

Recalling Sexual Partners: The Accuracy of Self-reports

JAMES JACCARD

Florida International University, USA

ROBERT MCDONALD

CHOI K. WAN

University at Albany, State University of New York

VINCENT GUILAMO-RAMOS

School of Social Work and Department of Population and Family Health, Columbia University

PATRICIA DITTUS & SHANNON
QUINLAN

University at Albany, State University of New York

JAMES JACCARD is a Professor of Psychology at Florida International University.

PATRICIA DITTUS and CHOI WAN were Senior Research Scientists at the University at Albany, State University of New York when contributing to this project. They are both currently Health Scientists at the Centers for Disease Control and Prevention.

VINCENT GUILAMO-RAMOS is an Assistant Professor of Social Work and Population and Family Health at Columbia University

SHANNON QUINLAN was a Research Associate at the University at Albany, State University of New York when contributing to this project

ROBERT MCDONALD was a Senior Research Scientist at the University at Albany, State University of New York when contributing to this project. He is now deceased.

ACKNOWLEDGEMENTS. This research was supported by grant R01-HD20346-68 from the National Institute of Health.

COMPETING INTERESTS: None declared.

ADDRESS. Correspondence should be directed to:

JAMES JACCARD, Department of Psychology, Florida International University, 11200 S.W. 8th Street, University Park, Miami, FL33199, USA. [email: James.Jaccard@fiu.edu]

Journal of Health Psychology
Copyright © 2004 SAGE Publications
London, Thousand Oaks and New Delhi,
www.sagepublications.com
Vol 9(6) 699–712
DOI: 10.1177/1359105304045354

Abstract

Accuracy of recall of the number of sexual partners individuals had over a period of one month, three months, six months and one year was studied in a group of 285 young, single, heterosexual adults. Self-reports of the number of partners were obtained on a weekly basis and then compared with recall of behavior over longer time periods that overlapped the weekly measures. For individuals who claimed abstinence or who claimed to be monogamous, accuracy of recall was relatively high, especially at the shorter time frames. Level of education was related to accuracy for claimed abstainers, such that lower levels of education were associated with lower accuracy of recall. Accuracy rates for individuals who reported having multiple sexual partners tended to be lower and were found to be related to one's propensity to engage in casual sex.

Keywords

recall, self-reports, sexual partners, sexual risk behavior

SELF-REPORTS of sexual risk behavior are used extensively in behavioral research on unintended pregnancies and sexually transmitted diseases. Numerous methodologists have stressed the need to determine the accuracy of self-reports of sexual risk behavior. Strategies to do so have included the analysis of convergent validity, predictive validity, concurrent validity, discriminant validity and construct validity (Weinhart, Forsyth, Carey, Jaworski, & Durant, 1998). One difficulty with studies of self-report accuracy of sexual risk behavior is that there is no 'gold standard' against which self-reports can be compared. For many risk behaviors, such as drug use, there exist biological markers that can be used to validate the accuracy of a self-report. However, no unambiguous markers exist in the sexual domain that can serve as a standard for such behaviors as condom use and the number of sexual partners that an individual has. The present study describes an approach that offers a viable (though not perfect) standard. The strategy involves interviewing individuals once a week for a period of one year. Each week, the individual is asked to report the number of times he or she had sexual intercourse, the number of times that condoms were used, and the number of different sexual partners that he or she had. At selected points during the year (e.g. after six months), the individual is asked to recall his or her sexual behavior over the time period in question (e.g. recall the number of different sexual partners he or she had during the past six months). This 'long' recall is then compared against an aggregate of the weekly reports across the time period. The aggregated weekly reports serve as the 'standard' against which the long recall measures are validated.

Paradigm assumptions

One assumption of the above paradigm is that individuals can recall accurately global features of their sexual activity during the past seven days. This is likely for many but not all risk populations. National studies suggest that single, young, heterosexual adults engage in sexual intercourse about once or twice every week or two (Leigh, Temple, & Trocki, 1993). In such cases, it is reasonable to assume that individuals can accurately recall the number of times they have had sex, who the sex was with

and whether a condom was used. However, for risk populations that engage in sex with considerable frequency (e.g. prostitutes), the assumption of accurate recall during the past seven days is dubious. In these cases, shorter recall periods (e.g. one day) probably are necessary to ensure accurate recall. Empirical studies support the proposition that most people can recall their sexual activities over short time intervals, such as one week (see Jaccard, McDonald, Wan, Dittus, & Quinlan, 2002; Jaccard & Wan, 1995 for reviews).

A second assumption is that individuals provide truthful reports of their behavior. Evidence suggests that truthful reporting tends to occur if individuals do not have to provide their responses in a face-to-face interview, if they are provided proper orientating instructions that motivate them to give truthful answers and if they are assured of confidentiality (Turner, Miller, & Rogers, 1997). For example, in a pilot study for this research, we compared response distributions for weekly reports of sexual behavior for: (a) respondents who were given self-administered questionnaires with instructions that stressed honest responding and guaranteed confidentiality, with (b) a group of respondents given the same instructions but who were led to believe that they were being monitored with a lie detector test as they answered questions (see Jaccard & Wan, 1995, for details). Following traditional bogus-pipeline methodology, it was assumed that individuals who thought they were being physiologically monitored for lying were truthful in their responses. If the response distributions in the two groups are comparable, then this suggests that those in the standard instruction condition are being truthful. This was indeed the case. Thus, given complete confidentiality, incentive to respond truthfully and privacy of responding, most individuals will be truthful about their sexual activity.

A third assumption is that there are no testing effects, i.e. that participation in the weekly assessments does not bias either the sexual activity of the individual or the individual's ability to recall sexual activity on a long-term basis. Evidence for the absence of such testing effects on risk behavior has been provided in several studies, both in the sexual domain (Fujita, 1971; Kunin & Ames, 1981; Rubin &

Mitchell, 1976) and in other areas of study (Gersovitz, Madden, & Smickilas-Wright, 1978; Smith, Jobe, & Mingay, 1991). For example, Halpern, Udry and Suchindran (1991) found no evidence for monitoring effects when comparing adolescents who were monitored on a weekly basis for two years with non-monitored controls. The present study included similar control groups and also observed no significant testing effects (see Jaccard, McDonald, Wan, & Dittus, 1998).

In sum, the paradigm that we propose uses aggregated weekly reports of behavior as a standard to evaluate the accuracy of recall of behavior over an extended period of time. Although some critics may argue that we are merely comparing one self-report of behavior with another self-report of behavior, we believe that there is convincing empirical evidence that the weekly reports are accurate for the study populations that we focus upon and thus, can be used as a valid proxy for behavior. There may be some error in the standard, but the amount of error is probably sufficiently small that the aggregated weekly reports can be used as a basis for evaluating the accuracy of recall over extended periods of time.

The accuracy of reports about sexual partners

Jaccard et al. (1998; see also Jaccard et al., 2002) used the above paradigm to analyze the accuracy of self-reports of the frequency of sex, the use of condoms and the frequency of unprotected sex among a group of 285 unmarried, heterosexual young adults living in an urban setting. Overall, their analyses affirmed the relative accuracy of self-reports of these behaviors, although differences in accuracy were observed as a function of the time duration studied, the question format and individual difference variables. The present article uses this database to explore the accuracy of self-reports of the number of sexual partners over a given period of time.

There is an extensive literature on the psychological processes that individuals use when making frequency judgments (Blair & Burton, 1987; Bradburn, Rips, & Shevell, 1987; Gigerenzer, 1996; Menon, 1993, 1997; Menon, Raghurib, & Schwarz, 1995; Strube, 1987). An important

distinction in this literature is that between episodic strategies and rule-based strategies. With episodic strategies, when asked to recall the number of instances of a behavior over a given time period, the individual thinks about the relevant time interval and tries to recall each occurrence of the event in question, counting them up as each event is recalled. The nature of the judgment process is episodic in that the individual tries to recall specific episodes of the event. A second strategy is to use a rule-based judgment process. Here, the individual invokes a stored rule or algorithm from memory that is used to generate the requested frequency, without recourse to an analysis of specific episodes of the event. For example, if asked to judge the number of times that a condom was used during sexual intercourse over the past six months, the individual might invoke a rule that he always uses condoms and then report a frequency of condom use that is equal to the recalled number of instances of sexual intercourse. Other strategies involve a combination of episodic and rule-based processes. For example, when making a judgment of the frequency of sex during the past 12 months, the individual may think episodically of the number of times he or she had sex during the past month and then adopt a generating rule that multiplies this result by 12. One might cognitively invoke 'correction' factors to account for months when unusual events occurred (e.g. being away on vacation).

It seems likely that judgments of the number of sexual partners that one has had over a given time period are governed by several of these processes. For individuals who judge that they are sexually inactive over the time period in question, a rule-based heuristic is probably invoked, whereby they simply report zero sexual partners. For individuals who see themselves as monogamous and 'partner loyal', a report of a single sexual partner is dictated without any cognitive review of individual sexual episodes. For individuals who have more than one partner, an episodic-like strategy is probably used where the focus is on recalling distinct partners through time rather than each individual sexual episode. The cognitive processes in the former two cases are straightforward and probably lead to high degrees of accuracy if the individuals are accurate in

recalling their sexual inactivity or accurate in characterizing themselves as having a 'monogamous' orientation. For the case of multiple partners, the judgments are more complex, involving an accurate dating of the time interval in question, as well as recall of those episodes where new partners occurred in the sexual sequence. Because of this increased complexity, we predicted that accuracy of recall will be lower for individuals with multiple sexual partners as opposed to monogamous or sexually inactive individuals.

We also examined accuracy of recall as a function of different time periods. We focused on the accuracy of judgments over a period of 1 month, 3 months, 6 months and 12 months. For judgments of the number of sexual partners, it seems reasonable to predict that individuals will be more accurate for shorter as opposed to longer time durations due to the lowered cognitive demands of a shorter time interval on episodic-based processes. This would favor the use of short time durations over which self-reports should be obtained. However, a disadvantage of a short time frame is that it may be insensitive to general behavioral tendencies to engage in sex with many partners due to too narrow a window for individual differences in behavior to emerge (i.e. there will be range restriction). Studies suggest that individual difference variables tend to predict behavioral patterns over time as opposed to a single behavior at a single point in time (Ajzen & Fishbein, 1977; Jaccard & Wilson, 1991). Short time frames may be too limited to adequately sample such behavioral patterns. Thus, there probably is an optimal 'balancing' of recall accuracy (for shorter time periods) against behavioral representativeness (for longer time periods), suggesting that moderate length recall periods may be optimal for studies of sexual risk taking. Another purpose of the present study was to test this possibility.

Method

Respondents

Respondents were 285 males and females living in the Capital District area of Albany, New York. Respondents were selected so that they would be heterosexual, not living with a partner, single (unmarried, separated or divorced) and

who said they would be residing in the community for the coming 12 months. Only heterosexuals were included because of our experience with such populations and because of practical constraints that precluded including an adequate number of gay men and women. The sampling strategy was *not* designed to yield a representative sample of the community, but rather a relatively diverse sample that would permit us to test for differential accuracy in recall as a function of gender, age and education. Table 1 presents a demographic profile of the study respondents, which is roughly comparable to that of the general population in the study area, but with a slight overrepresentation of European Americans.

Respondents were recruited through initial phone interviews based on random samples from the community at large. Potential respondents were asked a few general opinion questions followed by background questions. From answers to the background questions, it was determined if the respondent met the conditions for the larger study and if so, a packet describing the study was sent, followed by a telephone call inviting the respondent to participate in the larger study. Our desire was to restrict participants to sexually active heterosexuals, but we could not explicitly screen on sexual activity because of the sensitivity of the questions that would need to be asked. We included a question that determined if the individual had dated a member of the opposite sex in the past six months and used this as a criterion for study inclusion. In the main study, we asked individuals their sexual orientation and only included in the analysis those individuals who indicated that they were heterosexual.

Procedures

Because of cost constraints, the weekly interviews were conducted via a mail survey. Respondents were sent a two-page self-administered questionnaire that arrived on Monday of each week, which they were to complete with respect to the previous seven days. In the small number of instances where a respondent did not return his or her survey in the provided envelope within two days, the individual was contacted by project staff and given a reminder. At months 1, 3, 6 and 12, respondents were sent a more detailed questionnaire that assessed

Table 1. Demographic characteristics of the sample

<i>Statistic</i>	<i>Value</i>	<i>Statistic</i>	<i>Value</i>
Percent females	56	Percent high school or less	15
Mean age	29	Percent with some college	26
		Percent college degree	39
Percent Caucasian	93	Percent advanced degree	20
Percent Hispanic	2		
Percent African American	3	Percent full-time employed	80
Percent other ethnicity	1	Percent part-time employed	10
		Percent unemployed	3
Percent never married	83	Percent other employment status	7
Percent separated	2		
Percent divorced	13	Median salary (in thousands)	27
Percent other marital status	2	Percent earning < \$10,000	9
		Percent earning < \$20,000	30
Percent Catholic	56	Percent earning > \$40,000	14
Percent Protestant	22		
Percent Jewish	4		
Percent other religion	19		

self-reports of their sexual activity since study inception and other variables. Respondents were not told initially that these longer questionnaires would be sent to them and they were paid an additional \$5 for their completion. Thus, they were not expecting the extended recall questions and were unaware of the fundamental purpose of the study. Rather, they were told that the purpose was to survey the sexual activities of people in the community.

Respondents were paid \$135 each to be in the study. Approximately 70 percent of the respondents who were asked to be in the study agreed to do so. Of the 70 percent who agreed to be in the study, 70 percent completed the full 12 months of study participation. Bias analyses were performed to isolate selection effects and bias due to attrition. Data from the phone surveys were compared for individuals who agreed to be in the study versus those who did not on the following variables: (1) open-ended responses to what the individual thinks is the biggest problem facing American youth today; (2) the median age at which teenagers should begin dating; (3) the percent who felt that people are dating less now because of AIDS; (4) religion; (5) the percent who consider themselves to be politically liberal; and (6) the level of education completed. No significant differences were observed between groups on these variables. Attrition analyses for those who remained in the study for the full 12 months

versus those who dropped out of the study examined group differences on the above variables as well as background data collected before the respondents left the study. These variables included age, gender, education, frequency of sexual intercourse over a 1-month period, condom use consistency over a 1-month period, religious affiliation, religiosity, ethnicity, age at first intercourse, 22 self-concept items and 45 items on attitudes toward sex. No meaningful bias was evident in these analyses (see Jaccard et al., 1998 for details).

Prior to the study, respondents were provided with a description of the tracking system to be used and methods that would ensure the confidentiality of their responses. Instructional sets were provided to create high motivation for honest responding. A measure of social desirability response tendency was obtained at the one-month recall and, consistent with our pilot research, showed no significant correlation with any of the major variables reported here.

Measures

For the weekly survey, respondents were asked to write a number in response to the question: 'How many different sexual partners did you have this week?' It was made clear to respondents that a sexual partner was being defined as someone they had sexual intercourse with. Respondents were then asked to write the initials of each partner in a designated area and

to answer a few short questions about that partner. The initials and questions were used to identify the partner from one week to the next and to determine the number of different partners the individual had over the course of the time period in question. For the long recall sessions, respondents were provided instructions orienting them to the time interval and then were asked (illustrated here using the six-month time period): 'How many different sexual partners did you have during the past six months?' These formats are typical of measures used in behavioral studies of sexual risk behavior.

In addition, demographic and attitudinal variables were assessed at each long recall session. For analyses of individual difference variables, we focused on three constructs that should be related to the number of sexual partners an individual has. These were a measure of sensation seeking/thrill seeking as measured by agreement (on a five-point rating scale) with two items (I am a thrill seeker; I take more risks than most people), the propensity to engage in casual sex as measured by agreement or disagreement (on a five-point rating scale) with three items (I only have sex with someone I know pretty well; I have to be in a love relationship to have sex; I can enjoy sex even if I don't love the person in a committed way) and the perceived susceptibility to a sexually transmitted disease as measured by agreement or disagreement (on a five-point rating scale) of two items (Given my current lifestyle and habits, I run a risk of getting AIDS; Given my current lifestyle and habits, I run a risk of getting herpes). Higher scores indicated greater agreement with the statement. Items within a measure were summed to yield an overall score, with reverse scoring of the first two items of the propensity for casual sex. All items within a measure were highly correlated (minimum $r = 0.55$).

Results

Throughout this section, the term 'self-report' will refer to the number of partners reported at the long recall sessions and the term 'behavior' will refer to the behavioral index based on the weekly aggregated measures. Because of the non-normal distributions that were prevalent in the data, estimation of standard errors relied on

the bootstrapping methods discussed by Wilcox (1997). Detailed analysis of the accuracy of reports of sexual intercourse and condom use for this sample are presented in Jaccard et al. (2002). The present study focuses on the accuracy of self-reports of the number of sexual partners.

Methodological checks

A number of methodological issues were addressed, including the analysis of missing data patterns, attrition bias and testing effects. The results of these analyses are summarized briefly here and reported in detail in Jaccard et al. (1998). In any given week, approximately 2 percent of the sample failed to complete their weekly survey. Data values were imputed based on reported activity during the two weeks prior to and the two weeks after the missing survey. Analyses revealed no significant bias in the pattern of missing data as a function of a wide range of individual difference variables, suggesting that it was missing at random. Attrition bias and testing effects also were minimal. For example, a control group was included that completed the four long recall surveys at months 1, 3, 6 and 12 but who did not participate in the weekly monitoring phase of the study. These individuals did not differ significantly from respondents in the main study in their mean number of partners, mean frequency of sex and mean use of condoms. There also was no evidence that participation in the weekly surveys increased or decreased the accuracy of recall (Jaccard et al., 1998).

Number of sexual partners

The frequency distribution for the number of sexual partners at each of the four time periods is presented in Table 2. We focus comments on the 12-month time interval (second to last column), which is a commonly used interval in research. The majority of the sample (52% or $N = 141$) had only one sexual partner over the course of the 12 months, and 16 percent of the sample ($N = 45$) did not engage in sexual intercourse at all. About a third of the sample had multiple sexual partners, with nearly 8 percent having four or more sexual partners. Although these estimates suggest a population that is at low risk, the distribution actually reflects distributions observed in national studies of the

Table 2. Frequency distribution for number of sexual partners recalled and percent of accurate recall within the number of partners recalled

Number of partners recalled	One month		Three months		Six months		Twelve months	
	Frequency	Percent accurate	Frequency	Percent accurate	Frequency	Percent accurate	Frequency	Percent accurate
0	96	93.8	94	87.8	74	78.4	45	86.7
1	146	94.5	158	96.7	148	95.3	141	86.5
2	14	50.0	16	68.8	21	52.4	40	67.5
3	3	66.7	10	40.0	12	41.7	23	34.8
4	2	50.0	3	66.7	6	16.7	10	40.0
5	0	-	0	-	1	0.0	5	40.0
6	0	-	3	0.0	1	100.0	0	-
7	0	-	1	0.0	1	0.0	2	100.0
8	0	-	0	-	2	0.0	0	-
10	0	-	0	-	0	-	1	0.0
12	0	-	0	-	0	-	1	0.0
15	0	-	0	-	0	-	1	0.0

number of sexual partners of heterosexual, young adults (Michael, 1994). We believe it is important to study the accuracy of self-reports across the full spectrum of risk activity, including those who report relatively few sexual partners as well as those who report larger numbers of sexual partners. Social scientists often construct and test theories of risk behavior by comparing high-risk individuals with low-risk individuals. If the self-reports of low-risk individuals are inaccurate, then this can be just as damaging to effective theory building and theory testing as when self-reports of high-risk individuals are distorted. Our sample exhibits a reasonable range of both low-risk and high-risk individuals and thus is useful for studying issues in recall accuracy.

A notable characteristic of the sample was that it was primarily composed of European Americans. It is important to note that sexually transmitted diseases are common in such populations. Chlamydia, human papillomavirus (HPV) and herpes are widespread across all racial and ethnic groups (Centers for Disease Control and Prevention (CDC), 2002; Mertz, Trees, Levine, Lewis, Lichfield, & Pettus, 1998). Prevalence rates of HPV for European American women under the age of 25 is about 1 of every 3 or 4 women (Burk, Ho, Beardsley, Lempa, Peters, & Bierman, 1996). Just under 1 in 5 European American young adults in the United States are infected with herpes (CDC, 2002). Over the course of the 12-month study period, our sample reported engaging in a total of 17,156 acts of sexual intercourse, of which 13,223 were not protected by a condom. If one excludes those who did not engage in sexual intercourse, the average number of unprotected acts of sexual intercourse (in terms of condom use) over the year was approximately 56, indicating a sample clearly at risk for STDs. Thus, although the study population does not represent the traditional categories of high-risk individuals as reflected by being minority or being prostitutes, the sample does represent a population that characteristically has high rates of STDs and that is engaging in risk behavior.

Accuracy of claimed abstainers

At months 1, 3, 6 and 12, a non-trivial minority of the sample claimed not to have had any sexual partners since study inception. The

percentage of individuals making this assertion was 37 percent, 31 percent, 28 percent and 17 percent at each of the 4 time periods, respectively. Of those people who said they had zero sexual partners, between 5 percent to 20 percent actually did have a sexual partner, depending on the length of the recall period (see Table 2). The accuracy rate was best at one month and worst at six months with intermediate values for three and twelve-month recall periods. Statistical tests of the accuracy differences showed a statistically significant effect ($p < 0.05$) only for the one-month versus six-month comparison, using a modified Bonferroni method to control experiment-wise error rates (Jaccard, 1998).

Almost all of the respondents who gave inaccurate self-reports under-reported the number of sexual partners by one. We defined a dichotomous variable for accuracy (provided accurate response versus those that did not) and correlated this measure with gender, age and education. There was a reasonable correlation between education and the tendency to provide an accurate response for the 3, 6 and 12-month periods. The correlations were 0.28, 0.41, 0.48 and 0.54 for the 1, 3, 6 and 12-month periods, respectively (all $p < 0.01$). Thus, self-reported abstinent individuals with lower levels of education seemed to have greater difficulty recalling their sexual partners at the longer time intervals.

In sum, claimed abstainers were relatively accurate in characterizing their (lack of) sexual activity, although at the longer timer intervals there was a non-trivial minority who seemed to forget about a sexual partner they had had. This tended to be truer of individuals with lower levels of education.

Accuracy of claimed monogamists

A large segment of the sample claimed to be monogamous. The percentage of such individuals at the recall periods of 1 month, 3 months, 6 months and 12 months was 55.9, 55.4, 55.6 and 52.4, respectively. In general, the accuracy of recall for such individuals was high, with nearly 95 percent accurately characterizing their monogamy. The one exception was the 12-month recall period, which yielded levels closer to 85 percent accuracy (see Table 2). The 12-month accuracy rate was statistically significantly

($p < 0.05$) lower than the rates for the other time periods.

Almost all errors were within plus or minus one partner. There was no significant correlation between age, gender, education and the tendency to provide an accurate response for claimed monogamists. There was, however, a slight tendency toward inaccurate responses by those who expressed a propensity for casual sex. The correlations between the propensity for casual sex and accuracy were -0.18 , -0.16 , -0.07 and -0.17 at months 1, 3, 6 and 12, respectively. The correlation at the six-month period was not statistically significant, but all others were ($p < 0.05$).

In sum, most individuals who claimed to be monogamous were, in fact, monogamous. The claims tended to be slightly less accurate for the 12-month recall period and somewhat less accurate for those with a propensity to engage in casual sex.

Accuracy of individuals reporting multiple partners

Table 2 shows that the accuracy rates tended to be lower for individuals who reported having multiple partners as opposed to claimed abstainers and claimed monogamists, with only about 50 percent of those individuals reporting having multiple sexual partners providing accurate self-reports (as opposed to accuracy rates in the 90 percent range for claimed abstainers and claimed monogamists). The first column of Table 3 presents the mean absolute discrepancy between the reported number of sexual partners and the actual number of sexual partners. At one month, self-reports were, on average, inaccurate by approximately half a sexual partner, at three months self-reports were inaccurate by approximately two-thirds of a sexual partner, at six months the inaccuracy was by approximately three-fourths a sexual partner and for twelve months the inaccuracy was by approximately one full partner. Table 3 also presents the Pearson correlation between self-reports and behavior and a percentage bend correlation (which is an outlier resistant index of correlation, see Wilcox, 1997). All correlations were moderate in magnitude. The last column of Table 3 presents the percentage of individuals whose self-report was within one partner of their actual behavior.

Table 3. Indices of self-report-behavior correspondence for individuals reporting multiple partners

	<i>Mean absolute discrepancy</i>	<i>Pearson correlation</i>	<i>Percentage bend correlation</i>	<i>Percent within one partner</i>
1 month	0.53*	0.57*	0.50*	94.7
3 months	0.67*	0.56*	0.70*	90.9
6 months	0.75*	0.81*	0.79*	86.4
12 months	0.95*	0.59*	0.69*	81.9

* $p < 0.05$

Analysis of slopes If there is a one-to-one correspondence between self-reports and behavior, then a change of a single unit in one variable should be associated with a change of a single unit in the other variable. Table 4 presents the mean number of partners that individuals actually had as a function of the reported number of partners. This table provides perspectives on how changes in the number of partners recalled corresponds to changes in the actual number of partners. For example at month 1, when the number of partners reported changed from 0 to 1, the actual number of partners changed, on average, 0.99 units (from 0.06 to 1.05). When the number of partners reported changed from 1 to 2, the actual number of partners changed, on average, 0.60 units (1.64–1.05). Care must be taken when interpreting mean changes at the higher numbers of reported sexual partners because the means are based on small sample sizes. Table 5 reverses the predictor and criterion and presents the mean number of reported sexual partners as a function of the actual number of partners the individual had. Both tables show clear ordinal trends between

the variables: as the number of actual partners increases, so does the recalled number of partners and as the number of recalled partners increases, so does the number of actual partners (with small sample size exceptions occurring within the 3, 6 and 12-month intervals). However, there also is some evidence for non-interval level properties in the recall measures at the upper end of the scale (i.e. for multiple partners).

Correlates of accuracy Because claimed abstainers and monogamists had higher accuracy rates than respondents claiming multiple partners, any variable that is related to the discrimination of the former two groups from the latter group will show a spurious relationship with self-report accuracy. We previously reported individual difference analyses in accuracy for claimed abstainers and claimed monogamists. The present analyses focus on individual differences in accuracy only for respondents who claimed multiple sexual partners. We first calculated an absolute discrepancy score between the self-reported number of

Table 4. Mean number of actual sexual partners as a function of number of partners recalled

<i>Number of partners recalled</i>	<i>1-month mean</i>	<i>3-month mean</i>	<i>6-month mean</i>	<i>12-month mean</i>
0	0.06 (96)	0.16 (82)	0.30 (74)	0.18 (45)
1	1.05 (146)	1.02 (152)	1.05 (148)	1.15 (141)
2	1.64 (14)	2.00 (16)	2.00 (21)	2.30 (40)
3	2.67 (3)	3.20 (10)	3.67 (12)	3.65 (23)
4	3.00 (2)	4.33 (3)	4.50 (6)	5.10 (10)
5	–	–	6.00 (1)	5.80 (5)
6	–	5.00 (3)	6.00 (1)	–
7	–	3.00 (1)	8.00 (1)	7.00 (2)
8	–	–	7.00 (2)	–
10	–	–	–	3.00 (1)
12	–	–	–	9.00 (1)
15	–	–	–	8.00 (1)

(Sample sizes are in parentheses)

Table 5. Mean number of recalled sexual partners as a function of actual number of sexual partners

Number of partners	1-month mean	3-month mean	6-month mean	12-month mean
0	0.01 (91)	0.01 (73)	0.03 (59)	0.00 (39)
1	1.00 (150)	0.97 (158)	0.96 (156)	0.99 (130)
2	1.80 (15)	1.78 (18)	1.41 (22)	1.73 (48)
3	2.25 (4)	3.38 (8)	2.46 (13)	2.67 (21)
4	4.00 (1)	3.14 (7)	3.00 (5)	3.20 (15)
5	–	5.00 (2)	3.40 (5)	4.20 (5)
6	–	–	5.50 (2)	4.00 (1)
7	–	6.00 (1)	6.67 (3)	5.20 (5)
8	–	–	7.00 (1)	10.00 (2)
9	–	–	–	8.00 (2)
11	–	–	–	3.00 (1)

(Sample sizes are in parentheses)

sexual partners and the actual number of sexual partners. We then correlated this with gender, education, age and frequency of sex and observed no significant differences in accuracy as a function of these variables. We next examined if the absolute discrepancy score was associated with thrill seeking, the propensity for casual sex and the perceived susceptibility to STDs. The only statistically significant correlations were with the propensity for casual sex. The correlations were 0.52, 0.23, 0.21 and 0.39 at months 1, 3, 6 and 12 respectively. Only the correlations at months 1 and 12 were statistically significant ($p < 0.05$). In general, individuals with a higher propensity to engage in casual sex were less accurate in their self-reports.

Aggregate level analyses of accuracy

Moderate levels of accuracy at the individual level do not necessarily mean that measures will be inaccurate at the aggregate level. It may be the case that the number of individuals who over-report their number of sexual partners is balanced by the number of individuals who under-report their number of sexual partners, so that the mean number of reported sexual partners is close to the actual mean number of sexual partners. Table 6 presents measures of central tendency and variability for each of the time periods, focusing only on individuals who claimed to have multiple partners. A bias index is provided which reflects the percent by which the self-report overestimated or underestimated behavior. For the measures of central tendency, the one-month recall period tended to produce the largest bias. For the most part, the self-report central tendency for all time periods was

within \pm a third of a sexual partner of the behavioral central tendency, with the bias being smallest for the three-month period.

Sensitivity to individual difference variables

To evaluate sensitivity to individual difference variables, each of the self-reports and the measures of behavior were correlated with thrill seeking, the propensity for casual sex and perceived susceptibility to sexually transmitted diseases. Because most studies of sexual partners focus only on sexually active individuals, the correlations were computed excluding individuals who reported they were sexually inactive. Table 7 presents the correlations. All of the individual difference variables were statistically significantly correlated with both the reported and the actual number of sexual partners at months 3, 6 and 12. However, the correlations tended to be lower at one month and were non-significant for thrill seeking and perceived susceptibility to a STD.

Discussion

Self-reports of the number of sexual partners an individual has had over the course of a given time period are central to investigations of sexual risk behavior. The accuracy of such reports is an important issue. The present study revealed a somewhat mixed picture in terms of accuracy rates. For individuals who reported abstinence or who claimed to be monogamous, accuracy was quite high, especially at shorter time frames. For individuals who reported having multiple sexual partners, the

Table 6. Analysis of selected measures of location and scale for subjects with more than one partner

	<i>Mean</i>	<i>Median</i>	<i>Trimmed mean</i>	<i>SD</i>
Behavior 1	1.79	1.88	1.69	0.92
Self-report	2.32	2.00	2.08	0.67
Difference	0.53*	0.12	0.39*	-0.25
Bias	29%	6%	23%	-37%
Behavior 3	2.75	2.50	2.65	1.11
Self-report	2.91	2.50	2.50	1.30
Difference	0.16	0.0	-0.15	-0.19
Bias	6%	0%	-6%	17%
Behavior 6	3.21	2.91	2.92	1.79
Self-report	2.96	2.50	2.54	1.40
Difference	-0.25	-0.41*	-0.38	-0.39*
Bias	-8%	-16%	-15%	-28%
Behavior 12	3.31	2.89	2.86	1.98
Self-report	3.02	2.55	2.55	1.68
Difference	-0.29	-0.34	-0.31	-0.30
Bias	-10%	-13%	-12%	-16%

Trimmed mean is 20 per cent trimmed; ‘Difference’ is the self-report minus the behavior at the month indicated numerically after the behavior label; ‘Bias’ is $[100 * (\text{self-report measure}/\text{behavior})] - 100$; * indicates difference between self-report and behavior is statistically significant, $p < 0.05$

Table 7. Correlations with individual difference variables

	<i>Thrill seeking</i>	<i>Casual sex propensity</i>	<i>Perceived susceptibility</i>
Behavior 1	0.15	0.18*	0.04
Self-report 1	0.23*	0.26*	0.12
Behavior 3	0.28*	0.30*	0.21*
Self-report 3	0.30*	0.37*	0.19*
Behavior 6	0.27*	0.34*	0.25*
Self-report 6	0.30*	0.40*	0.21*
Behavior 12	0.28*	0.37*	0.32*
Self-report 12	0.38*	0.40*	0.38*

* $p < 0.05$

self-reports tended to be within plus or minus one partner of actual behavior and this was especially true at the shorter durations. Overall, these data suggest a reasonable degree of accuracy for the self-reports. On the other hand, accuracy rates tended to lessen for individuals who claimed to have multiple sexual partners and this was more true of individuals who expressed an attitudinal propensity for casual sex. The correlations between self-reports and behavior were about 0.70 for individuals who reported having multiple partners, except for the 1-month time frame, where the correlation was closer to 0.50 (probably due to a restricted range).

Choice of time frames

The amount of absolute error in self-reports tended to be lowest for the one-month time frame, but the benefits of this were offset by the fact that there was much less variability in the number of sexual partners. This reduced variability attenuated correlations with other variables. In addition, the restricted range makes the index less forgiving of error in terms of correctly ranking individuals along the dimension of number of sex partners.

The patterning of correlations with individual difference variables and the patterning of bias in the aggregate level measures tended to favor the use of the longer time intervals when studying

attitudinal and personality correlates of the number of sexual partners. This result is consistent with theoretical analyses that suggest that general individual difference variables are more likely to be correlated with behavioral patterns over extended periods of time rather than specific behaviors over relatively short time intervals (Ajzen & Fishbein, 1977; Jaccard & Wilson, 1991).

Taken together, these results suggest that applications in the behavioral sciences probably should not use a one-month time frame when assessing the tendency to engage in sex with multiple partners, at least for populations similar to those studied here. This may not be true of other risk populations where the frequency of sex is higher.

In terms of the three, six and twelve-month time periods, no one time frame shows consistently superior performance to the other. The amount of absolute error decreases with the shorter time frame (see Table 3), but this is accompanied by more restricted ranges of variability. Correlations between self-reports and actual behavior tended to be highest for the six-month time interval when analyzing individuals claiming multiple sexual partners, but the six-month time frame also revealed lower levels of accuracy for claimed abstainers. Thus, it is difficult to unequivocally recommend a duration of three versus six versus twelve months when assessing the number of sexual partners an individual has. Perhaps a strategy that probes all three durations and then uses the relative strengths of each is best.

Aggregate level analyses

Many studies are designed to determine both the short-term and long-term effects of an intervention. Most such studies compare the mean number of sexual partners between an experimental group and a control group. One concern of methodologists is whether self-reports of sexual risk behavior can be accurate over a substantial period of time following the intervention. The present data suggest that this is probably possible, as the amount of bias in the measures of central tendency was relatively small for longer time frames (see Table 6). Even if bias were present, this would not be problematic if the bias does not operate differentially in the two groups. For example, if interest is in

comparing mean shifts in the number of sexual partners for an experimental and control group, and self-reports tend to underestimate actual behavior by a factor of 10 percent, then this bias will be present in both groups and should not affect the estimate of the mean *difference* between groups. It is somewhat reassuring that mean self-reports of the number of sexual partners map fairly closely onto the mean actual number of sexual partners, even over extended periods of time.

Individual differences in self-report accuracy

We observed three notable trends with respect to individual differences in the accuracy of self-reports of the number of sexual partners. The first was the tendency for individuals who report multiple partners to be less accurate in their reports than claimed abstainers or claimed monogamists. This result was consistent with our predictions and probably represents greater cognitive demands of the recall task for individuals who have multiple partners. The second trend was the tendency for claimed abstainers to show lower accuracy as a function of education level. At the longer time frames, individuals with lower education were more likely to forget or fail to report that they did, in fact, have a sexual partner over the course of the time frame. Although this may reflect a general tendency for individuals with lower education to have difficulty with longer time durations, the educational differences did not emerge for claimed monogamists or individuals claiming to have multiple partners. Thus, there appears to be something operating that is unique to claimed abstainers. Future research needs to replicate and explore this potentially interesting dynamic.

The third trend was the tendency for lower levels of accuracy for individuals who express an attitudinal propensity for casual sex. This result is somewhat troubling because it is individuals who are prone to casual sex that are of most interest when studying the transmission of STDs. We suspect that those with higher propensities for casual sex are more likely to have multiple partners and that individuals with more partners are more likely to be subject to recall error, hence the observed association. Future research needs to explore this dynamic.

Psychologists have shown that recall of events is improved if the individual is personally involved in the events at the time that information is encoded and stored in memory (Smyth, Morris, Levy, & Ellis, 1987). Individuals with a propensity for casual sex are probably less psychologically invested with their partners, hence the lower levels of recall. To maximize accurate recall, psychologists have stressed the importance of matching cues in the recall questions to contextual cues that were present when the individual encoded the event in question (Smyth et al., 1987). This suggests that accuracy of recall might be improved by conducting careful studies of the contexts that lead to sexual intercourse on the part of those with a propensity toward casual sex and then incorporating these cues into the structuring of the recall questions. The present research used an assessment strategy that was simple and devoid of contextual cues, but which is quite common in existing behavioral research. We suspect that a better strategy will be one in which the researcher probes the number of sexual partners an individual has had in multiple ways using more than one question and then derives an estimate of the true number of sexual partners using a weighted combination of the multiple questions. Different questions would emphasize different contextual cues and time frames, with the weights that are ultimately applied being determined by individual difference variables such as the individual's education and propensity for casual sex.

Scaling properties

Tables 4 and 5 suggest that the self-report measure of the number of sexual partners does not have interval level characteristics. If the measure was interval level, then a one-unit change in the measure would be associated with, on average, a constant change in behavior no matter where on the scale the unit change occurred. Care must be taken in interpreting Tables 4 and 5 because of the small sample sizes upon which some of the means are based. However, at the longer time durations in the range between zero and four partners, non-intervalness is suggested even where the sample sizes are larger. This finding does not mean that parametric statistics (which assume interval level properties) cannot be applied to such data.

The degree of non-intervalness observed may not be extreme enough to cause problems. But researchers should keep in mind that a one-unit change in the reported number of sexual partners may reflect differing amounts of true change depending on the range of the scale.

Study limitations

It is important to keep in mind the limitations of our research when drawing conclusions about the validity of self-reports. The sample was primarily European American in a relatively large metropolitan area. It is likely that extremely high-risk individuals (e.g. prostitutes) were not represented in the sample. Although the sample is probably a reasonable reflection of young, single, white adults in metropolitan areas, generalizations cannot be extended to older populations or adolescents nor to populations that are solely characterized by high-risk behavior. We suspect some of the same trends we observed will be evident in these populations, but care must be taken when generalizing.

References

- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical literature. *Psychological Bulletin*, *84*(5), 888-918.
- Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, *14*, 280-288.
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*, 157-161.
- Burk, R. D., Ho, G. Y., Beardsley, L., Lempa, M., Peters, M., & Bierman, R. (1996). Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women. *Journal of Infectious Diseases*, *174*(4), 679-689.
- Centers for Disease Control and Prevention (CDC). (2002). *Tracking the hidden epidemics: Trends in the United States, 2000*. Atlanta, GA: Centers for Disease Control and Prevention.
- Fujita, B. (1971). The effects of an interview on attitudes and behavior. *Journal of Sex Research*, *7*, 138-152.
- Gersovitz, M., Madden, J., & Smickilas-Wright, H. (1978). Validity of the 24-hour dietary recall and

- seven-day record for group comparisons. *Journal of the American Dietetic Association*, 73(1), 48–55.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, 103(3), 592–596.
- Halpern, C., Udry, R., & Suchindran, C. (1991). Effects of repeated questionnaire administration in longitudinal studies of adolescent male sexual behavior. Unpublished manuscript, Carolina Population Center, University of North Carolina.
- Jaccard, J. (1998). *Interaction effects in factorial analysis of variance*. Newbury Park, CA: Sage Publications.
- Jaccard, J., McDonald, R., Wan, C., & Dittus, P. (1998). Methodological and measurement issues in the study of self report accuracy of sexual risk behaviors. Department of Psychology, University at Albany, State University of New York.
- Jaccard, J., McDonald, R., Wan, C., Dittus, P., & Quinlan, S. (2002). The accuracy of self reports of condom use and sexual behavior. *Journal of Applied Social Psychology*, 32(9), 1863–1905.
- Jaccard, J., & Wan, C. (1995). A paradigm for studying the accuracy of self reports of risk behavior relevant to AIDS: Empirical perspectives on stability, recall bias and transitory influences. *Journal of Applied Social Psychology*, 25, 1831–1858.
- Jaccard, J., & Wilson, T. (1991). Personality factors influencing risk behavior. In J. Wasserheit, S. Aral, & K. Holmes (Eds.), *Research in human behavior and sexually transmitted diseases in the AIDS era*. Washington, DC: American Society for Microbiology.
- Kunin, C. N., & Ames, R. (1981). Methods for determining the frequency of sexual intercourse and activities of daily living in young women. *American Journal of Epidemiology*, 113(1), 55–61.
- Leigh, B. C., Temple, M., & Trocki, K. (1993). The sexual behavior of U.S. adults: Results from a national survey. *American Journal of Public Health*, 83(10), 1400–1408.
- Menon, G. (1993). The effects of accessibility of information in memory on judgments of behavioural frequencies. *Journal of Consumer Research*, 20(3), 431–440.
- Menon, G. (1997). Are the parts better than the whole? The effects of decompositional questions on judgments of frequent behaviors. *Journal of Marketing Research*, 34, 335–346.
- Menon, G., Raghurir, P., & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility-diagnostics framework. *Journal of Consumer Research*, 22(2), 212–228.
- Mertz, K. J., Trees, D., Levine, W. C., Lewis, J. S., Lichfield, B., & Pettus, K. S. (1998). Etiology of genital ulcers and prevalence of human immuno-deficiency virus co-infection in ten U.S. cities. *Journal of Infectious Diseases*, 178(6), 1795–1798.
- Michael, R. (1994). *Sex in America: A definitive survey*. Boston, MA: Little, Brown.
- Rubin, Z., & Mitchell, C. (1976). Couples research as couples counseling: Some unintended effects of studying close relationships. *American Psychologist*, 31, 17–25.
- Smith, A. F., Jobe, J. B., & Mingay, D. M. (1991). Retrieval from memory of dietary information. *Applied Cognitive Psychology*, 5, 269–296.
- Smyth, M., Morris, P., Levy, P., & Ellis, A. (1987). *Cognition in action*. Hillsdale, NJ: Erlbaum.
- Strube, G. (1987). Answering survey questions: The role of memory. In H. J. Hippler, N. Schwartz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 86–101). New York: Springer-Verlag.
- Turner, C., Miller, H., & Rogers, S. (1997). Survey measurement of sexual behavior: Problems and progress. In J. Bancroft (Ed.), *Researching sexual behavior: Methodological issues* (pp. 237–260). Bloomington, IN: Indiana University Press.
- Weinhart, L. S., Forsyth, M., Carey, B. C., Jaworski, J., & Durant, L. E. (1998). Reliability and validity of self report measures of HIV related sexual behavior: Progress since 1990 and recommendations for research and practice. *Archives of Sexual Behavior*, 27(2), 155–180.
- Wilcox, R. (1997). *Introduction to robust estimation and testing*. New York: Academic Press.