with them. All of a sudden, I felt my sheets start to rise and I felt him touching me. I was so scared, I just couldn't move. As he starts to crawl up onto the bed, his daughter felt the motion of the bed and she turned around and just looked at him and he would have raped me if it wasn't for his little girl because his intent was to do that. He had sexual intentions on me and I felt so violated.

It is important to note that most legal statutes do not distinguish between attempted rape and sexual contact, reflecting a difficulty in determining the perpetrator's intent apart from his actual behaviors. Case law establishes that the difference between attempted rape and contact is whether the perpetrator had the intent to rape. The authors of the

original SES believed that the phrase "attempted to insert his penis" was a behavioral manifestation of intent to rape. More direct reference to the perpetrator's intent would be another strategy to differentiate attempted rape from contact; however, a perpetrator's intentions may not always be apparent to his victim.

Finally, we found modest support for the predicted linear relationship between objective severity of the SES continuum and subjective ratings of trauma. Rape incidents were rated as marginally more traumatic at the time and significantly more traumatic at present, consistent with other research showing long-lasting trauma associated with rape (e.g., Saunders, Kilpatrick, Hanson, Resnick, & Walker, 1999). However, degree of trauma associated with other types of sexual assault did not conform to the expected pattern. These findings are exploratory and require replication, given that they are based on a novel measure of trauma that may be subject to retrospective biases. Hence, findings do not necessarily negate the validity of the continuum scoring method, which is based upon objective severity. Nonetheless, they serve as a reminder that the continuum score is an imperfect approximation of subjective severity and caution is necessary when considering level of victimization as a predictor of subsequent outcomes or psychological sequelae.

We view this research as largely exploratory and, as noted above, recognize several limitations to our method. Findings were based on a subset of transcripts, coded by a limited number of coders. Moreover, our sample, while large, was limited to women from a single community who were willing to participate in a lengthy in-person interview at an inner-city location. It is not possible to gauge how such women may differ from those who did not participate. Nonetheless, findings provide unique psychometric data regarding the ability of the SES to assess women's sexual aggression experiences. Results affirm the use of multiple, behaviorally specific questions as a means of assessing the range of sexually aggressive incidents that women experience. Coercion and rape items appear to function well in assessing these experiences; however, noncompleted experiences are not as well captured or distinguished by existing items. Researchers may wish to modify these items so that the precision with which contact and attempted rape experiences are measured would equal that of coercion and rape. This may be unrealistic, however, given that perpetrators typically use multiple strategies, verbal and physical, to attempt to have intercourse and the point at which unwanted touching becomes attempted rape is hard to specify. It may be more fruitful to consider all noncompleted sexual aggression experiences as representing a single category, distinct from verbally coerced intercourse and from rape. Modifications of items, ideally, would use language that better reflects the terms that women use (e.g., "he went too far") and the way that they view these incidents (e.g., as a potential sexual assault that was stopped). We urge researchers interested in altering the SES to test empirically the effects of these modifications, since changes in wording may have a significant impact on women's recall and hence on prevalence estimates.

Initial submission: August 13, 2003
Initial acceptance: October 24, 2003
Final acceptance: January 6, 2004

NOTES

1. Similar coding was undertaken using two undergraduate-level RIA research assistants with no connection to the study. The pattern of results was identical to that reported here.
2. Women who reported more than one item at the most recent age were subsequently asked whether these experiences represented a single incident or multiple incidents. The majority (164/206 or 80%) indicated that these represented a single incident, therefore we believe that most women who reported multiple items occurring at the same age had experienced a single incident in which multiple types of sexual aggression occurred.
3. We failed to obtain event-level data for the most recent incident of sexual aggression for 22 respondents due to the following reasons: technical or interviewer error (10), time conflict for interviewer (2) or respondent (4), respondent refused interview (4) or could not think of a specific incident to discuss (1), or interview was discontinued because the event had actually occurred before age 14 (1). An additional six women completed interviews; however, we did not have a usable interview transcript for the following reasons: refused tape recording (3), incomplete transcript (1), or inaudible tape (2). Of the 71 women who reported a close call, all interviews were completed; however, five incidents involved experiences that happened before age 14, and one incident involved only physical, not sexual threat. These cases were dropped from all subsequent analyses. An additional four women refused to be taped; these cases were included in quantitative analyses.
4. All three coders agreed that one particular incident (classified as rape by the respondent) did not involve unwanted sex. The respondent indicated on the SES that she had experienced intercourse when incapacitated. In describing the incident she revealed that she used poor judgment after drinking heavily but that intercourse did not occur against her will. The other three cases in which a coder viewed an incident as not unwanted sex appeared to be idiosyncratic.

REFERENCES

Abbey, A., Clinton-Sherrod, A. M., McAuslan, P., Zawacki, T., & Buck, P. O. (2003). The relationship between quantity of alcohol consumed and the severity of sexual assaults committed by college men. *Journal of Interpersonal Violence, 18,* 813–833.

Abbey, A., Ross, L. T., McDuffie, D., & McAuslan, P. (1996). Alcohol and dating risk factors for sexual assault among college women. *Psychology of Women Quarterly, 20,* 147–169.

Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire*. Beverly Hills, CA: Sage.

Fisher, B. S., Cullen, F. T., & Turner, M. G. (2000). *The sexual victimization of college women* (Rep. No. NCJ 182369). U.S. Department of Justice.

Gidycz, C. A., Coble, C. N., Latham, L., & Layman, M. J. (1993). Sexual assault experience in adulthood and prior victimization experiences: A prospective analysis. *Psychology of Women Quarterly, 17,* 151–168.

Gidycz, C. A., & Koss, M. P. (1989). The impact of adolescent sexual victimization: Standardized measures of anxiety, depression, and behavioral deviancy. *Violence and Victims, 4,* 139–149.

Greenfield, T. K., Graves, K. L., & Kaskutas, L. A. (1999). Long-term effects of alcohol warning labels: Findings from a comparison of the United States and Ontario, Canada. *Psychology & Marketing, 16,* 261–282.

Gribble, J. N., Miller, H. G., Rogers, S. M., & Turner, C. F. (1999). Interview mode and measurement of sexual behaviors: Methodological issues. *Journal of Sex Research, 36,* 16–24.

Gylys, J. A., & McNamara, J. R. (1996). A further examination of validity for the Sexual Experiences Survey. *Behavioral Sciences and the Law, 14,* 245–260.

Humphrey, J. A., & White, J. W. (2000). Women's vulnerability to sexual assault from adolescence to young adulthood. *Journal of Adolescent Health, 27,* 419–424.

Koss, M. P. (1988). Hidden rape. In A. W. Burgess (Ed.), *Rape and sexual assault* (pp. 3–25). New York: Garland.

Koss, M. P. (1993). Detecting the scope of rape: A review of prevalence research methods. *Journal of Interpersonal Violence, 8,* 198–222.

Koss, M. P. (1996). The measurement of rape victimization in crime surveys. *Criminal Justice and Behavior, 23,* 55–69.

Koss, M. P., & Gaines, J. A. (1993). The prediction of sexual aggression by alcohol use, athletic participation, and fraternity affiliation. *Journal of Interpersonal Violence, 8,* 94–108.

Koss, M. P., & Gidycz, C. A. (1985). Sexual Experiences Survey: Reliability and validity. *Journal of Consulting and Clinical Psychology, 53,* 422–423.

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 Consulting and Clinical Psychology, 55,* 162–170.

Koss, M. P., & Oros, C. J. (1982). Sexual Experiences Survey: A research instrument investigating sexual aggression and victimization. *Journal of Consulting and Clinical Psychology, 50,* 455–457.

Krahe, B., Reimer, T., Scheinberger-Olwig, R., & Fritsche, I. (1999). Measuring sexual aggression: The reliability of the Sexual Experiences Survey in a German sample. *Journal of Interpersonal Violence, 14,* 91–100.

Layman, M. J., Gidycz, C. A., & Lynn, S. J. (1996). Unacknowledged versus acknowledged rape victims: Situational factors and posttraumatic stress. *Journal of Abnormal Psychology, 105,* 124–131.

Livingston, J. A., Buddie, A. M., Testa, M., & VanZile-Tamsen, C. (in press). Incidents of verbally coerced sexual intercourse in a representative community sample. *Psychology of Women Quarterly*.

Rodzinka, K. J., Petretic-Jackson, P. A., & Broach, J. L. (1999, November). Sexual assault acknowledgment and substance use. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, Toronto, Ontario, Canada.

Saunders, B. E., Kilpatrick, D. G., Hanson, R. F., Resnick, H. S., & Walker, M. E. (1999). Prevalence, case characteristics, and long-term psychological correlates of child rape among women: A national survey. *Child Maltreatment, 4,* 187–200.

Schwartz, N. (2003). Cognitive research into survey measurement: Its influence on survey methodology and cognitive theory. In M. G. Sirken, D. J. Herrman, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Eds.), *Cognition and survey research* (pp. 65–75). New York: Wiley.

Scott, K. D., & Aneshensel, C. S. (1997). An examination of the reliability of sexual assault reports. *Journal of Interpersonal Violence, 12,* 361–374.

Sudman, S., & Bradburn, N. M. (1982). *Asking questions: A practical guide to questionnaire design*. San Francisco: Jossey-Bass.

Testa, M., & Livingston, J. A. (1999). Qualitative analysis of women's experiences of sexual aggression: Focus on the role of alcohol. *Psychology of Women Quarterly, 23,* 573–589.

Testa, M., Livingston, J. A., VanZile-Tamsen, C., & Frone, M. R. (2003). The role of women's substance use in vulnerability to forcible and incapacitated rape. *Journal of Studies on Alcohol, 64,* 756–764.

Welte, J., Barnes, G., Wieczorek, W., Tidwell, M.-C. O., & Parker, J. (2001). Alcohol and gambling pathology among U. S. adults: Prevalence, demographic patterns and comorbidity. *Journal of Studies on Alcohol, 62,* 706–712.

Wolfgang, M., Figlio, R., Tracey, P., & Sincer, S. (1985). *Wolfgang Crime Severity Index*.Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.