



Using Composite Scores to Summarize Adolescent Sexual Risk Behavior: Current State of the Science and Recommendations

David H. Barker^{1,2} · Lori A. J. Scott-Sheldon^{1,3,4} · Daniel Gittins Stone⁵ · Larry K. Brown^{1,2}

Received: 2 February 2018 / Revised: 23 July 2019 / Accepted: 24 July 2019 / Published online: 19 August 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Composite scores offer the advantage of summarizing across multiple sexual risk behaviors to both simplify results and better capture the influence of core contextual, interpersonal, and intrapersonal dynamics that affect multiple sexual risk behaviors. There is inconsistency in how researchers utilize composite scores with minimal guidance on the advantages and disadvantages of frequently used approaches. Strengths and weaknesses of each approach are discussed in the context of assessing adolescent sexual risk behavior. A latent variable model and three commonly used composites were applied to data combined across four clinical trials ($n = 1322$; 50% female). Findings suggested that the latent variable approach was limited due to minimal correlations among sexual risk behaviors, that choice of composite had minimal impact on cross-sectional results so long as there is sufficient variability in risk behavior in the sample, but composite choice could impact results from clinical trials particularly for subgroup analyses. There are unique challenges to creating composites of adolescent risk behavior, including the fluidity and infrequency of adolescent sexual relationships that result in many participants reporting no sexual behavior at any given assessment and a low correlation between the number of partners and condomless sex acts. These challenges impede application of data-driven approaches to defining sexual risk composites. Recommendations to improve consistency in reporting include: (1) reporting each type of risk behavior separately prior to forming a composite, (2) aggregating across assessments to increase the chance of observing sexual risk behaviors, and (3) continued work toward a unified definition of adolescent sexual risk behavior that can guide the development of appropriate measurement models.

Keywords Adolescence · Sexual risk behavior · HIV · Clinical trials

Introduction

Sexual risk prevention during adolescence focuses on a number of behavioral targets, including delaying onset of sexual activity (Tortolero et al., 2010), reducing the number

of condomless sex acts (Coyle et al., 2006), reducing the number of partners (Jemmott et al., 2010), and increasing condom use (DiClemente et al., 2009). Some programs focus on a specific behaviors such as condom use and abstinence (Boekeloo et al., 1999), while others focus on developmental and contextual processes that influence sexual risk behavior such as emotion regulation (Raffaelli & Crockett, 2003), parent–adolescent communication (Huebner & Howell, 2003), parental monitoring (Huebner & Howell, 2003; Li et al., 2000), and social networks and interpersonal connectedness (Markham et al., 2010). Programs that focus on these important but more distal processes likely impact more than one behavioral target of sexual risk. This impact, however, may be diffuse and may not be apparent when considering a single sexual risk behavior. These important, but diffuse, effects may only be detected when all behaviors are considered (Brown et al., 2011). While there has been guidance on how to assess sexual risk behavior (Dolezal et al., 2012; Fenton, 2001; Fonner et al., 2013; Graham et al., 2005; Schroder

✉ David H. Barker
David_Barker@brown.edu

¹ Department of Psychiatry and Human Behavior, The Warren Alpert Medical School, Brown University, Box-G-BH, Providence, RI 02906, USA

² Division of Child and Adolescent Psychiatry, Rhode Island Hospital, Providence, RI, USA

³ Centers for Behavioral and Preventive Medicine, The Miriam Hospital, Providence, RI, USA

⁴ Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, USA

⁵ Department of Applied Psychology, Bouvé College of Health Sciences, Northeastern University, Boston, MA, USA

et al., 2003a, b), there has been limited guidance on how to integrate information across multiple adolescent risk behaviors. As a result, investigators have used multiple approaches complicating the accumulation of knowledge across studies (Mullen et al., 2002). The intent of this article is to discuss the challenges of integrating information across multiple adolescent sexual risk behaviors and examine how different approaches to forming composites influence conclusions drawn from randomized trials.

Adolescent sexual risk prevention trials typically measure specific risk behaviors and, to a lesser extent, the interpersonal and intrapersonal processes that influence the behaviors (e.g., beliefs and cognitions, emotion regulation, family processes, peer process, partner characteristics). Specific behaviors commonly assessed in adolescent prevention trials that are linked to increased HIV risk include age of sexual debut (Falasinnu et al., 2015; Heywood et al., 2015; Slater and Robinson, 2014), recency of sexual activity (Lightfoot, 2012), number of partners (Ashenhurst et al., 2017; Pequegnat et al., 2016), number of sexual acts, and number of condomless acts (Fonner et al., 2013). Other important, but less frequently assessed, behaviors include partner concurrency (Ashenhurst et al., 2017; Le Pont et al., 2003) and alcohol or other drug use before sex. Because there is more consistency among studies in which sexual risk behaviors are included than in the underlying interpersonal and intrapersonal processes contributing to risk, we will focus on quantifying the constellation of sexual behaviors that place individuals at risk of contracting HIV.

Assessing and Summarizing Sexual Risk During Adolescence

Adolescent sexual risk behavior differs from adult risk behavior. For example, adolescent relationships are not as stable as adult relationships (Carver et al., 2003; Connolly & McIsaac, 2009). Compared to adults, adolescents report lower rates of sexual behavior (Herbenick et al., 2010) and higher rates of condom use (Reece et al., 2010) but also have higher rates of sexually transmitted infections (STIs) (Slater & Robinson, 2014). Examining sexual risk behaviors among adolescents presents unique challenges to the assessment of these behaviors as well as the operational definition of a composite index that combines information across multiple risk behaviors. These challenges will be discussed below.

Assessment of Sexual Risk Behaviors

Regardless of the population, researchers studying sexual risk behavior make several important decisions in how to assess behavior, including whether to ask about global estimates of sexual behaviors across a specified recall period (e.g., number of partners, number of sexual encounters),

whether to ask about event-level behaviors (e.g., condom use at last sex, alcohol or other drug use before last sex), whether to use categorical response options or ask for open-ended reports of frequency, and how much structure should be provided to assist with recall (e.g., ask questions partner-by-partner, timeline followback, computer-assisted self-interview). There have been a number of quality reviews that have addressed these measurement considerations (Dolezal et al., 2012; Fenton, 2001; Fonner et al., 2013; Graham et al., 2005; Schroder et al., 2003a, b; Scott-Sheldon et al., 2010), and it is not our intent to repeat these recommendations other than to highlight a couple of considerations specific to adolescent populations.

Recall and Numeracy Skills

Compared to adults, adolescents are not fully developed in terms of numeracy (e.g., estimation, casting, counting) or recall (Fenton, 2001; McAuliffe et al., 2010; Napper et al., 2010; Scott-Sheldon et al., 2010). They are also more sensitive to social norms and perceived expectations (Blakemore, 2008; Somerville, 2013). Together, these factors may lead to exaggerated or inaccurate estimates of sexual behavior when they are asked open-ended questions about sexual behavior. Moreover, the amount of error increases as a function of the number of events reported because recall is often more challenging as participants may use different cognitive strategies to report frequent (estimation, casting) versus infrequent (counting) events (McAuliffe et al., 2010). Thus, adolescents engaging in fewer behaviors provide more accurate reports of their behaviors than those who engage in more frequent behaviors. For these reasons, it is suggested that sexual risk assessments among adolescents include clear, familiar, and nonjudgmental language and that the recall period be no greater than 3 months (Scott-Sheldon et al., 2010). Providing additional scaffolding, such as asking partner-by-partner, timeline followback, or providing structured response options, will help with accuracy of the recall (Crosby et al., 1996; Weinhardt et al., 1998). Structured response options convey information about expectations and norms (Fenton, 2001), and care is required to ensure that the choice of response options appropriately reflect the expected rate of behavior in the population being assessed so not to suggest expected or normative responses. Ultimately, response options should be piloted and evaluated using cognitive interview strategies to help ensure that the structure of the question is not unduly influencing responses (Webb et al., 2015).

Fluidity and Sparsity of Adolescent Relationships

The fluidity of adolescent relationships and relatively low frequency of sexual partnerships are additional challenges to assessing change in sexual risk behavior during adolescence

(Manning et al., 2014). The fluidity of adolescent relationships complicates attempts to categorize partner status (e.g., steady, friends with benefits, hookups, acquaintance). Therefore, instead of asking adolescents to classify their relationships, it is advisable to directly assess pertinent partner characteristics (e.g., duration, quality, time spent with partner). Beyond the fluidity of adolescent relationships, the low frequency of sexual partnerships can be challenging to analyze and can obscure the benefit of an intervention approach. Even among high-risk samples, a significant number of adolescents will not engage in sexual behavior during a study's follow-up period and others will have extended periods of time without a sexual partner. The meaning behind the partnerless periods is not always clear. It may be that an intervention enabled some adolescents to choose not to partner or enabled them to partner but not engage in sexual behavior. Alternatively, some adolescents may not have changed their propensity for engaging in sexual risk behavior but may not have had the opportunity to partner during the observation period. The first two explanations would lend support to the efficacy of the intervention while the third would not. Asking solely about sexual partners does not allow researchers to differentiate among such possibilities. Measuring additional contextual information such as the number of nonsexual romantic relationships or about declining or avoiding sexual opportunities may help provide the needed contextual information to interpret the meaning of partnerless periods.

Sexual Risk Composites

Appropriately assessing sexual risk behaviors during adolescence helps improve the quality of the collected data as do choices about summarizing information across multiple types of sexual risk behavior. In the broader adolescent and adult literature, several approaches have been applied to quantifying the constellation of sexual risk behaviors. This variety is partially related to the challenges of defining a measurement model for adolescent sexual risk behavior. During adolescence, sexual behaviors are determined by a myriad of intrapersonal, interpersonal and contextual factors (Cooper, 2010). These factors differ depending on the decision being made by the adolescent and one decision may influence another. For example, the determinants driving the decision to engage in sex may differ from those driving the decision to use a condom or have multiple partners. Research also suggests complex interactions among gender, partner characteristics, relationship context, and condom use (Lescano et al., 2006; Senn et al., 2014; Staras et al., 2013). Differential determinants suggest that sexual behaviors may not be strongly related to one another when other important drivers and dynamics are not taken into account. In addition to having multiple differential determinants, sexual behaviors also relate differentially to health outcomes. For example, the behaviors that place

an adolescent at risk for pregnancy are not identical to those that place them at risk for acquiring HIV. These issues pose significant challenges to fitting latent variable measurement models where a latent construct is thought to give rise to the observed indicators. Such models require that the indicators be correlated and assume that the correlation is due to a common construct (Borsboom et al., 2003; Edwards, 2011). The low correlations among sexual behaviors may be why researchers attempting to fit such models tend to expand the indicators beyond behavior to include intentions and attitudes (Siegel et al., 2001). The limitations of conventional measurement models have resulted in researchers utilizing composites where behaviors are not presumed to be correlated or related (Bollen & Bauldry, 2011), or pre-defined categories of risk (e.g., low, medium, high). Some of the composite weights or category definitions are defined a priori while others are sample-dependent. Moreover, the resulting composite can be analyzed as a continuous variable or as discrete categories. These advantages and disadvantages of each approach as well as examples from the literature are listed in Table 1.

Using a priori weights and definitions are determined either by investigators or through validation of composites in samples other than the one being analyzed. Because using a priori definitions do not require advanced statistical models, they are the most commonly used approach to creating sexual risk composites. Unfortunately, there is little consistency in the definitions and limited work on developing and validating a unifying definition (Webb et al., 2015). In adolescent sexual risk behavior, composites are typically created by investigators using coding definitions and thresholds (e.g., two or more partners, any condomless sex) to create ordered categories (Bowleg et al., 2014; Brown et al., 2011; Epstein et al., 2014; Graham et al., 2013; Murphy et al., 2009). Definitions vary in the sexual behaviors that are included (e.g., number of partners, number of sexual acts, number of condomless acts, sex under the influence of alcohol or other drugs, participation in the sex trade, oral, anal, or vaginal sex) and in the thresholds that are used to define risk. This variability presents researchers with many options and the temptation to tweak the composite to provide results consistent with their expectations.

Compared with investigator-determined definitions, there are very few examples of investigators-using weights that were externally validated in other samples. One example of such a composite is the Vaginal Episode Equivalent Index (VEE; Susser et al., 1998). The VEE is a weighted composite of condomless oral, vaginal, and anal sex acts with weights calibrated to the relative risk of HIV transmission. Unfortunately, the VEE does not include other sexual risk behaviors important to adolescent sexual risk prevention such as multiple concurrent partners and substance use prior to sex. To our knowledge, there is no consistently used, empirically

Table 1 Organization of sexual risk composite scores

	Continuous	Categorical
A priori or sample-independent weights	<p><i>Examples</i></p> <p>Create an a priori composite risk score that captures overall risk on a one-dimensional continuous scale (Bancroft et al., 2003; Brown et al., 2011; Carey et al., 2004; Carmona et al., 2014; Fergus et al., 2007; Huang et al., 2012; Murry et al., 2011; Pearlman et al., 2002; Raffaelli & Crockett, 2003; Wilson & Widom, 2011)</p> <p><i>Pros</i></p> <ul style="list-style-type: none"> Easily created Simple to use in subsequent analyses Provides a single estimate for treatment response Can be applied to small samples <p><i>Cons</i></p> <ul style="list-style-type: none"> Meaning of the underlying dimension is not clear May miss important patterns in the observed data Meaning of the underlying dimension may change across studies and populations 	<p><i>Examples</i></p> <p>A priori categories based on patterns of risk behaviors (Bowleg et al., 2014; Epstein et al., 2014; Graham et al., 2013; Huebner & Howell, 2003; Li et al., 2000; Murphy et al., 2009; Mustanski, 2008; Siegel et al., 2001; Wu et al., 2005)</p> <p><i>Pros</i></p> <ul style="list-style-type: none"> Easily created Transparent mapping of behaviors to category assignment Provides conceptually meaningful categories of behavior that have been linked to increased risk of STIs, including HIV <p><i>Cons</i></p> <ul style="list-style-type: none"> Challenging to analyze particularly when each category is treated as discrete rather than ordered Results may differ depending on each category providing a more complex picture of results May miss important features of the observed data
	<p><i>Examples</i></p> <p>Latent variable measurement models (e.g., item response theory, exploratory factor analysis, principal components analysis; (Mattson et al., 2008; Siegel et al., 2001; Williams et al., 2001)</p> <p>Create a weighted composite based on predictive models of a specific outcome (e.g., contracting HIV; (Susser et al., 1998)</p> <p><i>Pros</i></p> <ul style="list-style-type: none"> Simple to use in subsequent analyses Single estimate for treatment response Captures information from the observed variance and covariance among sexual behaviors <p><i>Cons</i></p> <ul style="list-style-type: none"> Challenging to create Measurement model likely to vary across populations and ages Meaning of the underlying dimension may change across studies and populations Requires covariance among sexual behaviors The meaning of the underlying dimension is not clear Requires larger samples 	<p><i>Examples</i></p> <p>Latent class (Beadnell et al., 2005; Hipwell et al., 2011; Lanza & Collins, 2008; Vasilenko et al., 2014)</p> <p>Cluster analysis (Newman & Zimmerman, 2000)</p> <p><i>Pros</i></p> <ul style="list-style-type: none"> Potential for clear mapping of behavior to categories (depends on the quality of the classification model) Uses information from the observed variance and covariance among sexual behaviors Potential for identifying important categories of behaviors that may not be being considered in the literature <p><i>Cons</i></p> <ul style="list-style-type: none"> Challenging to create Classification model likely to vary across populations and ages Requires covariance among sexual behaviors Challenging to analyze particularly when each category is treated as discrete rather than ordered Results may differ depending on each category Requires larger samples
Data-driven or sample-dependent weights	<p><i>Examples</i></p> <p>Latent variable measurement models (e.g., item response theory, exploratory factor analysis, principal components analysis; (Mattson et al., 2008; Siegel et al., 2001; Williams et al., 2001)</p> <p>Create a weighted composite based on predictive models of a specific outcome (e.g., contracting HIV; (Susser et al., 1998)</p> <p><i>Pros</i></p> <ul style="list-style-type: none"> Simple to use in subsequent analyses Single estimate for treatment response Captures information from the observed variance and covariance among sexual behaviors <p><i>Cons</i></p> <ul style="list-style-type: none"> Challenging to create Measurement model likely to vary across populations and ages Meaning of the underlying dimension may change across studies and populations Requires covariance among sexual behaviors The meaning of the underlying dimension is not clear Requires larger samples 	<p><i>Examples</i></p> <p>Latent class (Beadnell et al., 2005; Hipwell et al., 2011; Lanza & Collins, 2008; Vasilenko et al., 2014)</p> <p>Cluster analysis (Newman & Zimmerman, 2000)</p> <p><i>Pros</i></p> <ul style="list-style-type: none"> Potential for clear mapping of behavior to categories (depends on the quality of the classification model) Uses information from the observed variance and covariance among sexual behaviors Potential for identifying important categories of behaviors that may not be being considered in the literature <p><i>Cons</i></p> <ul style="list-style-type: none"> Challenging to create Classification model likely to vary across populations and ages Requires covariance among sexual behaviors Challenging to analyze particularly when each category is treated as discrete rather than ordered Results may differ depending on each category Requires larger samples

derived approach that captures all the sexual risk behaviors that are targeted by adolescent sexual risk prevention programs. Using a consistent composite index that captures these behaviors could help improve interpretability across studies and populations, but only if the composite values hold the same meaning from one population to the next, which requires replication and cross-validation.

An alternative to using a priori definitions and weights is applying statistical techniques to generate weights or thresholds from the analytic sample. There are numerous strategies for creating sample-dependent weights including ad hoc definitions such as standardizing and then summing rates of behaviors (Fergus et al., 2007; Huang et al., 2012; Wilson & Widom, 2011), unsupervised classification or data reduction algorithms (e.g., latent class analysis, random forest, k-means, latent variable models, principal components analysis), and supervised prediction models where the goal is to predict a particular outcome such as contracting HIV or other STIs (e.g., support vectors, regularized regression, random forest; (Kuhn & Johnson, 2013). Using these approaches can help maximize information in data from any given sample and thus help identify novel and potentially important patterns that may be missed by a priori definitions. Data-driven approaches, however, are also prone to capitalizing on idiosyncrasies of the data to which they are fit, which, if not addressed, can limit generalizability. Protecting against overfitting typically requires cross-validation within a sample, replication within a population, and careful extension to new populations (Kuhn & Johnson, 2013). The time, energy, and sample size requirements for careful and accurate generalization are prohibitive and, thus, often not completed. Consequently, there are very few data-driven or empirically based composites of sexual risk behavior that have been created and validated for use across samples.

There is no consistency in how the various composites are analyzed, with some investigators treating the score as a continuous measure, while others treat the score as categorical. Treating the composite as a continuous risk score assumes a single dimension of sexual risk, which simplifies data summary and analyses. However, a single dimension also obscures the meaning of the composite score due to the potential for different factors predicting the different behaviors that comprise the composite. For example, predictors for higher risk behaviors, such as condomless sex with multiple partners, may differ from factors that predict lower risk behaviors such as decision to have sex or condom-protected sex. These challenges are similar to those of applying latent variable measurement models to sexual risk behavior. While the composite score does not require correlation among indicators, the resulting index lacks an intuitive interpretation, which makes it difficult to discern clinically meaningful change (Denison et al., 2008; Noar, 2008).

Composites can also be treated as categorical. Well-validated classification models provide face-valid, often intuitive, groupings of individuals (e.g., high, moderate, low risk), which facilitates communication of findings to clinicians and policy makers. The intuitive meaning of a well-developed classification comes from highlighting large and typically clinically meaningful transitions between patterns or classes of behavior. For example, transitioning from engaging in condomless sex with multiple partners to consistent condom use with a single partner represents a marked reduction in risk of HIV infection, whereas the benefits of reducing the monthly number of condomless acts from three to two are less clear. However, small changes in risk behaviors may be important and categorical definitions are less sensitive to potentially meaningful change within each class. Treating composites as categorical also introduces additional analytic complexity when examining transitions among multiple classes. To make such analyses interpretable, researchers will often reduce the number of categories to a binary (risk vs. no risk) or focus on transitions that have theoretical relevance such as moves from high-risk to lower risk categories.

Worked Examples

To further illustrate the use of various approaches to calculating a sexual risk composite and to better understand the influences of different approaches on substantive conclusions, we attempted to fit a conventional measurement model with a single continuous latent variable along with three approaches to constructing sexual risk composites using pooled data from four adolescent risk prevention trials for adolescents dealing with mental health challenges.

Method

We used data from four clinical trials of HIV prevention programs designed for adolescents with mental health concerns (R01NR011906; R01MH066641; R01MH63008; and R01MH61149). These programs involved three general approaches to reducing risk: (1) improving HIV prevention skills (condom use, partner negotiation), (2) improving emotion regulation, and (3) improving parent–adolescent communication and parental monitoring. Each trial also included a health promotion control condition that was matched for time and attention to the respective active treatment conditions. The combined dataset included 1735 participants, 1322 of whom reported complete sexual risk data at both baseline and extended follow-up assessment (9–12 months). Demographics for the sample by study are presented in Table 2. Additional details about the studies are included in “Appendix 1” section. We focus our analyses on how findings differ depending on which risk composite is used. Therefore,

Table 2 Baseline characteristics by study

Characteristics mean (SD) or % (n)	Study 1 (n = 117)	Study 2 (n = 345)	Study 3 (n = 276)	Study 4 (n = 584)
Age (in years)	15.66 (1.16)	12.92 (0.54)	15.55 (1.45)	15.39 (1.31)
Female	64% (75)	48% (165)	32% (88)	57% (330)
Hispanic	26% (30)	37% (129)	19% (52)	9% (55)
Race				
African American	16% (19)	29% (99)	28% (78)	58% (339)
American Indian/Alaskan Native	3% (4)	3% (9)	3% (8)	1% (8)
Asian	2% (2)	1% (4)	0% (0)	1% (4)
Multiple Races	16% (19)	16% (56)	14% (38)	10% (56)
Native Hawaiian or other Pacific Islander	3% (3)	3% (11)	2% (5)	1% (4)
White	46% (54)	33% (115)	52% (144)	28% (166)
Mental health diagnosis by category				
No diagnosis	41% (48)	70% (242)	45% (124)	36% (209)
Internalizing	11% (13)	4% (14)	9% (25)	11% (67)
Externalizing	23% (27)	14% (47)	24% (66)	22% (126)
Internalizing and externalizing	17% (20)	4% (15)	10% (28)	14% (84)
Any Mania	5% (6)	8% (27)	9% (25)	14% (84)
Columbia Impairment	15.16 (9.06)	12.15 (8.01)	15.13 (7.75)	16.78 (8.93)
Baseline risk				
Ever had sex	45% (53)	8% (29)	51% (141)	53% (312)
Recent sex (within recall period)	34% (40)	4% (15)	35% (97)	32% (187)
Number of partners, past 3–6 months	0.71 (1.33)	0.10 (0.76)	0.90 (2.58)	0.64 (1.34)
Number of acts	5.29 (15.61)	0.36 (3.57)	11.22 (31.95)	3.5 (10.73)
Number of condomless acts	3.45 (13.36)	0.12 (1.48)	6.09 (23.41)	1.54 (7.29)

subsequent analyses used only complete cases ($n = 1322$) and pooled all active treatment conditions.

Measures

All studies used audio computer-assisted self-interviews to assess adolescents' sexual behaviors. Two of the studies used a recall period of 6 months, and two used a recall period of 3 months. All studies included questions for the following sexual behaviors: (1) ever engaged in vaginal or anal sex, (2) recently engaged in vaginal or anal sex (i.e., during the recall period), (3) number of recent sexual partners, (4) number of total vaginal and/or anal sex acts, and (5) number of condomless acts. The behavioral counts were transformed into rates across 3 months to account of differences in recall periods among studies. Behavioral count questions were open responses. As mentioned previously, adolescents' developing sense of numeracy along with their sensitivity to social desirability may lead to inaccurate and inflated estimates of their behavior. Evidence of potential inflation was seen in the combined dataset where 22 (2%) adolescents reported rates equivalent to one or more sexual events per day in the 3-month period and 23 (2%) reported rates equivalent to more than 23 partners per year. Although such rates are possible, they most likely reflect a significant overestimation of the

behavior. To limit the influence of inflated estimates, open-ended responses were transformed into ordered categorical variables using 1, 2, 5, and 8 as cut points for number of partners and 1, 2, 5, 10, 20, and 30 as cut points for both number of protected acts and number of condomless acts. Each of the four trials also assessed functional impairment due to psychiatric symptoms using the Columbia Impairment Scale (CIS; Bird et al., 1993), with 98% of included participants completing the measure at baseline.

Formative Measurement Model

A latent measurement model with ordinal indicators was fit using number of partners, number of protected acts, and number of condomless acts using Mplus 7.3. The model fitting process was complicated by the zero-inflation in the sexual risk variables that resulted in several ill-behaved models. We first attempted to fit the model with a robust weighted least square estimator (WLSMV). This model produced a nonpositive definite residual covariance matrix, a problem that persisted after simplifying the model to assume equal spacing between categories. We also attempted using a robust maximum likelihood estimator (MLR) with a logit link function, as well as only running the models using participants who were recently active using both WLSMV and MLR

estimators, all with similar difficulties. Finally, we ran the model using a Bayesian framework with semi-informative priors. These models showed poor mixing of the MCMC chains. There may be an analytic approach that would address these challenges, but many readily available tools for a latent measurement model were not successful. It appears that using number of partners, number of protected acts, and number of condomless acts did not provide sufficient information to easily fit a latent measurement model. Indeed, studies that have employed measurement models to summarize sexual risk have used many more behaviors (e.g., carrying a condom, pregnancy, unwanted sex, HIV testing) and have included aspects of the intrapersonal and interpersonal processes driving sexual behavior such as intentions or partner communication (Siegel et al., 2001). Because we were not able to generate a reliable model, we did not include results from this measurement model in subsequent analyses.

Risk Composites

We selected three approaches to forming composites from the literature. Each composite used the same variables as the formative measurement model. The first composite (C1) used procedures similar to those used in previous studies that form a continuous risk score (Fergus et al., 2007; Huang et al., 2012; Wilson & Widom, 2011), where each behavioral count was z-scored using the grand mean and SD from the entire dataset (baseline and extended follow-up), thus preserving differences among assessments, treatment conditions, and studies. The resulting z-scores were then summed to form a composite score for each participant at each assessment. The second composite (C2) classified behavior into the following categories: never engaged in vaginal or anal sex (0), no recent sex (1), only one partner and no condomless acts (2), either multiple partners or any condomless acts (3), and both multiple partners and any condomless acts (4). This approach was also similar to previous studies (Bowleg et al., 2014; Epstein et al., 2014; Graham et al., 2013; Murphy et al., 2009).

Participants were assigned a category for each assessment. The third approach (C3) was to use a mixture model to define classes of participants based on the three ordinal variables used in creating the previous composites. Models were fit using Mplus 7.3. Class enumeration proceeded by fitting ten models while increasing the number of classes from 1 to 10. Each model was estimated using the MLR estimator with a logit link function. Both baseline and extended follow-up assessments were included in the class enumeration process. Nesting of assessment within participant was accounted for using the TYPE = COMPLEX utility. When using TYPE = COMPLEX, only the Bayesian information criterion (BIC) is valid when comparing models. The BIC values indicated that the three- and four-class solutions were similar and both outperformed the other models. The

three-class solution identified a class with no recent sex and two sexually active classes that differed primarily in terms of the amount of condomless sex. The four-class solution also produced a class with no recent sex (No Sex) but separated the sexually active participants into three classes: one with higher numbers of partners, condomless sex, and protected sex ($Prtns_{high}/Sex_{high}$), one with low number of partners and high amounts of condomless sex ($Prtns_{low}/CISex_{high}$), and one with low number of partners and low amounts of condomless sex ($Prtns_{low}/CISex_{low}$). Profiles for the three- and four-class solutions along with class enumeration statistics are listed in “Appendix 2” section. We retained the four-class solution because it provided a more nuanced description of number of partners and condomless sex.

Results

Comparing Risk Composites

The relationship among the three risk composites is depicted in Fig. 1. The association between the two categorical approaches (C2 and C3) was moderate (Cramer’s $V=0.65$) when considering the entire sample and somewhat smaller (Cramer’s $V=0.37$) when considering just the recently active. If it was assumed that the categorical composites are ordered, their associations with the continuously scaled composite (C1) were strong for the full sample (Spearman’s $\rho=0.99$ and 0.86 for C2 and C3, respectively) and somewhat less when only considering those who are sexually active (Spearman’s $\rho=0.61$ and 0.68). These associations indicate that overall the composites seem to be similar, but

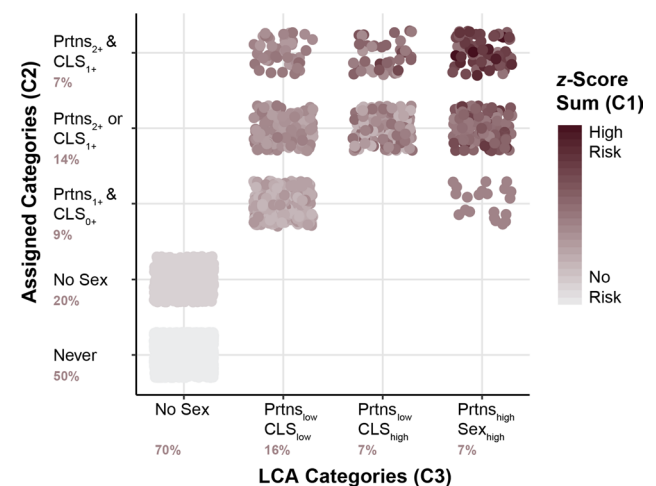


Fig. 1 For the z-score sum composite, behavioral counts were first recoded into categories, z-scored across participants, assessments, intervention, and study, and then summed. Assigned categories were defined a priori. LCA categories were defined using latent class analysis

there are marked differences in how they classify adolescents who were recently active. It is also clear that the transition into sexual activity has a considerable influence on the relationships among composites.

To examine how findings from observational studies are potentially influenced by the choice of composite, we examined the association between each composite and the CIS at the baseline assessment. Models were fit using Mplus 7.3 with the composite as the dependent variable with CIS and study as independent variables. CIS was standardized using the mean and standard deviation of the total sample. A linear model was fit to the *z*-scored composite (C1), and the two categorical composites (C2 and C3) were treated as ordered categorical outcomes. Individual study effects were estimated by including the interaction between CIS and study in the model. To examine possible influence on more complex models, we also fit a set of models that included gender and allowed gender to interact with all dependent variables. All models were estimated using MLR. Regression coefficients were standardized using the standard deviation of the outcome or the latent variable underlying the ordered or binary categories.

Standardized regression estimates by study as well as the moderation of female gender on the relationships between sexual risk and CIS are depicted in Fig. 2. Overall, there was a small positive association between baseline impairment and sexual risk behaviors with stronger associations for females

versus males. Generally, there were only minor differences in these associations based on which composite was used. One important exception was the use of the *z*-scored (C1) composite in Study 2, which had very little sexual behavior at baseline due to the younger age of the participants. These results suggest that the choice of composite may not greatly influence the estimation of associations in observational studies, so long as there is a reasonable amount of risk behavior observed in the sample. If risk behavior is low, it is likely that the categorical composites (C2 and C3) better represent the uncertainty around the associations.

We also examined how choice of composite may influence estimation of treatment effects in clinical trials and on gender as a possible treatment modifier. Each of the composites was defined separately using the baseline and extended follow-up (9–12 months) for all four clinical trials. The extended follow-up was used as the outcome, and each model included baseline, treatment condition, and study. Study-level results were estimated by including the interaction between study and treatment condition. Gender was also included as a potential treatment modifier in a separate set of analyses and allowed to interact with treatment and study. Models were estimated similarly to the CIS analyses. To highlight approaches that treat the categorical composites as nominal data, we included additional definitions for the two categorical composites. Specifically, we classified transitions among classes from baseline to the extended follow-up as

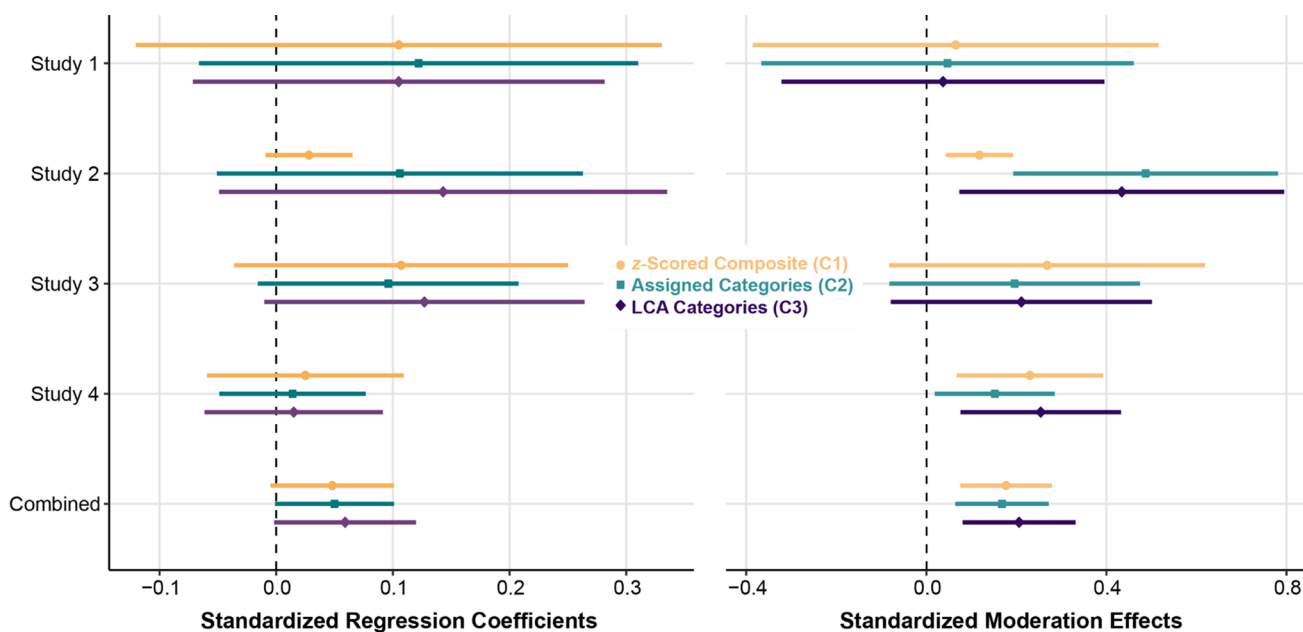


Fig. 2 Associations between the Columbia Impairment Scale (CIS) and sexual risk composites both within and across studies using data from the baseline assessment of the four clinical trials. The combined estimates included study as a fixed effect. Separate models were run that included female gender as a moderator. Moderation effects represent

the difference between female versus male in the association between CIS and the composite measures. Model coefficients were standardized by using the standard deviation of the outcome, or the latent variable underlying the ordered or binary categories

follows: (1) reporting low/decreased risk defined as maintaining a low-risk category (i.e., 0, 1 or 2 for C2; or no sex or $\text{Prtns}_{\text{low}}/\text{CISex}_{\text{low}}$ for C3) at follow-up or moving down a category from their baseline report, versus (2) reporting high/increased risk defined as a reported high-risk category (i.e., 3 or 4 for C2; or $\text{Prtns}_{\text{high}}/\text{Sex}_{\text{high}}$ or $\text{Prtns}_{\text{low}}/\text{CISex}_{\text{high}}$ for C3) at follow-up or moving up a category from baseline. This definition was analyzed as a binary outcome with a logit link function but did not include baseline in the analytic model. Results are depicted in Fig. 3.

None of the composites showed a significant treatment effect; however, the pattern of results differed depending on which composite was used, particularly among the smaller studies, with there being a moderate-to-large difference using the metric of Cohen's d between the highest and lowest point estimates for Study 1. Moreover, the SEs of the z -scored (C1) composite appeared to vary as a function of the proportion of sexually active participants. For example, only 12% of participants in Study 2 reported sexual activity at the extended follow-up assessment and the confidence intervals for the treatment effects varied considerably among the three composite measures for this study.

The between-composite differences were magnified in the gender analyses with Study 2 showing significant improvement for females versus males on two of the composites (C1 and C2) with large differences (i.e., > 0.5) in the point

estimates. There were also large differences among composites for Study 1 and small-to-moderate differences for the remaining two studies and for the combined sample.

Discussion

Creating a composite to summarize sexual risk behavior holds the promise of better understanding treatment effects of prevention programs, particularly for interventions targeting processes that may influence multiple risk behaviors. Although the promise of a composite is recognized in the literature, there is little consistency in how researchers combine dimensions of sexual risk behavior. Results from our worked example suggest that while these approaches may be similar for cross-sectional studies with sufficient variability in risk behaviors, they can influence estimates of treatment effects and their SEs, particularly among studies with smaller samples or low numbers of sexually active participants. These differences are magnified when used to identify treatment modifiers, which complicates the already challenging process of individualizing intervention approaches (Lagakos, 2006). Inconsistencies in how composites are defined further complicate efforts to aggregate findings across trials in a field that is already inconsistent in the assessment, analysis, and reporting of sexual risk behaviors. Improving consistency

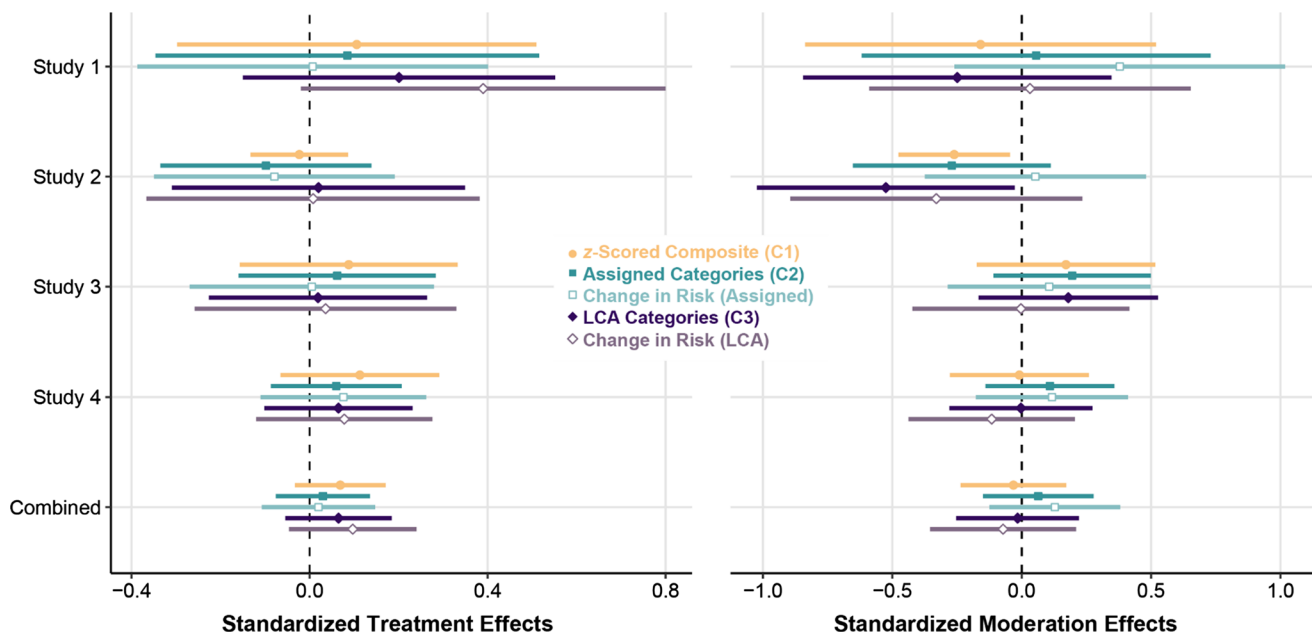


Fig. 3 Treatment effects were estimated within each study and across studies using the extended follow-up data, controlling for baseline. The combined estimates included study as a fixed effect. For both categorical composite definitions, we defined change in risk as follows: low/decreased risk=maintained a low-risk category or moved down a category; high/increased risk=reported a high-risk category

or moved up a category. Separate models were run that included female gender as a moderator. Moderation effects represent the difference between female versus male in the treatment effect. Model coefficients were standardized by using the standard deviation of the outcome, or the latent variable underlying the ordered or binary categories

in how composites are defined requires guidance on how to address the challenges unique to each population of interest.

Challenges to Summarizing Sexual Risk During Adolescence

The worked examples highlight challenges with aggregating data across multiple adolescent sexual risk behaviors. The two principal challenges to producing a single index of sexual risk behaviors are the relatively low proportion of adolescents who engage in sexual risk behavior and the relatively low observed correlations among risk behaviors for those who are sexually active. Even in at-risk samples like the ones used in the worked example, a high proportion of the sample will not have engaged in sexual activity during any given recall period. For example, in the combined dataset, only 59% had ever been sexually active with only 35% recently active. Among those who were active, the associations among number of partners, number of protected acts, and number of condomless acts were low ($\rho_{\text{Prtns/ClSex}} = 0.14$, $\rho_{\text{Prtns/PrSex}} = 0.26$, and $\rho_{\text{PrSex/ClSex}} = -0.25$). These low correlations suggest that each dimension provides unique information and may not be amenable latent variable measurement models. This is not to suggest that these dimensions are unrelated, only that without assessing the processes that link the dimensions (e.g., partner characteristics, knowledge, attitudes, relationship quality, relationship duration, risk propensity), the observed simple association among reported behaviors is small.

In adolescent populations, the low number of recent sexual events and low correlations among different types of sexual risk behaviors pose significant challenges to latent variable measurement models and to data-driven approaches to summarizing across sexual risk behaviors. When considering the full sample, data-driven approaches are primarily driven by recent activity versus no recent activity. When considering only the recently sexual active, the low correlations provide limited information for data-driven approaches to distill and summarize. The limited information among core risk behaviors is likely why studies that have employed latent variable measurement models often included additional behaviors such as alcohol and other drug use prior to sex, carrying a condom, unwanted sex, pregnancy, or STIs, and also included some of the drivers of the risk behaviors such as intentions, attitudes, or partner communication (Siegel et al., 2001). While including more behaviors may help with fitting a measurement model, the additional information may further complicate interpretation of the underlying risk score, particularly when behaviors are mixed with intentions or attitudes.

There are statistical models that explicitly address the high number of zeros. These approaches to zero-altered data (Atkins et al., 2013) jointly estimate the process that generates the zeros or excess zeros along with the process

that generates the behavioral counts. Typically, zero-altered models estimate two processes, but for adolescent behaviors there are at least three: processes influencing the transition into sexual activity, processes influencing recent activity, and processes influencing the behavioral count. Moreover, the count processes (number of partners, condomless sex, protected sex) share the same zero-altered processes (transition into sexual activity, recent sexual activity) and estimating separate models for each count process will produce different estimates for these shared zero-altered processes. Fitting a joint model of the zero-altered processes and each of the count processes is a formidable analytic challenge. Although zero-altered models help improve estimation of any given risk behavior, they can be difficult to analyze with no clear approach to aggregating across the zero-altered and count processes; consequently, they have not been used when forming risk composites (Aicken et al., 2013; Webb et al., 2015).

Recommendations and Future Directions

In this article, we have briefly outlined the most common approaches to forming sexual risk indices and presented worked examples. Results suggested that the choice of index may influence findings from clinical trials but does not indicate which approach is best suited for evaluating adolescent sexual risk behaviors. Part of the difficulty is that there is no clear criterion against which to compare the various approaches to forming a composite. Although it is possible to utilize adverse health outcomes such as STI incidence or unwanted pregnancies to weight the sexual risk behaviors, such models require large datasets with both sexual behaviors and the adverse outcome of interest. Moreover, the models will differ depending on which outcome is used to generate the weights. Ultimately, deciding on a composite approach is not a function of improving measurement accuracy or statistical methodology but of agreeing on common definitions and short of that, consistency in assessment, analysis, and reporting of each sexual risk behavior. We would like to forward some recommendations to help improve the quality and consistency in how sexual risk behavior is reported in the literature.

Adding Context

One of the most challenging aspects of studying adolescent sexual risk behavior is the fluidity of adolescent relationships and low frequency of sexual behavior which result in inconsistent behavior across time and zero-altered data. The high number of zeros makes it difficult to identify treatment effects because during any given assessment period, it is not clear whether an adolescent is actively lowering exposure to sexual risk behavior or has simply not had opportunity to demonstrate their propensity for risk due to lack of partner.

Enriching current assessment with additional context, such as asking about nonsexual relationships or asking about successful avoidance of sexual situations, will help provide additional information about those with low levels of risk behaviors (Manning et al., 2014). A richer assessment will help identify those who are actively avoiding risk from those who have limited opportunity to manifest risk behaviors. Collecting contextual information may require more time-intensive data collection approaches such as structured interviews, timeline followback, daily diaries, or experience sampling. These more time-intensive measures can be integrated with traditional longitudinal designs to add contextual information that will enhance the accuracy and interpretability of the data (Gioia et al., 2012; Sliwinski, 2008). Care is needed, however, when using daily measures as the relatively low frequency of sexual behavior during adolescence might result in high investment of resources and participants' time for limited yield. Studies using daily assessment approaches to study sexual behavior tend to sample from populations that regularly engage in sexual behavior (Blood & Shrier, 2013; Wray et al., 2016) or pool across participants and time to focus on predictors of sexual events (Blood & Shrier, 2013). It is not clear whether prospective daily assessment is a cost-effective approach to assess individual change in sexual risk behaviors over time.

Scaffolding Recall

Given adolescents' developing numeracy and heightened sensitivity to social norms and perceived expectations, it is important to appropriately scaffold their recall of sexual behaviors. Such scaffolding may include replacing open-ended responses with carefully calibrated response categories that have been piloted using cognitive interview techniques to ensure accurate communication of expectations and norms. Timeline followback methods (Weinhardt et al., 1998), which are widely used in the study of alcohol and other drug use, can provide much-needed structure to adolescent recall and can help track relationships overtime, thus providing contextual information about relationship duration and concurrent relationships (Rizzo et al., 2017). Self-administered administration of timeline followback methods has been developed which can help minimize bias due to social desirability and impression management often seen in face-to-face interviews (Collins et al., 2008; Rueger et al., 2012). More work is needed to evaluate the reliability of self-reported timeline followback methods in assessing adolescent sexual risk behaviors (Schroder et al., 2003b), but the strong performance of timeline followback methods across risk outcomes (Hjorthøj et al., 2012; Norberg et al., 2012), populations (Carey et al., 2001; Sobell et al., 2001), and modes of administration (Maisto et al., 2008; Pedersen

et al., 2012; Sobell et al., 1996) argues favorably for increased use in sexual risk prevention trials.

Aggregating Across Assessments

One approach to examining group differences in inconsistent and infrequent behaviors is to lengthen the time frame being considered, thus increasing the opportunity to observe the behavior. Expanding time frames can be done through lengthening the recall window or by aggregating across multiple assessments with shorter recall windows. Because expanding recall windows adversely affects accuracy of the recall, we recommend aggregating multiple shorter recall windows. Although aggregating across assessments helps to minimize the number of zeros in the data, it does so at the expense of temporal precision. Aggregated data do not contain information about when events happened only that they happened. Retaining temporal information requires reducing the time interval over which data are aggregated and reintroduces the challenges of sporadic and zero-altered data. Balancing the relative importance of temporal information versus observing sufficient behavior depends on the frequency of the behavior and the importance of timing to the research question. For clinical trials, *when* adolescents reduced their risk may not be as critical as *if* they reduced their risk following intervention, making aggregating across time a promising approach.

Report Each Sexual Risk Dimension

To improve consistency of data reported in peer-reviewed articles for adolescent sexual risk behavior, it is necessary to report each of the most common risk behaviors, such as age of onset (or at least proportion of the sample that has ever been sexually active), any sexual activity during the period of observation, condomless acts, number of partners, and high-risk behaviors (e.g., substance use before sex, sex with someone you just met). Reporting how an intervention influenced each risk behavior enables comparisons among trials, even if each study used different strategies for forming risk composites. Routinely reporting each behavior will also protect against potential bias introduced by selectively choosing which behaviors to report.

Develop Standard Sexual Risk Composites

Significant work is needed to develop and validate standard sexual risk composites. As stated previously, the work of developing a common sexual risk composite is ultimately a measurement problem that requires building consensus around what is considered sexual risk behavior. As a field, we need to answer some fundamental measurement questions to find consensus in how to form a sexual risk composite. These questions include the following: (1) Should

the underlying risk propensity be one-dimensional or categorical? (2) Should the model be calibrated to predict a specific health outcome such as STI incidence, or should it be a more general model of sexual risk behavior? (3) Should the model be strictly behavioral or should it include thoughts, attitudes, and beliefs? (4) Which behaviors should be routinely included as sexual risk behaviors? Note that these questions do not include those specific to the assessment of each risk behavior which are covered in several excellent reviews (Dolezal et al., 2012; Fenton, 2001; Fonner et al., 2013; Graham et al., 2005; Schroder et al., 2003a, b).

Statistical simulation might be able to assist in addressing questions about sensitivity to change. For example, it would be helpful to know how each composite responds to changes in one or more of the contributing behaviors. Understanding sensitivity to change could be evaluated by generating a “true” model for sexual life-history data including age of onset, number of relationships, relationship durations, number of sexual acts, number of condomless acts, and number of high-risk behaviors. Various competing composites could then be formed using these simulated data. Manipulating how the life-history data changes post-intervention would help inform the sensitivity of each composite to changes in components of the life history.

Agreement on a set of common composites will require more effort in harmonizing definitions which will take time; in the interim, researchers will continue to use composites. Beyond reporting each dimension of sexual risk, we recommend using a clearly defined, a priori classification of risk behavior. The categorical definition is easily defined and has high face validity that facilitates clear communication of results from one study to the next. When coupled with reports of each dimension, a theoretically defined, categorical risk composite will provide information about treatment changes in overall sexual risk behavior. Including this common composite definition along with outcomes for each risk

dimension would provide a foundation that would greatly facilitate comparisons across studies. Consistent reporting will enhance meta-analytic efforts to summarize the effectiveness of current prevention approaches, which have been limited by inconsistent reporting of changes in adolescent sexual risk behaviors following prevention trials.

Acknowledgements This work was supported by the National Institute of Mental Health (K23MH102131 to David H. Barker) and by the National Institute on Alcohol Abuse and Alcoholism (R01AA021355 to Lori A. J. Scott-Sheldon) of the National Institutes of Health. The work was facilitated by the Providence/Boston Center for AIDS Research (P30AI042853). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding This study was funded by the following grants: National Institute of Mental Health (K23MH102131); *National Institute* on Alcohol Abuse and Alcoholism (R01AA021355), and National Institute of Allergy and Infectious Diseases (P30 AI042853).

Compliance with Ethical Standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all parents/caregivers, and assent was obtained from all participants under the age of 18.

Appendix 1: Details About Treatment Studies

See Table 3.

Table 3 Overview of prevention trials

FED ID#	Age range (years)	Population	Treatment condition	Sample	Assessment schedule (in months)				
					0	3	6	9	12
Study 1 (STAR) R01MH61149	12–18	Youth in therapeutic school settings	Health promotion	89	•	•	•		
			Affect management	96	•		•	•	
Study 2 (TRAC-2) R01NR011906	12–14	Seventh graders with significant mood or conduct symptoms	Health promotion	222	•	•	•		•
			Affect management	198	•	•	•		•
Study 3 (BALANCE) R01MH066641	15–18	Youth in therapeutic school settings	Health promotion	124	•	•	•		•
			Affect management	157	•	•	•		•
			Skills training	136	•	•	•		•
Study 4 (STYLE) R01MH63008	13–18	Youth in mental health treatment	Health promotion	235	•	•	•		•
			Skills training	259	•	•	•		•
			Family based	227	•	•	•		•

Appendix 2: Class Enumeration Results

See Tables 4, 5, 6, 7, and 8.

Table 4 Class enumeration fit statistics used to decide on the number of classes

Model	Entropy	Negative log likelihood	Number of parameters	BIC	Smallest class size
1	NA	−9034.432	17	18,203.92	2820
2	1	−6471.562	35	13,221.18	818
3	0.999	−6338.384	53	13,097.83	168
4	0.937	−6275.599	71	13,115.26	236
5	0.953	−6251.629	89	13,210.32	97
6	0.937	−6237.301	107	13,324.66	96
7	0.965	−6215.517	125	13,424.1	44
8	0.962	−6209.702	143	13,555.47	24
9	0.951	−6193.232	161	13,665.53	35
10	0.815	−6193.037	179	13,808.14	42

Table 5 Estimated means for each sexual risk behavior by class (three-class model)

	Latent classes (three-class model)		
	Class 1	Class 2	Class 3
<i>Binary outcomes</i>			
Probability (95% CI)			
Vaginal/anal ever	1.00	1.00	0.28 (0.26; 0.30)
Vaginal/anal past 90 days	1.00	1.00	0.00
<i>Behavioral counts</i>			
Monthly rate (95% CI)			
Vaginal/anal acts	4.35 (4.27; 4.43)	4.81 (4.64; 4.98)	0.00
Vaginal/anal condomless acts	1.73 (1.68; 1.78)	4.81 (4.64; 4.98)	0.00
Vaginal/anal protected acts	2.69 (2.62; 2.75)	0.00	0.00
Vaginal/anal partners	0.53 (0.51; 0.56)	0.39 (0.34; 0.44)	0.00

Table 6 Average latent class probabilities for most likely latent class membership (3-class model)

	Most likely latent class membership		
	Class 1	Class 2	Class 3
Class 1	0.999	0.001	0.000
Class 2	0.000	1.000	0.000
Class 3	0.000	0.000	1.000

Table 7 Estimated means for each sexual risk behavior by class (four-class model)

	Latent classes (four-class model)			
	HPRT HSEX	LPRT HCLS	LPRT LCLS	NOSEX
<i>Binary outcomes</i>				
Probability (95% CI)				
Vaginal/anal ever	1.00	1.00	1.00	0.28 (0.26; 0.30)
Vaginal/anal past 90 days	1.00	1.00	1.00	0.00
<i>Behavioral counts</i>				
Monthly rate (95% CI)				
Vaginal/anal acts	11.08 (10.84; 11.32)	5.31 (5.14; 5.47)	1.20 (1.15; 1.25)	0.00
Vaginal/anal condomless acts	4.48 (4.33; 4.63)	5.25 (5.08; 5.41)	0.17 (0.15; 0.19)	0.00
Vaginal/anal protected acts	6.83 (6.64; 7.02)	0.06 (0.04; 0.08)	1.03 (0.99; 1.08)	0.00
Vaginal/anal partners	0.99 (0.92; 1.06)	0.40 (0.36; 0.45)	0.34 (0.31; 0.37)	0.00

Table 8 Average latent class probabilities for most likely latent class membership (four-class model)

	Most likely latent class membership			
	Class 1	Class 2	Class 3	Class 4
Class 1	0.893	0.089	0.018	0.000
Class 2	0.015	0.985	0.000	0.000
Class 3	0.133	0.073	0.794	0.000
Class 4	0.000	0.000	0.000	1.000

References

- Aicken, C. R. H., Gray, M., Clifton, S., Tanton, C., Field, N., Sonnenberg, P., et al. (2013). Improving questions on sexual partnerships: Lessons learned from cognitive interviews for Britain's Third National Survey of Sexual Attitudes and Lifestyles ("Natsal-3"). *Archives of Sexual Behavior, 42*, 173–185.
- Ashenhurst, J. R., Wilhite, E. R., Harden, K. P., & Fromme, K. (2017). Number of sexual partners and relationship status are associated with unprotected sex across emerging adulthood. *Archives of Sexual Behavior, 46*, 419–432.
- Atkins, D. C., Baldwin, S. A., Zheng, C., Gallop, R. J., & Neighbors, C. (2013). A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychology of Addictive Behaviors, 27*, 166–177.
- Bancroft, J., Janssen, E., Strong, D., Carnes, L., Vukadinovic, Z., & Long, J. S. (2003). Sexual risk-taking in gay men: The relevance of sexual arousability, mood, and sensation seeking. *Archives of Sexual Behavior, 32*, 555–572.
- Beadnell, B., Morrison, D. M., Wilsdon, A., Wells, E. A., Murowchick, E., Hoppe, M., ... Nahom, D. (2005). Condom use, frequency of sex, and number of partners: Multidimensional characterization of adolescent sexual risk-taking. *Journal of Sex Research, 42*, 192–202.
- Bird, H. R., Shaffer, D., Fisher, P., & Gould, M. S. (1993). The Columbia Impairment Scale (CIS): Pilot findings on a measure of global impairment for children and adolescents. *International Journal of Methods in Psychiatric Research, 3*, 167–176.
- Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience, 9*, 267.
- Blood, E. A., & Shrier, L. A. (2013). The temporal relationship between momentary affective states and condom use in depressed adolescents. *Archives of Sexual Behavior, 42*, 1209–1216.
- Boekeloo, B. O., Schamus, L. A., Simmens, S. J., Cheng, T. L., O'Connor, K., & D'Angelo, L. J. (1999). A STD/HIV prevention trial among adolescents in managed care. *Pediatrics, 103*, 107–115.
- Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods, 16*, 265–284.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110*, 203–219.
- Bowleg, L., Neilands, T., Tabb, L. P., Burkholder, G. J., Malebranche, D. J., & Tschann, J. M. (2014). Neighborhood context and Black heterosexual men's sexual HIV risk behaviors. *AIDS and Behavior, 18*, 2207–2218.
- Brown, L. K., Nugent, N. R., Houck, C. D., Lescano, C. M., Whiteley, L. B., Barker, D., ... Zlotnick, C. (2011). Safe Thinking and Affect Regulation (STAR): Human immunodeficiency virus prevention in alternative/therapeutic schools. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*, 1065–1074.
- Carey, M. P., Carey, K. B., Maisto, S. A., Gordon, C. M., Schroder, K. E. E., & Venable, P. A. (2004). Reducing HIV-risk behavior among adults receiving outpatient psychiatric treatment: Results from a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 72*, 252–268.
- Carey, M. P., Carey, K. B., Maisto, S. A., Gordon, C. M., & Weinhardt, L. S. (2001). Assessing sexual risk behaviour with the Timeline Followback (TLFB) approach: Continued development and psychometric evaluation with psychiatric outpatients. *International Journal of STD and AIDS, 12*, 365–375.
- Carmona, J., Slesnick, N., Guo, X., & Letcher, A. (2014). Reducing high risk behaviors among street living youth: Outcomes of an integrated prevention intervention. *Children and Youth Services Review, 43*, 118–123.
- Carver, K., Joyner, K., & Udry, J. R. (2003). National estimates of adolescent romantic relationships. In P. Florsheim (Ed.), *Adolescent romantic relations and sexual behavior: Theory, research, and practical implications* (pp. 23–56). Mahwah, NJ: Lawrence Erlbaum.
- Collins, R. L., Kashdan, T. B., Koutsky, J. R., Morsheimer, E. T., & Vetter, C. J. (2008). A self-administered timeline followback to measure variations in underage drinkers' alcohol intake and binge drinking. *Addictive Behaviors, 33*, 196–200.

- Connolly, J., & McIsaac, C. (2009). Adolescents' explanations for romantic dissolutions: A developmental perspective. *Journal of Adolescence*, *32*, 1209–1223.
- Cooper, M. L. (2010). Toward a person \times situation model of sexual risk-taking behaviors: Illuminating the conditional effects of traits across sexual situations and relationship contexts. *Journal of Personality and Social Psychology*, *98*, 319–341.
- 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*, *18*, 187–203.
- Crosby, G. M., Stall, R. D., Paul, J. P., Barrett, D. C., & Midanik, L. T. (1996). Condom use among gay/bisexual male substance abusers using the timeline follow-back method. *Addictive Behaviors*, *21*, 249–257.
- Denison, J. A., O'Reilly, K. R., Schmid, G. P., Kennedy, C. E., & Sweat, M. D. (2008). HIV voluntary counseling and testing and behavioral risk reduction in developing countries: A meta-analysis, 1990–2005. *AIDS and Behavior*, *12*, 363–373.
- DiClemente, R. J., Wingood, G. M., Rose, E. S., Sales, J. M., Lang, D. L., Caliendo, A. M., ... Crosby, R. A. (2009). Efficacy of sexually transmitted disease/human immunodeficiency virus sexual risk-reduction intervention for African American adolescent females seeking sexual health services: A randomized controlled trial. *Archives of Pediatrics and Adolescent Medicine*, *163*, 1112–1121.
- Dolezal, C., Marhefka, S. L., Santamaria, E. K., Leu, C.-S., Brackis-Cott, E., & Mellins, C. A. (2012). A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Archives of Sexual Behavior*, *41*, 401–410.
- Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research Methods*, *14*, 370–388.
- Epstein, M., Bailey, J. A., Manhart, L. E., Hill, K. G., & Hawkins, J. D. (2014). Sexual risk behavior in young adulthood: Broadening the scope beyond early sexual initiation. *Journal of Sex Research*, *51*, 721–730.
- Falasinu, T., Gilbert, M., Hottes, T. S., Gustafson, P., Ogilvie, G., & Shoveller, J. (2015). Predictors identifying those at increased risk for STDs: A theory-guided review of empirical literature and clinical guidelines. *International Journal of STD and AIDS*, *26*, 839–851.
- Fenton, K. A. (2001). Measuring sexual behaviour: Methodological challenges in survey research. *Sexually Transmitted Infections*, *77*, 84–92.
- Fergus, S., Zimmerman, M. A., & Caldwell, C. H. (2007). Growth trajectories of sexual risk behavior in adolescence and young adulthood. *American Journal of Public Health*, *97*, 1096–1101.
- Fonner, V. A., Kennedy, C. E., O'Reilly, K. R., & Sweat, M. D. (2013). Systematic assessment of condom use measurement in evaluation of HIV prevention interventions: Need for standardization of measures. *AIDS and Behavior*, *18*, 2374–2386.
- Gioia, C. J., Sobell, L. C., Sobell, M. B., & Simco, E. R. (2012). Shorter and proximal timeline followback windows are representative of longer posttreatment functioning. *Psychology of Addictive Behaviors*, *26*, 880–887.
- Graham, C. A., Crosby, R. A., Sanders, S. A., & Yarber, W. L. (2005). Assessment of condom use in men and women. *Annual Review of Sex Research*, *16*, 20–52.
- Graham, S. M., Raboud, J., McClelland, R. S., Jaoko, W., Ndinya-Achola, J., Mandaliya, K., ... Bayoumi, A. M. (2013). Loss to follow-up as a competing risk in an observational study of HIV-1 incidence. *PLoS ONE*, *8*(3), e59480. <https://doi.org/10.1371/journal.pone.0059480>.
- Herbenick, D., Reece, M., Schick, V., Sanders, S. A., Dodge, B., & Fortenberry, J. D. (2010). Sexual behavior in the United States: Results from a national probability sample of men and women ages 14–94. *Journal of Sexual Medicine*, *7*, 255–265.
- Heywood, W., Patrick, K., Smith, A. M. A., & Pitts, M. K. (2015). Associations between early first sexual intercourse and later sexual and reproductive outcomes: A systematic review of population-based data. *Archives of Sexual Behavior*, *44*, 531–569.
- Hipwell, A. E., Stepp, S. D., Keenan, K., Chung, T., & Loeber, R. (2011). Parsing the heterogeneity of adolescent girls' sexual behavior: Relationships to individual and interpersonal factors. *Journal of Adolescence*, *34*, 589–592.
- Hjorthøj, C. R., Hjorthøj, A. R., & Nordentoft, M. (2012). Validity of timeline follow-back for self-reported use of cannabis and other illicit substances—Systematic review and meta-analysis. *Addictive Behaviors*, *37*, 225–233.
- Huang, D. Y. C., Lanza, H. I., Murphy, D. A., & Hser, Y. I. (2012). Parallel development of risk behaviors in adolescence: Potential pathways to co-occurrence. *International Journal of Behavioral Development*, *36*, 247–257.
- Huebner, A. J., & Howell, L. W. (2003). Examining the relationship between adolescent sexual risk-taking and perceptions of monitoring, communication, and parenting styles. *Journal of Adolescent Health*, *33*, 71–78.
- Jemmott, J. B., 3rd, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics and Adolescent Medicine*, *164*, 152–159.
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. New York: Springer.
- Lagakos, S. W. (2006). The challenge of subgroup analyses—Reporting without distorting. *New England Journal of Medicine*, *354*, 1667–1669.
- Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. *Developmental Psychology*, *44*, 446–456.
- Le Pont, F., Pech, N., & Boelle, P. Y. (2003). A new scale for measuring dynamic patterns of sexual partnership and concurrency: Application to three French Caribbean regions. *Sexually Transmitted Diseases*, *30*, 6–9.
- Lescano, C. M., Vazquez, E. A., Brown, L. K., Litvin, E. B., & Pugatch, D. (2006). Condom use with “casual” and “main” partners: What's in a name? *Journal of Adolescent Health*, *39*, 443.e1–443.e7.
- Li, X., Stanton, B., & Feigelman, S. (2000). Impact of perceived parental monitoring on adolescent risk behavior over 4 years. *Journal of Adolescent Health*, *27*, 49–56.
- Lightfoot, M. (2012). HIV prevention for adolescents: Where do we go from here? *American Psychologist*, *67*, 661–671.
- Maisto, S. A., Conigliaro, J. C., Gordon, A. J., McGinnis, K. A., & Justice, A. C. (2008). An experimental study of the agreement of self-administration and telephone administration of the timeline followback interview. *Journal of Studies on Alcohol and Drugs*, *69*, 468–471.
- Manning, W. D., Longmore, M. A., Copp, J., & Giordano, P. C. (2014). The complexities of adolescent dating and sexual relationships: Fluidity, meaning(s), and implications for young adults' well-being. *New Directions for Child and Adolescent Development*, *144*, 53–69.
- Markham, C. M., Lormand, D., Gloppen, K. M., Peskin, M. F., Flores, B., Low, B., & House, L. D. (2010). Connectedness as a predictor of sexual and reproductive health outcomes for youth. *Journal of Adolescent Health*, *46*, S23–S41.
- Mattson, C. L., Campbell, R. T., Karabatsos, G., Agot, K., Ndinya-Achola, J. O., Moses, S., & Bailey, R. C. (2008). Scaling sexual behavior or “sexual risk propensity” among men at risk for HIV in Kisumu, Kenya. *AIDS and Behavior*, *14*, 162–172.

- McAuliffe, T. L., DiFranceisco, W., & Reed, B. R. (2010). Low numeracy predicts reduced accuracy of retrospective reports of frequency of sexual behavior. *AIDS and Behavior, 14*, 1320–1329.
- Mullen, P. D., Ramirez, G., Strouse, D., Hedges, L. V., & Sogolow, E. (2002). Meta-analysis of the effects of behavioral HIV prevention interventions on the sexual risk behavior of sexually experienced adolescents in controlled studies in the United States. *Journal of Acquired Immune Deficiency Syndromes, 30*, S94–S105.
- Murphy, D. A., Brecht, M.-L., Herbeck, D. M., & Huang, D. (2009). Trajectories of HIV risk behavior from age 15 to 25 in the national longitudinal survey of youth sample. *Journal of Youth and Adolescence, 38*, 1226–1239.
- Murry, V. M., Berkel, C., Chen, Y., Brody, G. H., Gibbons, F. X., & Gerrard, M. (2011). Intervention induced changes on parenting practices, youth self-pride and sexual norms to reduce HIV-related behaviors among rural African American youths. *Journal of Youth and Adolescence, 40*, 1147–1163.
- Mustanski, B. (2008). Moderating effects of age on the alcohol and sexual risk taking association: An online daily diary study of men who have sex with men. *AIDS and Behavior, 12*, 118–126.
- 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.
- Newman, P. A., & Zimmerman, M. A. (2000). Gender differences in HIV-related sexual risk behavior among urban African American youth: A multivariate approach. *AIDS Education and Prevention, 12*, 308–325.
- Noar, S. M. (2008). Behavioral interventions to reduce HIV-related sexual risk behavior: Review and synthesis of meta-analytic evidence. *AIDS and Behavior, 12*, 335–353.
- Norberg, M. M., Mackenzie, J., & Copeland, J. (2012). Quantifying cannabis use with the timeline followback approach: A psychometric evaluation. *Drug and Alcohol Dependence, 121*, 247–252.
- Pearlman, D. N., Camberg, L., Wallace, L. J., Symons, P., & Finison, L. (2002). Tapping youth as agents for change. *Journal of Adolescent Health, 31*, 31–39.
- Pedersen, E. R., Grow, J., Duncan, S., Neighbors, C., & Larimer, M. E. (2012). Concurrent validity of an online version of the timeline followback assessment. *Psychology of Addictive Behaviors, 26*, 672–677.
- Pequegnat, W., Hartwell, T. D., Green, A. M., Strader, L. C., & Group, N. I. of M. H. C. H. P. T. (2016). How many sexual partners of an individual need to be evaluated to capture HIV/STI risk behavior in a study? *AIDS and Behavior, 20*, 1353–1359.
- Raffaelli, M., & Crockett, L. J. (2003). Sexual risk taking in adolescence: The role of self-regulation and attraction to risk. *Developmental Psychology, 39*, 1036–1046.
- Reece, M., Herbenick, D., Schick, V., Sanders, S. A., Dodge, B., & Fortenberry, J. D. (2010). Condom use rates in a national probability sample of males and females ages 14 to 94 in the United States. *Journal of Sexual Medicine, 7*, 266–276.
- Rizzo, C. J., Joppa, M. C., Barker, D., Zlotnick, C., Warren, J., Saint-Eloi Cadely, H., & Brown, L. K. (2017). Individual and relationship characteristics of adolescent girls with histories of physical dating violence. *Journal of Interpersonal Violence*. <https://doi.org/10.1177/0886260517696859>.
- Rueger, S. Y., Trela, C. J., Palmeri, M., & King, A. C. (2012). Self-administered web-based timeline followback procedure for drinking and smoking behaviors in young adults. *Journal of Studies on Alcohol and Drugs, 73*, 829–833.
- Schroder, K. E. E., Carey, M. P., & Vanable, P. A. (2003a). Methodological challenges in research on sexual risk behavior: I. Item content, scaling, and data analytical options. *Annals of Behavioral Medicine, 26*, 76–103.
- Schroder, K. E. E., Carey, M. P., & Vanable, P. A. (2003b). Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Annals of Behavioral Medicine, 26*, 104–123.
- Scott-Sheldon, L. A. J., Kalichman, S. C., & Carey, M. P. (2010). Assessment of sexual behavior. In A. Steptoe (Ed.), *Handbook of behavioral medicine* (pp. 59–72). New York, NY: Springer.
- Senn, T. E., Scott-Sheldon, L. A. J., & Carey, M. P. (2014). Relationship-specific condom attitudes predict condom use among STD clinic patients with both primary and non-primary partners. *AIDS and Behavior, 18*, 1420–1427.
- Siegel, D. M., Aten, M. J., & Enaharo, M. (2001). Long-term effects of a middle school—And high school-based human immunodeficiency virus sexual risk prevention intervention. *Archives of Pediatrics and Adolescent Medicine, 155*, 1117–1126.
- Slater, C., & Robinson, A. J. (2014). Sexual health in adolescents. *Clinics in Dermatology, 32*, 189–195.
- Sliwinski, M. J. (2008). Measurement-burst designs for social health research. *Social and Personality Psychology Compass, 2*, 245–261.
- Sobell, L. C., Agrawal, S., Annis, H., Ayala-Velazquez, H., Echeverria, L., Leo, G. I., ... Zióikowski, M. (2001). Cross-cultural evaluation of two drinking assessment instruments: Alcohol timeline followback and inventory of drinking situations. *Substance Use and Misuse, 36*, 313–331.
- Sobell, L. C., Brown, J., Leo, G. I., & Sobell, M. B. (1996). The reliability of the alcohol timeline followback when administered by telephone and by computer. *Drug and Alcohol Dependence, 42*, 49–54.
- Somerville, L. H. (2013). The teenage brain: Sensitivity to social evaluation. *Current Directions in Psychological Science, 22*, 121–127.
- Staras, S. A. S., Livingston, M. D., Maldonado-Molina, M. M., & Komro, K. A. (2013). The influence of sexual partner on condom use among urban adolescents. *Journal of Adolescent Health, 53*, 742–748.
- Susser, E., Desvarieux, M., & Wittkowski, K. M. (1998). Reporting sexual risk behavior for HIV: A practical risk index and a method for improving risk indices. *American Journal of Public Health, 88*, 671–674.
- 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. *Journal of Adolescent Health, 46*, 169–179.
- Vasilenko, S. A., Kugler, K. C., Butera, N. M., & Lanza, S. T. (2014). Patterns of adolescent sexual behavior predicting young adult sexually transmitted infections: A latent class analysis approach. *Archives of Sexual Behavior, 44*, 705–715.
- Webb, M. C., Chaney, J. D., Chen, W. W., Dodd, V. J., Huang, I.-C., & Sanders, S. (2015). Assessing specific sexual behavior: Instrument development and validation techniques. *International Journal of Education and Social Science, 2*, 1–11.
- Weinhardt, L. S., Carey, M. P., Maisto, S. A., Carey, K. B., Cohen, M. M., & Wickramasinghe, S. M. (1998). Reliability of the timeline follow-back sexual behavior interview. *Annals of Behavioral Medicine, 20*, 25–30.
- Williams, M., McCoy, H. V., Bowen, A., Saunders, L., Freeman, R., & Chen, D. (2001). An evaluation of a brief HIV risk reduction intervention using empirically derived drug use and sexual risk indices. *AIDS and Behavior, 5*, 31–43.
- Wilson, H. W., & Widom, C. S. (2011). Pathways from childhood abuse and neglect to HIV-risk sexual behavior in middle adulthood. *Journal of Consulting and Clinical Psychology, 79*, 236–246.
- Wray, T. B., Kahler, C. W., & Monti, P. M. (2016). Using ecological momentary assessment (EMA) to study sex events among very high-risk men who have sex with men (MSM). *AIDS and Behavior, 20*, 2231–2242.
- Wu, Y., Burns, J. J., Stanton, B. F., Li, X., Harris, C. V., Galbraith, J., & Wei, L. (2005). Influence of prior sexual risk experience on response to intervention targeting multiple risk behaviors among adolescents. *Journal of Adolescent Health, 36*, 56–63.