Inconsistencies in Adolescent Self-Reported Sexual Behavior: Experience from Four Randomized Controlled Trials

Jill R. Glassman[1] · Elizabeth R. Baumler[2] · Karin K. Coyle[3]

Abstract

To examine the degree, correlates, and implications of inconsistent self-report data on sexual risk behaviors of adolescents. We analyzed data from four longitudinal group-randomized controlled trials of evidence-based HIV/STI/pregnancy prevention programs in Texas and California from 2000 to 2010. Across- and within-time logical inconsistencies in sexual behavior survey responses were analyzed using multilevel logistic regression. Rates of any inconsistencies ranged from 12 to 18% across the four trials. In all trials, rates were higher in males than in females. Age, normative beliefs, and race/ethnicity were most strongly associated with inconsistencies. We found substantial rates of inconsistencies in adolescents' self-reports of their sexual behavior, which did not occur at random. Studies should routinely report observed rates of inconsistencies and methods used to adjust for them so that any biases in the population to which the study generalizes are understood by public health practitioners and policy-makers looking to adopt programs for their particular population.

Keywords Self-report sexual risk behavior · Inconsistent reporting · Measurement

Self-report measures of adolescent health risk behaviors are essential for public health surveillance, intervention and policy development, and evaluating program effectiveness (McFarlane & St. Lawrence, 1999; Upchurch et al., 2002; Brener et al., 2003; McAuliffe et al., 2007; Palen et al., 2008; Rose et al., 2009; Hamilton & Morris, 2010; Brown et al., 2012). Yet, the accuracy of self-reports by adolescents of their own health risk behaviors is theorized to depend heavily on a variety of factors, including (1) ability to understand the questions, (2) ability to recall behaviors, (3) perceived social desirability of certain behaviors, (4) motivation to complete surveys accurately, and (5) salience of the behavior to current/changing identity (Rodgers et al., 1982; Rosenbaum, 2006, 2009; Palen et al., 2008; Brener

et al., 2003; Beguy et al., 2009; Napper et al., 2010). This may be especially true for sensitive sexual risk behaviors for which effects of social desirability and identity salience may be strong (Brener et al., 2003; Newcomer et al., 1988; O'Sullivan, 2008; Rodgers et al., 1982; Rosenbaum, 2006, 2009).

One of the particular challenges of measuring sexual risk behaviors in adolescents is the lack of objective measures against which to validate self-report responses (Brown et al., 2012; McAuliffe et al., 2007; Palen et al., 2008; Upchurch et al., 2002). A handful of measurement studies involving females have used biomarker testing (Rose et al., 2009; Rosenbaum et al., 2017; DiClemente et al., 2013; DiClemente, 2016), but this method is resource-intensive and has ethical and practical barriers, and therefore is not widely used. Thus, researchers relying on self-report surveys to collect data on teen sexual behavior continue to depend on survey administration strategies (e.g., restricting recall periods, using built-in skip patterns) as well as building in validity checks to examine inconsistencies in responses to repeated or related questions within and across time (Brown et al., 2012; Weinhardt et al., 1998).

Unfortunately, inconsistency rates are not routinely reported in published surveillance and program effectiveness studies, possibly due to space limitations. However, a

✉ Elizabeth R. Baumler
elbaumle@utmb.edu

[1] Clinical Excellence Research Center (CERC), Stanford University School of Medicine, 291 Campus Drive, Stanford, CA, USA

[2] Center for Violence Prevention, The University of Texas Medical Branch, 301 University Blvd, Galveston, TX, USA

[3] Education Training and Associates, 100 Enterprise Way, Suite G300, Scotts Valley, CA, USA

number of studies have examined inconsistencies in reports of sexual behaviors across time using test–retest data. In particular, data from the Youth Risk Behavior Survey (YRBS), the National Longitudinal Study of Adolescent Health (Add Health), and other surveys administered reported inconsistency rates ranging from 6 to 15% for variables representing "ever had sex" and age of first intercourse. (Brener et al., 2002; Rosenbaum, 2006, 2009; Upchurch et al., 2002).

A handful of studies further examined variability across demographic factors in the rates of inconsistency of adolescent self-reported sexual behavior with mixed results. Rodgers et al. (1982) found that black males had higher rates of inconsistent reporting on lifetime sexual experience, but Alexander et al. (1993) found that black males had lower relative rates of inconsistent reporting on this behavior. Upchurch et al. (2002) found that inconsistencies across time in reports of sexual behaviors on the Add Health survey varied by gender and ethnicity, but in another analysis of these data, Rosenbaum (2006) found this was no longer the case after controlling for religious factors. Brener et al. (1995) found an association between age and inconsistent reporting. Several studies found inconsistencies were higher in males than in females (Beguy et al., 2009; Tenkorang, 2021). Brown et al. (2012) found higher inconsistencies among black female adolescents.

Variations in rates of inconsistencies by demographic and other risk factors may bias prevalence estimates (Siddiqui et al., 1999) and impact the generalizability of program evaluation findings. For example, if the most high-risk youth tend to have the highest rates of inconsistencies and their observations are removed, estimates of health risk behaviors may be biased toward more "normative behaviors" (Ramos et al., 2017). Furthermore, conclusions about program effectiveness based on a sample excluding inconsistent responses may not apply to those youth most at risk of HIV/STI or pregnancy. More information is needed to fill in the gaps about variations in inconsistencies across key risk and protective factors—especially those that are markers for higher risk such as susceptibility to peer norms. This information is critical for continuing to guide future research to improve the measurement of teen sexual risk, intervention evaluation (e.g., sample size) planning, data cleaning, and other analysis strategies to reduce and account for inaccuracies resulting from inconsistent responses.

The goals of the present study were to update and add new information to the literature on rates and previously unexplored correlates of inconsistent reporting of sexual risk behaviors among adolescents. We accomplished these goals by examining the variation in a variety of different types of inconsistencies within and across demographic and risk subgroups from four randomized controlled trials of adolescent HIV/STI/pregnancy prevention studies conducted in two geographic locations in the USA. In some cases, we were able to pool data across studies to enable examination of the relative contribution of multiple

correlates at once. We further sought to understand how this information might be used to inform strategies for optimal study design and analysis procedures that maximize power and minimize bias.

Methods

We conducted secondary analyses of data from four previous group-randomized controlled trials (GRTs) conducted in Texas and California from 2000 to 2010 (Table 1). All four studies were federally funded (CDC or NIH) in the area of HIV/STI and pregnancy prevention in middle and high school youth. All studies collected student self-report survey data on sexual behaviors including vaginal sexual intercourse, number of partners with whom had vaginal sex, condom use, birth control use, and substance use before/during vaginal sex. All four studies focused on populations of youth that experience disparities in sexual health outcomes, such as sexually transmitted infections (STIs) and/or unplanned pregnancies. Further details of the studies can be found in the original manuscripts describing their primary evaluation results (Coyle et al., 2006, 2013; Markham et al., 2012; Tortolero et al., 2010).

For all datasets, both within-time and across-time inconsistencies in self-reports were calculated overall and by subgroups. Factors known to be associated with sexual risk behaviors were examined as potential correlates of inconsistency using multilevel regression analyses in which the outcome variables had inconsistencies or were free of inconsistencies. Given the heterogeneity in study samples, analyses were conducted separately for each study dataset.

Measures

Definitions of Inconsistencies

Within-Time Inconsistency

Five sexual behaviors commonly measured in surveillance and evaluation studies of adolescent sexual risk behaviors were used to assess within-time inconsistent cases: lifetime sexual experience (ever had vaginal sex), number of sexual partners with whom had vaginal sex, frequency of vaginal sex, number of sexual partners with whom had vaginal sex without a condom, and frequency of vaginal sex without a condom (all within past 3 months except lifetime sex). In particular, an inconsistent case was defined as one in which any of the following were true: (1) number of partners with whom had vaginal sex exceeded the number of times with whom had vaginal sex, (2) number of partners with whom had vaginal sex without a condom exceeded the number of

Table 1 Overview of four previous group-randomized controlled trials (GRTs) conducted in Texas and California from 2000 to 2010

	It's Your Game	All About Youth	All4You!	All4You2!
Location	Texas	Texas	California	California
Time period	2004–2007	2005–2009	2000–2003	2006–2010
Intervention setting	Middle schools	Middle schools	Alternative high schools	Alternative high schools
Number of schools	10	15	24	11
Sample size[a]	907	1258	988	765
Baseline data collection[b]	7th grade	7th grade	N/a	N/a
Number of observations	4	5	4	3
Treatment arms	2	3	2	4
Number of cohorts	1	1	4	Multiple, continued recruitment
Main study outcomes	Initiation of vaginal sex	Initiation of vaginal sex	Frequency of sex, vaginal without a condom, use of condom at last intercourse, number of partners without a condom, use of effective pregnancy prevention	Frequency of unprotected vaginal sex
Follow-ups	Baseline, 6 months, 18 months, 24 months	Baseline, 6 months, 18 months, 24 months, 36 months	Baseline, 6 months, 12 months, 18 months	Baseline, 7 months, 19 months
Data collection mode	Audio CASI	Audio CASI	Paper and pencil	Paper and pencil
% Female	59.1	59.8	37.3	46.7%
% African-American	42.3	39.3	27.7	22.1%
% Hispanic	44.0	48.4	29.1	37.6%
Baseline mean age (SD)	13.0	12.6	16.3	16.2

[a]Sample size reported in the main outcome manuscript

[b]Students attending alternative schools were not classified in grades; rolling recruitment was utilized

partners with whom had vaginal sex, and (3) number of times with whom had vaginal sex without a condom exceeded the number of times with whom had vaginal sex. Based on the above definition, a 0, 1 flag was created to identify each case as 1 = inconsistent within time, or 0 = consistent within time for each time point in each study.

Across-Time Inconsistency

Sexual experience (i.e., "ever had vaginal sex") was used to identify cases that were inconsistent across the entire study period. If a respondent provided contradictory information across time with regard to this outcome, they were flagged as inconsistent. For example, if someone responded, "yes, I have had vaginal sex" at the baseline observation but responded "no, I have never had vaginal sex" at a subsequent follow-up, this was coded as an inconsistent case. A 0, 1 flag variable indicating 1 = at least one inconsistency across time and 0 = no inconsistencies across time was created for each respondent in each of the four datasets. Thus, there were four possible patterns of inconsistent responses that would lead to this flag variable receiving a value of 1 for a given respondent in the study with 3 survey waves (Coyle et al., 2013)—YNN, YNY, YYN, NYN; nine patterns for the two

studies with 4 survey waves (Coyle et al., 2006, Tortolero et al.)—YNNN, NYNN, NNYN, YNNY, YNYN, YYNN, YYYN, YNYY, YYNY; and 16 patterns for the study with 5 survey waves (Markham et al., 2012)—e.g., YNNNN, NYNNN, NNYNN.

Any Inconsistency

The number of respondents with an inconsistent response either within time at any time point or across time was also calculated to examine the number of observations that might be lost in the worst case (recoding any inconsistency to missing). This was accomplished for each of the studies by creating a flag that denoted if a case was identified as inconsistent within time for any of the time points within the study, or inconsistent across time.

Potential Correlates

We examined the relationship between inconsistent responders and the following variables: age (in years), sex (female = 1), race/ethnicity (black, Hispanic, white, Asian, others), GRT treatment arm (1 = intervention), and perceived normative beliefs about abstinence (1 = strongly disagree to 5 = strongly

agree). These were selected because they are known correlates of sexual risk behaviors in adolescents (Kirby et al., 2007; Scott et al., 2011) and because they were available in the datasets. Perceived peer norms about abstinence were measured by the question "Most of my friends think it is OK for people my age NOT to have sex," with response options ranging from 1 = "strongly disagree" to 5 = "strongly agree."

Analysis

First, the overall rates of inconsistencies were calculated using the flag variables described above. Second, these rates were computed for age, gender, race/ethnicity, and intervention condition. Two-sample *t*-tests were used to evaluate whether inconsistency rates differed significantly across subgroups.

Multilevel logistic random effect regression models were used to identify and quantify the degree to which hypothesized correlates (age, sex, race/ethnicity, perceived peer norms, and intervention exposure) were related to inconsistent reporting both within time point, across time, and for the combined outcome of any inconsistency. Two-level models were fit (level 1 student, level 2 school) with a random intercept estimated at the school level for potential intra-class correlation present among students within the same school. To increase power for the regression analyses, the datasets from the two middle school studies were combined, and the datasets from the two high school studies were combined. In concert with the original study analyses, we used multilevel modeling to adjust for the intra-class correlation that may be present between students attending the same schools. For all models, the dependent variable (within-time consistency, across-time consistency, any inconsistency) was coded 0, 1 with 1 indicating the presence of inconsistent reporting.

Results

The rates of any inconsistencies combined (within or across time) ranged from 12 to 18% across the four studies (Table 2). Rates were higher for across-time (6–11%) than for within-time (5–8%) inconsistencies in all but one trial (13%). Within-time inconsistency rates decreased at later measurement occasions (Table 2). None of the inconsistency rates differed significantly by treatment group. In all trials, rates were higher for males than for females, though differences were not statistically significant in the high school studies. Any inconsistency rates were as high as 25% for males in one of the middle school trials (Table 2).

At the high school level, age, normative beliefs about abstinence, and self-identification as Asian race/ethnicity were most strongly associated with across-time inconsistencies in

multiple logistic regression analyses. Age was measured in years at baseline and ranged from 14 to 18. Of the two high school studies combined (Table 3), with older high school-aged students ($OR = 0.80$, $p < 0.01$), and students who were higher on the normative peer beliefs about abstinence scale ($OR = 0.77$, $p < 0.01$) having lower rates of inconsistencies, and students who self-identify as being Asian having higher rates of inconsistencies ($OR = 1.9$, $p < 0.05$). Only age was strongly associated with within-time inconsistencies for this age group, with older students being more likely to report within-time inconsistencies ($OR = 1.2$, $p < 0.05$). Only normative peer beliefs about abstinence were significantly associated with any inconsistencies for the high school groups ($OR = 0.86$. $p < 0.05$).

For the middle school studies (Table 3), age (range 12–15) and race/ethnicity were significantly associated with across-time inconsistencies, with students more likely to report inconsistencies if they were older ($OR = 1.5$, $p < 0.01$) or identified as black ($OR = 2.6$, $p < 0.001$) or "other" race ($OR = 2.3$, $p < 0.01$). Similarly, age and identifying as black were significantly associated with within-time inconsistencies as was gender, with older students ($OR = 1.5$, $p < 0.01$), and those who identified as black ($OR = 2.5$, $p < 0.001$) more likely to report within time inconsistencies and students identifying as females ($OR = 0.22$, $p < 0.001$) less likely to report these inconsistencies. Overall, older students ($OR = 1.6$. $p < 0.001$), black ($OR = 2.6$, $p < 0.001$), and "other" ($OR = 1.8$. $p < 0.05$) race students were more likely to report any inconsistencies in the middle school groups, while females ($OR = 0.44$, $p < 0.001$) were less likely to report similar inconsistencies. Peer norms regarding abstinence were not included in the regression analyses for the middle school groups since the normative belief survey items differed sufficiently so as to prevent harmonization.

Discussion

In each of four different studies conducted in California and Texas, more than 1 in 10 adolescents in middle and high school reported at least one type of inconsistent response to questions about their sexual behaviors. Across-time inconsistencies, which focused on responses to the question "have you ever had vaginal sex?" at different measurement occasions, made up the bulk of the inconsistencies for three of the four studies. The methodology used by definition only allows for the detection of students "recanting" their report of ever having had vaginal sex (changing from "yes" to "no") but does not allow for identification of inaccurate reporting that produces a plausible pattern of behavior.

Within-time inconsistencies, which focus on such problems as reporting more sexual partners than times had sex, were less prevalent except in the "All4You2!" study, which

Table 2 Across-time and within-time inconsistency

	It's Your Game	All About Youth	All4You!	All4You2!
Sample size	1207	1682	709	650
Any inconsistency (within or across time)	11.9%	14.4%	14.7%	18.4%
Across-time inconsistency				
Overall	8.5%	10.0%	11.0%	6.2%
Male	10.4%	13.6%	12.2%	7.4%
Female	7.0%	7.5%	9.4%	4.8%
Test for difference[a]	$p = 0.04$	$p < 0.001$	$p = 0.27$	$p = 0.16$
African-American	11.5%	12.7%	8.4%	5.1%
Hispanic	4.3%	7.8%	10.3%	6.4%
White	8.1%	4.7%	9.0%	2.5%
Others	12.4%	12.3%	14.5%	9.3%
Asian/Pacific Islander[c]	n/a	n/a	14.5%	12.5%
Test for difference[b]	$p < 0.001$	$p = 0.008$	$p = 0.33$	$p = 0.19$
Control group	8.0%	9.9%	12.7%	6.9%
Treatment group	9.1%	10.0%	9.9%	5.9%
Test for difference[a]	$p = 0.49$	$p = 0.94$	$p = 0.24$	$p = 0.62$
Within-time inconsistency				
Overall	4.8%	6.1%	7.5%	13.4%
Male	8.2%	9.4%	7.9%	13.5%
Female	2.1%	3.8%	6.5%	13.2%
Test for difference[a]	$p < 0.001$	$p < 0.001$	$p = 0.42$	$p = 0.91$
African-American	6.5%	7.7%	8.0%	14.8%
Hispanic	3.3%	4.8%	7.3%	12.9%
White	2.6%	6.7%	7.4%	10.7%
Others	3.9%	6.0%	9.7%	16.4%
Asian/Pacific Islander[c]	n/a	n/a	4.6%	12.5%
Test for difference[b]	$p = 0.07$	$p = 0.13$	$p = 0.55$	$p = 0.73$
Control group	4.2%	7.1%	8.2%	9.2%
Treatment group	5.5%	5.7%	7.0%	15.0%
Test for difference[a]	$p = 0.40$	$p = 0.25$	$p = 0.50$	$p = 0.03$

All samples are restricted to a valid response to "ever had sex" for 2 or more time points

[a] t-test for the difference between the two groups

[b] Chi-squared test for difference across the group

[c] n/a when the sample size is too small to estimate

involved high school-aged youth attending district-run alternative schools, including students navigating academic and discipline challenges (Coyle et al., 2013). This sample had higher rates of sexual activity, in general, and it is possible that their reports of more frequent sexual behavior could have contributed to recall challenges (Dareng et al., 2017). Literacy and survey format may have also played a role, as many of these young people were working to catch up on their school credits, and the survey was paper-based and appeared long because of the formatting. The use of online surveys with automatic skip patterns and audio options that are now routine in behavioral studies may address some of these issues.

The overall rates of inconsistencies were somewhat higher for the high school youth relative to the middle school youth. Again, this could be a result of high school students simply having higher overall rates of sexual activity, making it more difficult for them to have an accurate recall. Additionally, these trends may reflect developmental differences in influencing factors that may affect sexual behavior and its reporting. For example, developmental neuroscience highlights that being in the presence of peers, whether in physical or virtual spaces, activates reward circuitry in the brain (Suleiman & Brindis, 2014). This "peer effect" increases the rewarding feelings that are generated from engaging in a wide range of risk-taking or sensation-seeking behaviors.

Table 3 Factors associated with inconsistencies for middle school- and high school-combined datasets in logistic regression models

Dependent variable	Combined middle school studies IYG+AAY (Houston, TX)			Combined high school studies All4You! + All4You2! (Northern CA)		
	Model 1: across-time inconsistency	Model 2: within-time inconsistency	Model 3: any inconsistency	Model 1: across-time inconsistency	Model 2: within-time inconsistency	Model 3: any inconsistency
Covariate (response option)	OR (95% CI)	OR (95%CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Treatment arm (1=treatment)	1.3 (0.89, 2.0)	0.93 (0.53, 1.7)	1.1 (0.82, 1.5)	0.81 (0.54, 1.2)	1.2 (0.83, 1.8)	1.1 (0.79, 1.5)
Age (13, 14, …, 20)	1.4** (1.1, 1.8)	1.5** (1.2, 1.9)	1.6*** (1.3, 1.9)	0.80** (0.69, 0.94)	1.2* (1.0, 1.4)	1.0 (0.92, 1.2)
Gender (1=female)	0.72 (0.52, 1.0)	0.22*** (0.16, 0.37)	0.44*** (0.34, 0.59)	0.75 (0.49, 1.2)	1.0 (0.70, 1.4)	1.1 (0.72, 1.3)
Normative beliefs about abstinence (1=strongly disagree;5=strongly agree)[a,b]				0.77** (0.65, 0.92)	0.91 (0.80, 1.1)	0.86* (0.77, 1.0)
Hispanic (ref)						
Black (1=yes)	2.6*** (1.8, 3.8)	2.4*** (1.5, 3.9)	2.6*** (1.9, 3.6)	0.82 (0.46, 1.5)	1.0 (0.66, 1.6)	1.1 (0.80, 1.6)
Asian (1=yes)[c]				1.9* (1.1, 3.4)	0.66 (0.34, 1.3)	1.1 (0.68, 1.7)
White (1=yes)	0.43 (0.06, 3.2)	1.1 (0.25, 4.8)	0.73 (0.22, 2.4)	0.67 (0.33, 1.4)	0.90 (0.54, 1.5)	0.83 (0.54, 1.3)
Others (1=yes)	2.3** (1.3, 4.0)	1.4 (0.67, 2.9)	1.8* (1.2, 3.0)	1.6 (0.92, 2.7)	1.3 (0.79, 2.0)	1.4 (0.96, 2.0)

OR denotes the odds ratio of the likelihood of an inconsistent response

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

[a] Normative beliefs about abstinence were measured by the item "Most of my friends think it is OK for people my age NOT to have sex"

[b] Items not consistent across middle school studies, could not be combined

[c] Asian was combined with others in the middle school combined sample due to the small sample size in this classification

School-based studies typically collect data in group settings; further research could help explore the potential influences of peer presence on the consistency of reporting sexual behaviors. Other developmental factors, such as the heightened importance of social acceptance, may also contribute (Crone & Dahl, 2012); however, these influences are at play across the adolescent period suggesting there are other factors that may be contributing to the higher rates of inconsistent reporting among the high school-aged young people in these studies.

The findings have implications for study power since almost all studies use some form of data cleaning procedure to resolve logical inconsistencies (Brener et al., 2003, 2004). While it may be tempting to regard recoding to missing all inconsistent responses as the most conservative/accurate method, some authors note the possibility that inconsistencies due to reporting "yes" to ever had vaginal sex at one time point and "no" at a follow-up time point may balance inaccuracies in "the other direction" (i.e., "no" then "yes," where the yes is actually false) that are not detectable through inconsistency checks (Rodgers et al., 1982; Steuve & O'Donnell, 2000). On the other hand, different methods of recoding, such as using the most recent responses as gold standards and carrying forward "yes" responses, have been shown to affect estimates of the prevalence of sexually active adolescents, and generally can only increase inaccuracy estimates (Upchurch et al., 2002; Steuve & O'Donnell, 2000). Furthermore, for these reasons, while studies of survey administration mode have found higher overall reports of sensitive behaviors on electronic surveys (e.g., audio CASI) (Romer et al., 1997), it is not clear that electronic surveys, which can facilitate skip provide a wholesale solution to the problem of inconsistent reporting of behaviors (Bloom, 1998; McAuliffe et al., 2007; Steuve & O'Donnell, 2000). There is a delicate balance between maintaining teens' trust in the confidentiality and veracity of their own survey responses and building in forcing functions to make them resolve inconsistencies online in real time.

It is therefore worth acknowledging and planning for certain inevitable data cleaning procedures to evaluate sexual health interventions. Removing inconsistent cases impacts power due to loss of sample size. Based on our results, when determining the needed sample size for school-based studies of sexual health programs, an additional inflation factor of at least 1.14 should be considered to account for the loss of on average 12% of the sample due to inconsistencies $(1/(1-0.12) = 1.14)$.

We found no association of inconsistent reporting with treatment condition assignment, therefore the impact on the evaluation of intervention effectiveness would not be a primary concern; however, cleaning or recoding cases may impact the ability to generalize results to the intended population, given inconsistency rates do not occur at random, as found in all four studies examined here. Age, race/ethnicity,

and sex all were associated significantly with inconsistent responses, primarily at the middle school level Goldberg et al. (2014) also reported variation in reporting consistency by race/ethnicity and sex and discussed the potential impact of racial/ethnic and gender norms on reporting. Cultural norms may affect the timing of dating and sexual experiences, expectations of partners, and related beliefs and values. There is a need for further research on the role of cultural and gender norms on inconsistencies. During elicitation focus groups for another study with urban middle school young people (Coyle et al., 2019), the young men talked about the gendered "milestones" and pressures they felt to have sex by the end of 8th grade. If these types of expectations are wider spread, they may account for why we saw higher rates of inconsistent reporting for older youth in middle school. It is also possible that for males, overreporting may be a strategy for navigating these types of pressures or a perceived opportunity to gain social status. Finally, it is also possible that higher inconsistencies among older youth reflect an increased understanding of the behaviors being assessed, which has been discussed for other behaviors, such as substance use (Broman et al., 2022).

Given these findings, if data are cleaned to delete inconsistent observations, the sample may no longer be representative of the priority population. However, if the data are retained without adjustment and they are not accurate, they may pose a threat to the external study validity. One solution is to add an examination of correlates of inconsistencies to data cleaning processes, and then re-weight analytic samples after deleting inconsistent cases. For example, if male participants have a higher rate of inconsistent reporting and were therefore cleaned out of the analytic sample at a higher rate than females, the remaining male respondents might need to be weighted more heavily to bring the analytic sample back into alignment with the overall sample. Thus, the application of sample weights that account for the differential loss of subjects following the removal of unevenly distributed inconsistent cases could be used to realign a cleaned analytic study sample back to the population it was intended to represent. While attrition analyses are commonplace in understanding the impact of subject retention on study findings, rarely do studies include or report on the analysis of cases lost to data cleaning efforts, which generally result in outcome-specific missing values rather than subject-wide exclusion. In the absence of investigation into subject loss from the analyses solely due to data cleaning protocols, impacts on the analytic sample go largely unreported. At the very least, evaluations should acknowledge the differences in the demographic distributions of cases impacted by data cleaning efforts and how these impact the representation in the analyzed sample compared to the recruited sample. This is particularly important for the main study outcomes.

Our finding that high school adolescents who believed more strongly that their peers thought it was ok not to have

sex were less likely to report inconsistencies in their sexual activity over time ($OR = 0.77$) is consistent with the social desirability theory of inconsistencies (Brown et al., 2012). It is also consistent with the literature on how resistance to peer influences changes from early to later adolescence (e.g., Steinberg & Monahan., 2007). Adolescents who feel less pressure to report they were more sexually active than they were would have less cause to be inconsistent in their reports over time. We did not include this measure in the middle school analyses for this study because of item variation, but we would expect to see a similar pattern among middle school adolescents given developmental changes pushing adolescents to prioritize issues such as social networks, belonging, and autonomy that begin with the onset of puberty, which typically occurs between ages 9 and 12 among US adolescents (typically earlier in girls) (Crone & Dahl, 2012).

This finding suggests new potential leverage points for reducing inconsistencies in the first place, which as additional assurances as to the privacy of students' responses, and using a developmental framing when asking questions to address any potential shame or embarrassment. It also highlights the need for further research on the motivational biases that may be at play and the need to engage young people in solutions for potentially addressing these biases in survey methodology. Finally, these findings also highlight the potential of examining other potential correlates, such as other risk behaviors or peer influences that may inform study design and/or analyses.

Limitations

It is important to note that our definition of inconsistencies is an incomplete proxy for measuring all inaccuracies in the data. This study used secondary analyses and relied on existing survey items rather than constructing new ones to assess the consistency and accuracy of self-report. We also did not have other indicators to validate self-report data, such as biomarker testing. Given the age and sensitivity of the study sample, skip patterns were employed so that if a respondent reported they were not sexually experienced, they were not asked about subsequent sexual behaviors such as number of partners, number of times, or condom use. This did not impact the across-time consistency assessment, given all respondents were asked about sexual experience; however, this limited assessment of within-time inconsistency measurement because a participant's initial response impacts exposure to other sexual behavior questions. For example, we would not detect instances of within-time inconsistency in which a participant answered "no" to the ever had vaginal sex question and was not asked to report on other measures of sexual behavior. Additionally, this study did not measure the accuracy of self-report, only inconsistent

reporting as a proxy for suspected inaccuracy. Furthermore, when inconsistencies were noted, the study was not able to identify with any precision which measure was likely to be the most accurate.

Public Health Implications

Public health surveillance of sexual risk behaviors and evaluations of sexual health interventions routinely rely on self-report data that may have a non-random subset of invalid data due to inconsistent response. Studies can improve the reliability of their sample size estimates by using additional inflation factors when conducting sample size calculations at the study-planning stage to account for potential data loss. They can further improve sample accuracy by examining the correlates of inconsistencies in their sample and using re-weighting schemes to account for this source of bias in the analyzable compared to the intended sample. Finally, they should attempt to reduce inconsistent reporting of adolescent sexual risk behaviors by focusing on novel ways to reduce the influence of peer normative beliefs in data collection processes and settings. These measures, together with the routine reporting of observed rates of inconsistencies, may adjust for biases in the population to which a study generalizes, thereby aiding public health practitioners and policy-makers looking to adopt programs for their particular population.

Declarations

Ethics Approval Ethical approval was waived by the local Ethics of Education, Training, and Research in view of the retrospective nature of the study that utilized deidentified secondary data collected from previous randomized controlled trials all of which had ethical approval.

Consent to Participate Informed consent was obtained from legal guardians.

Conflict of Interest The authors declare no competing interests.

References

Alexander, C. S., Somerfield, M. R., Ensminger, M. E., Johnson, K. E., & Kim, Y. J. (1993). Consistency of adolescents' self-report of sexual behavior in a longitudinal study. *Journal of Youth and Adolescence, 22*(5), 455–471.

Beguy, D., Kabiru, C. W., Nderu, E. N., & Ngware, M. W. (2009). Inconsistencies in self-reporting of sexual activity among young people in Nairobi, Kenya. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 45*, 595–601. https://doi.org/10.1016/j.jadohealth.2009.03.014

Bloom D. E. (1998). Technology, experimentation, and the quality of survey data. *Science (New York, N.Y.), 280*(5365), 847–848. https://doi.org/10.1126/science.280.5365.847

Brener, N. D., Billy, J. O., & Grady, W. R. (2003). Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: Evidence from the scientific literature. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 33*, 436–457. https://doi.org/10.1016/s1054-139x(03)00052-1

Brener, N. D., Collins, J. L., Kann, L., Warren, C. W., & Williams, B. I. (1995). Reliability of the youth risk behavior survey questionnaire. *American Journal of Epidemiology, 141*(6), 575–580.

Brener, N. D., Grunbaum, J. A., Kann, L., McManus, T., & Ross, J. (2004). Assessing health risk behaviors among adolescents: The effect of question wording and appeals for honesty. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 35*, 91–100. https://doi.org/10.1016/j.jadohealth.2003.08.013

Brener, N. D., Kann, L., McManus, T., Kinchen, S. A., Sundberg, E. C., & Ross, J. G. (2002). Reliability of the 1999 youth risk behavior survey questionnaire. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 31*, 336–342. https://doi.org/10.1016/s1054-139x(02)00339-7

Broman, M. J., Bista, S., & Broman, C. L. (2022). Inconsistency in self-reporting the use of substances over time. *Substance Use & Misuse, 57*, 1356–1364. https://doi.org/10.1080/10826084.2022.2083168E

Brown, J. L., Sales, J. M., DiClemente, R. J., Salazar, L. F., Vanable, P. A., Carey, M. P., Brown, L. K., Romer, D., Valois, R. F., & Stanton, B. (2012). Predicting discordance between self-reports of sexual behavior and incident sexually transmitted infections with African American female adolescents: Results from a 4-city study. *AIDS and Behavior, 16*, 1491–1500. https://doi.org/10.1007/s10461-012-0163-8

Coyle, K., Anderson, P., Franks, H., Walker, J., & Glassman, J. (2019). You-me-us: Results of a cluster randomized trial of a health relationships approach to sexual risk reduction. *The Journal of Primary Prevention, 40*, 607–629. https://doi.org/10.1007/s10935-019-00569-w

Coyle, K. K., Glassman, J. R., Franks, H. M., Campe, S. M., Denner, J., & Lepore, G. M. (2013). Interventions to reduce sexual risk behaviors among youth in alternative schools: A randomized controlled trial. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 53*, 68–78. https://doi.org/10.1016/j.jadohealth.2012.12.012

Coyle, K. K., Kirby, D. B., Robin, L. E., Banspach, S. W., Baumler, E., & Glassman, J. R. (2006). All4You! A randomized trial of an HIV, other STDs, and pregnancy prevention intervention for alternative school students. *AIDS Education and Prevention: Official Publication of the International Society for AIDS Education, 18*, 187–203. https://doi.org/10.1521/aeap.2006.18.3.187

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience, 13*, 636–650. https://doi.org/10.1038/nrn3313

Dareng, E. O., Adebamowo, S. N., Eseyin, O. R., Odutola, M. K., Pharoah, P. P., & Adebamowo, C. A. (2017). Test-retest reliability of self-reported sexual behavior history in urbanized Nigerian women. *Frontiers in Public Health, 5*, 172. https://doi.org/10.3389/fpubh.2017.00172

DiClemente, R. J. (2016). Validity of self-reported sexual behavior among adolescents: Where do we go from here? *AIDS and Behavior, 20*, S215–S217. https://doi.org/10.1007/s10461-015-1217-5

DiClemente, R. J., Swartzendruber, A. L., & Brown, J. L. (2013). Improving the validity of self-reported sexual behavior: No easy

answers. *Sexually Transmitted Diseases, 40*, 111–112. https://doi.org/10.1097/OLQ.0b013e3182838474

Goldberg, S. K., Haydon, A. A., Herring, A. H., & Halpern, C. T. (2014). Longitudinal consistency in self-reported age of first vaginal intercourse among young adults. *Journal of Sex Research, 51*(1), 97–106.

Hamilton, D. T., & Morris, M. (2010). Consistency of self-reported sexual behavior in surveys. *Archives of Sexual Behavior, 39*, 842–860. https://doi.org/10.1007/s10508-009-9505-7

Kirby, D. B., Laris, B. A., & Rolleri, L. A. (2007). Sex and HIV education programs: Their impact on sexual behaviors of young people throughout the world. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 40*, 206–217. https://doi.org/10.1016/j.jadohealth.2006.11.143

Markham, C. M., Tortolero, S. R., Peskin, M. F., Shegog, R., Thiel, M., Baumler, E. R., Addy, R. C., Escobar-Chaves, S. L., Reininger, B., & Robin, L. (2012). Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 50*, 279–288. https://doi.org/10.1016/j.jadohealth.2011.07.010

McAuliffe, T. L., DiFranceisco, W., & Reed, B. R. (2007). Effects of question format and collection mode on the accuracy of retrospective surveys of health risk behavior: A comparison with daily sexual activity diaries. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association, 26*, 60–67. https://doi.org/10.1037/0278-6133.26.1.60

McFarlane, M., & St Lawrence, J. S. (1999). Adolescents' recall of sexual behavior: Consistency of self-report and effect of variations in recall duration. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 25*, 199–206. https://doi.org/10.1016/s1054-139x(98)00156-6

Napper, L. E., Fisher, D. G., Reynolds, G. L., & Johnson, M. E. (2010). HIV risk behavior self-report reliability at different recall periods. *AIDS and Behavior, 14*, 152–161. https://doi.org/10.1007/s10461-009-9575-5

Newcomer, S., Udry, J. R., & Hill, C. (1988). Adolescents' honesty in a survey of sexual behavior. *Journal of Adolescent Research, 3*, 419–423.

O'Sullivan, L. F. (2008). Challenging assumptions regarding the validity of self-report measures: The special case of sexual behavior. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 42*, 207–208. https://doi.org/10.1016/j.jadohealth.2008.01.002

Palen, L. A., Smith, E. A., Caldwell, L. L., Flisher, A. J., Wegner, L., & Vergnani, T. (2008). Inconsistent reports of sexual intercourse among South African high school students. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 42*, 221–227. https://doi.org/10.1016/j.jadohealth.2007.08.024

Ramos, D. O., Daly, M., Seidl-de-Moura, M. L., Jomar, R. T., & Nadanovsky, P. (2017). Inconsistent reports of risk behavior among Brazilian middle school students: National School Based Survey of Adolescent Health (PeNSE 2009/2012). *Cadernos De Saude Publica, 33*, e00145815. https://doi.org/10.1590/0102-311X00145815

Rodgers, J. L., Billy, J. O., & Udry, J. R. (1982). The rescission of behaviors: Inconsistent responses in adolescent sexuality data. *Social Science Research, 11*, 280–296.

Romer, D., Hornik, R., Stanton, B., Black, M., Li, X., Ricardo, I., & Feigelman, S. (1997). "Talking" computers: A reliable and private method to conduct interviews on sensitive topics with children. *Journal of Sex Research, 34*, 3–9.

Rose, E., Diclemente, R. J., Wingood, G. M., Sales, J. M., Latham, T. P., Crosby, R. A., Zenilman, J., Melendez, J., & Hardin, J. (2009). The validity of teens' and young adults' self-reported condom use.

Archives of Pediatrics & Adolescent Medicine, 163, 61–64. https://doi.org/10.1001/archpediatrics.2008.509

Rosenbaum, J. E., Zenilman, J. M., Rose, E., Wingood, G. M., & DiClemente, R. J. (2017). Semen says: Assessing the accuracy of adolescents' self-reported sexual abstinence using a semen Y-chromosome biomarker. *Sexually Transmitted Infections, 93*, 145–147. https://doi.org/10.1136/sextrans-2016-052605

Rosenbaum, J. E. (2009). Truth or consequences: The intertemporal consistency of adolescent self-report on the Youth Risk Behavior Survey. *American Journal of Epidemiology, 169*, 1388–1397.

Rosenbaum, J. E. (2006). Reborn a virgin: Adolescents' retracting of virginity pledges and sexual histories. *American Journal of Public Health, 96*, 1098–1103. https://doi.org/10.2105/AJPH.2005.063305

Scott, M. E., Wildsmith, E., Welti, K., Ryan, S., Schelar, E., & Steward-Streng, N. R. (2011). Risky adolescent sexual behaviors and reproductive health in young adulthood. *Perspectives on Sexual and Reproductive Health, 43*, 110–118. https://doi.org/10.1363/4311011

Siddiqui, O., Mott, J. A., Anderson, T. L., & Flay, B. R. (1999). Characteristics of inconsistent respondents who have "ever used" drugs in a school-based sample. *Substance Use & Misuse, 34*, 269–295. https://doi.org/10.3109/10826089909035646

Steinberg, L., & Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental Psychology, 43*, 1531–1543. https://doi.org/10.1037/0012-1649.43.6.1531

Stueve, A., & O'Donnell, L. (2000). Inconsistencies over time in young adolescents' self-reports of substance use and sexual intercourse. *Substance Use & Misuse, 35*, 1015–1034. https://doi.org/10.3109/10826080009148430

Suleiman, A. B., & Brindis, C. D. (2014). Adolescent school-based sex education: Using developmental neuroscience to guide new directions for policy and practice. *Sexuality Research and Social Policy, 11*, 137–152. https://doi.org/10.1007/s13178-014-0147-8

Tenkorang, E. Y. (2021). Inconsistent reports of sexual intercourse by adolescents in Edo State. *Nigeria. Sexuality Research and Social Policy, 18*(2), 271–280.

Tortolero, S. R., Markham, C. M., Peskin, M. F., Shegog, R., Addy, R. C., Escobar-Chaves, S. L., & Baumler, E. R. (2010). It's your game: Keep it real: Delaying sexual behavior with an effective middle school program. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 46*, 169–179. https://doi.org/10.1016/j.jadohealth.2009.06.008

Upchurch, D. M., Lillard, L. A., Aneshensel, C. S., & Fang Li, N. (2002). Inconsistencies in reporting the occurrence and timing of first intercourse among adolescents. *Journal of Sex Research, 39*, 197–206. https://doi.org/10.1080/00224490209552142

Weinhardt, L. S., Forsyth, A. D., Carey, M. P., Jaworski, B. C., & 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*, 155–180. https://doi.org/10.1023/a:1018682530519