A Psychometric Evaluation of the Core Baseline Questionnaire Used in the

Oregon Youth Development Project

Claudia G. Vincent
University of Oregon

Lan Xue
Oregon State University

Tary J. Tobin
University of Oregon

Rose Fuller
Northwest Family Services

ACKNOWLEDGMENTS

Correspondence concerning this paper should be addressed to Claudia Vincent, Ph.D., College of Education, University of Oregon, Eugene, OR 97403. Email: clavin@uoregon.edu

Abstract

This paper focuses on the challenges of measuring youth behaviors associated with onset of sexual activity. We examine the psychometric properties of the Adolescent Family Life Prevention Programs Core Baseline Questionnaires administered within the context of a pilot study for a youth development project that assessed the effect of a classroom intervention intended to delay sexual initiation of middle school students. A total of 608 8th graders participating in the project completed the Core Baseline Questionnaire in the fall of 2009. Following the *Standards for Educational and Psychological Testing* (1999) put forth by the American Educational Research Association, we examine evidence based on test content, scale reliability, and construct validity. Results indicate that the existing instruments from which items on the Core Baseline Questionnaire were derived had limited documented psychometric properties. High Cronbach alpha values and narrow 95% confidence intervals around subscale means indicated good reliability of subscales. Items measuring social environments, parental communication, and access to adult support clustered into well-defined constructs, while items measuring sexual activity and intentions appeared to align less easily into identifiable factors.

A Psychometric Evaluation of the Core Baseline Questionnaire Used in the Oregon Youth

Development Project

It has been repeatedly demonstrated that teenage sexual activity is associated with a number of risks and undesirable outcomes. The most prominent immediate risks include sexually transmitted diseases and unintended pregnancy (Carter-Jessop, Franklin, Heath, Jimenez-Irizarry, & Peace, 2000; Genuis & Genuis, 2004; Kirby & Laris, 2009). Teen pregnancy and parenthood have been associated with academic failure, school drop-out, delinquency, and increased risk of incarceration (Allen, Philliber, Herrling, & Kupermine, 1997; Unruh, Bullis, & Yovanoff, 2004). In addition to the immediate negative effects on individual lives, these outcomes also have long-term and large-scale consequences as they represent lost earning potentials and tax health care systems and social services (Genuis & Genuis, 2004).

Interventions designed to address the potentially far-reaching problem of teenage sexual activity commonly blend strategies designed to increase youth's knowledge of risky sexual behaviors, self-regulation skills, and resiliency, as well as enhance parent-adolescent communication (Card, Lessard, & Benner, 2007; Carter-Jessop et al., 2000; Coley, Votruba-Drzal, & Schindler, 2009; Denny, Young, Rausch, & Spear, 2002; DuRant et al., 2008; Lederman, Chann, & Roberts-Gray, 2008; Moore & Sugland, 1997; Weed, Ericksen, Lewis, Grant & Wibberly, 2008; White, 1996). Establishing trusting relationships with adult and peer mentors has also been identified as an important intervention component (Moore & Sugland, 1997; Sipe, 2002). To maximize the effectiveness of these intervention components for the target population, Greenberg et al. (2003) recommend carefully coordinated integration of these strategies into school operations and activities to create environments supporting positive youth development.

Evaluation of the effectiveness of individual intervention components as well as entire youth development programs has been identified as a critical need to guide best practice as well as increase accountability to stakeholders (Coffman, Guerin, & Gottfried, 2006; Greenberg et al., 2003; Moore & Sugland, 1997). Instruments used to evaluate the effectiveness of programs should not only be easy to administer, but also have known psychometric properties. The interpretability of evaluation data depends very much on the reliability and validity of the instruments used within a given context and with a given population (Greenberg et al., 2003; Thompson et al., 2005). In addition, psychometric information about instruments used to measure teen attitudes about sexual activity could advance theoretical frameworks on which interventions are built (Gorsuch, 1983). Given that teen attitudes about sexual activity are influenced by a myriad of biological, social, and psychological factors, knowledge of the dimensionality and factor structure of measures might provide useful clues about how teens respond to these multiple factors, and how interventions might be best designed to shape those responses to produce maximally positive outcomes.

The Oregon Youth Development Project (OYDP) conducted by Northwest Family Services (NWFS) in Portland, Oregon, and funded by the Office of Adolescent Pregnancy Prevention (OAPP) combines the intervention components identified in the literature to encourage 8th grade students to delay sexual initiation and activity through heightened awareness of healthy life choices, the risks of engaging in premarital sex, the benefits of choosing abstinence, and increased parent-adolescent communication. Within the context of a randomized controlled trial, the OYDP tests the effectiveness of a standard intervention compared to an enriched intervention. The curriculum for the standard intervention consists, in part, of a series of lessons delivered by public school personnel in health education classes on topics including self-

respect, managing peer pressure, healthy friendships and relationships, and sexually transmitted

diseases. To supplement these teacher-delivered lessons, Youth Educators from NWFS provide

additional classroom lessons, using an interactive style, while high school peer leaders serve as

mentors and role models, present socio-dramas, and engage middle schools students in

presentations and discussions. The curriculum of the enriched intervention adds a parent

education component to the standard intervention protocol. The parent education component

encourages parents to engage their child in conversations about teen sexual activity and

abstinence.

The effectiveness of these intervention components is evaluated through the *CORE*

Prevention Questionnaire—Baseline (CQ-B) completed by students prior to intervention

delivery and the *CORE Prevention Questionnaire—Follow-up (*CQ-F*)* completed by students

after intervention delivery. Both questionnaires were developed by the federal office of

Adolescent Family Life (AFL) and are posted at *http://www.hhs.gov/opa/familylife/core_*

instruments/index.html. Projects focused on teen pregnancy prevention funded through AFL and

potentially informing program decisions at the school and community level are required to use

these instruments. Therefore, an examination of their psychometric properties appears warranted.

The CQ-B consists of a total of 43 items arranged into the following 7 subscales: (a) activities (1

item), (b) what you think, (2 items) (c) your family (11 items), (d) what young people think and

do (5 items), (e) about you (4 items), (f) more questions (6 items), and (g) what you think and do

(14 items). The CQ-F consists of a total of 35 items, arranged into the following 6 subscales: (a)

activities (1 item), (b) what you think, (2 items) (c) your family (11 items), (d) what young

people think and do (2 items), (e) more questions (5 items), and (f) what you think and do (14

items). Subscale "about you" is omitted from the follow-up questionnaire, because demographic

information about the respondents does not change between survey administrations.

Some items that were on the CQ-B were not on the CQ-F. On the CQ-F subscale "what young people think and do," the items assessing respondents' ability to say "no" to risky activities, ability to distance themselves from others engaging in risky behavior, and frequency of tobacco, alcohol, and marijuana use are omitted. On the CQ-F subscale "more questions," the item assessing friends' perceived sexual behavior is omitted. In addition, the item asking respondents to record the degree to which they agree with common teen opinions omits the opinion "Having a good marriage seems realistic to me."

The subscales of the CQ-B and CQ-F reflect potential aspects of the overall construct of "teen attitude towards sexuality." These aspects appear to be "activities," "future plans," "family relationships and rules," "responses to peer pressure," "self-awareness," and "sexual behaviors." These subscales might provide initial guidance for psychometric analysis. Psychometric testing might start at the item level and assess items' relevance to the overall construct of "teen attitudes towards sexuality" based on available literature, proceed to examining the cohesion of those items into the subscales represented on the measures as well as the total scale, and the number of latent constructs contained in the total scale based on items' individual factor loadings.

Currently in its second year (2009-2010), the OYDP is now engaged in a randomized controlled trial testing the effectiveness of the standard intervention against the effectiveness of the enriched intervention. Within the context of this trial, the CQ-B and CQ-F were administered to 8th graders in participating middle schools. The goal of our study is to examine (a) the psychometric properties of the instruments from which items for the CQ-B originated, (b) the scale reliability of the CQ-B, and (c) evidence supporting its construct validity to guide interpretation of the evaluation data gathered during the randomized trial, as well as provide

potentially useful information about components of "teen attitudes about sexuality" that might inform the design of interventions. Because the CQ-F differs from the CQ-B only minimally, our analyses focused on the CQ-B only, the more comprehensive initial assessment.

<div align="center">Method</div>

Participants

Seven middle schools located in suburban areas of the Pacific Northwest participated. Within those schools, student participants were recruited from 8th grade health classes. A total of 608 8th graders completed the CQ-B in the fall term of the 2009-2010 academic year.

Psychometric Properties of Source Instruments

To assess the extent to which items included on the CQ-B were representative of the construct of interest, we carefully reviewed the source instruments from which items were derived. A table compiled by AFL provided a list of the sources of specific items. These sources were carefully reviewed for data on psychometric properties. Because only scales measuring youth perceptions are relevant to the CQ-B, psychometric properties for scales measuring parent perceptions were not included in our review. While psychometric properties are rarely specific to one specific item in isolation and therefore are not transportable from one instrument to another, the overall psychometric quality of the source instrument provides some confidence in the relevance of specific items.

Scale Reliability Analyses

We assessed the cohesiveness of the subscales of the CQ-B by calculating Cronbach's alpha as an index of a measure's internal consistency (Cronbach, 1951; Nunnally, 1967). Because one subscale contained a small number of items, we used the Spearman-Brown prophecy formula to calculate alpha (Spearman, 1910). Subscale 1 ("Activities") consisted of 1

item only; therefore, reliability calculations were not possible and it was omitted from the analyses. Subscale 5 ("About you") contained items measuring respondents' demographic information (age, gender, ethnicity, race) rather than a psychological construct; therefore it also was omitted from the analyses. Items included on the remaining subscales differed in scoring range and directionality. To achieve interpretable outcomes, we rescaled and recoded items in an effort to align variance proportions and directionality of all items within a subscale. For example, scoring ranges for the items on subscale 3 ("Your family") varied from 0-1 (2 points) to 1-3 (3 points) to 0-3 (4 points). Because the widest scoring range was a 4-point scale, all items were rescaled on a 4-point scale. For the majority of items, higher scores were more desirable than lower scores. For items 9a, 9b, 12a, and 12b, however, the opposite was true. Therefore we reversed the coding for those items to match the directionality of the overall subscale. Table 1 provides an overview of which items were rescaled and reverse-coded by subscale.

We omitted items whose means were uninterpretable, did not appear to pertain directly to the construct being measured by the subscale, or did not apply to the majority of our sample. On subscale 3 ("Your Family"), item 4 ("who do you live with now") was omitted from the analysis because its mean (average of choices of adults) was not meaningfully interpretable. Items 7 ("I have/do not have a mother/mother figure") and 10 ("I have/do not have a father/father figure") were also omitted because they represented demographic information and were less relevant to the construct of interest. On subscale 7 ("What You Think and Do"), items 41, 42, 43, and 44 measuring past sexual activity and birth control choices were omitted because they pertained only to a small portion of our sample. A total of 8.9% of respondents reported ever having had sex, and 5.3% reported engaging in sexual activity in the last 6 months. Questions regarding birth control methods were not applicable to the majority of our sample.

Not all respondents completed all items on all subscales. Because missing data affected all variables, imputing missing data for given variables seemed difficult to justify. Instead, we relied on pairwise deletion, which yielded sample sizes ranging from n = 567 (Subscale 2) to n = 478 (Subscale 6), well above the recommended minimum of 5 responses per item (Nunally & Bernstein, 1994). Subscale 7 contained 1 item (#36) that was specific to boys, and 1 item (#37) that was specific to girls. To maintain a maximal sample size for this subscale, we conducted separate reliability analyses for boys (n = 253) and girls (n = 271). Table 1 provides an overview of sample sizes by subscale.

Following Cronbach's recommendation to interpret coefficient alpha only when data meet the assumptions of (a) approximately equal inter-item correlations, (b) random sampling of respondents and items, and (c) differential decision-making (Cronbach & Shavelson, 2004), we calculated mean inter-item correlations for all subscales. Because inter-item correlations varied widely, we also calculated the standard error of measurement for each subscale mean (Salvia & Ysseldyke, 2001) to assess the amount of error contained in the mean. To evaluate our degree of confidence in the range of values within which the "true" mean lies, we constructed 95% confidence intervals around each subscale mean.

It is difficult to meet the assumption of random sampling of respondents and items in a natural setting. Strictly speaking, participating students within participating schools were not randomly selected from all possible participants, because schools were selected from a given geographical area and needed to agree to participate, and students within schools needed to agree to complete the surveys. However, within the parameters of the design of the randomized trial, the assumption of random sampling of respondents appears minimally met. Similarly, the assumption of random sampling of items might be met by the creation of the instrument based on

available literature and existing surveys. The assumption of differential decision-making is

clearly met, because no criterion score exists for the CQ-B; instead survey results are used to

obtain a baseline measure of the attitudes of students prior to the intervention. Differences

between baseline and follow-up surveys allow decisions regarding the effectiveness of the

intervention.

Construct Validity Analysis

Because little is known about the latent factors underlying the CQ-B, we conducted an

exploratory factor analysis to assess the dimensionality of the instrument as well as its overall

cohesion. Exploratory factor analysis requires that (a) the extraction method, (b) the number of

factors, and (c) the rotation method be specified (Russell, 2002). Following the recommendation

of Gorsuch (1983), we used principal axis factoring to extract the minimum number of factors

accounting for the variance shared by the items. This iterative process relies on a reduced

correlation matrix taking into account naturally occurring measurement error. We relied on the

findings of Tobin, Vincent, Xue, and Fuller (in review) to specify the expected number of

factors. Tobin et al. used a small sample of surveys (n = 72) collected during the pilot study

preceding the randomized trial conducted by NWFS for a principal component analysis with

Varimax rotation for an initial and tentative assessment of the number of components contained

in the measures. The principal component analysis identified a total of 9 factors with an

eigenvalue of >1. Six of these factors were clearly interpretable: (1) attitudes toward marriage

and abstinence, (b) past sexual experience, (c) boys' sexual desire and lack of self-regulation, (d)

plans for the future, (e) refusal skills, and (f) current sexual activity (Tobin et al., in review).

Based on these findings we specified 6 factors for our exploratory factor analysis. We used

Varimax rotation to capture the most parsimonious representation of all respondents within an

orthogonal factor structure, and reported variance explained for both unrotated and rotated

factors to evaluate shifts in explained variance from the first to later factors (Russell, 2002).

Varimax rotation is recommended if more than one factor is expected to emerge from the data,

because it identifies the factor solution maximizing explained variance across all available

factors (Gorsuch, 1983). Because we hypothesized the existence of 6 distinctive factors based on

Tobin et al. (in review), the Varimax rotation method seemed appropriate.

To achieve stable and interpretable outcomes, the recommended sample size for an

exploratory factor analysis is 5 to 10 cases per item (Gorsuch, 1983; Russell, 2002). Given that

the CQ-B has 43 items and was completed by n = 608 students, we exceeded the recommended

sample size and expected our analysis to yield interpretable outcomes.

Results

Psychometric Properties of Source Instruments

The CQ-B was constructed based on existing assessments and surveys. Items were

derived from 13 published instruments; the psychometric properties of some of these instruments

were reported in published studies. A summary of our review of these studies follows.

Youth Asset Survey (YAS). The YAS was created by a research team at the University of

Oklahoma Health Sciences Center to examine how family, peer, and neighborhood influence

adolescents' behavioral choices. A factor analysis of the YAS showed that it had good internal

reliability (Cronbach alpha values exceeded .60), and measured six theorized developmental

assets: (a) family communication, (b) peer role models, (c) future aspirations, (d) responsible

choices, (e) community involvement, and (f) non-parental role models (Oman et al., 2002).

Saving Sex for Later surveys. The youth survey used to evaluate "Saving Sex for Later"

consisted of three scales: family support (items assessing comfort level with trusted adults),

family monitoring (items assessing parents' knowledge of their child's whereabouts and

activities), and behavioral risk (items assessing petting behavior). No psychometric data for the

youth survey are reported (O'Donnell et al., 2005).

Family management variables. Based on consistent evidence of the association between

family management practices and youth delinquent behavior and extensive clinical work,

Patterson and Stouthamer-Loeber (1984) identified 4 variables to measure the impact of

interventions: (a) parental monitoring of the child's whereabouts, (b) effective discipline, (c)

effective problem-solving skills, and (d) supporting the child's acquisition of prosocial skills.

These four variables were measured via interviews, questionnaires, direct observation at home

and in the laboratory, and telephone contacts and were found to be associated with youth's police

contacts and delinquency. Test-retest correlation of the measures was r = .77 (Patterson &

Stouthamer-Loeber, 1984).

Authoritative Parenting Index (API). The API measured children's perceptions of

parenting behaviors. A factor analysis revealed that the API contained two factors: parental

responsiveness and parental demandingness. Across two studies the responsiveness scales

showed excellent internal consistency with alpha values ranging from .71 to .90, as did the

demandingness scale with alpha values ranging from .65 to .83. Construct validity of the API

was assessed by comparing responses of children whose parents were authoritative to responses

of children whose parents were lenient. Reponses were correctly classified in about 50% of

cases, which is consistent with previous research (Jackson, Henriksen, & Foshee, 1998).

National Longitudinal Study of Adolescent Health (Add-Health) questionnaires. Add-

Health was a comprehensive study administering a series of measures to students, parents, and

school administrators. The adolescent in-home interview consisted of 38 sections collection data

on constructs ranging from daily activities perceptions about pregnancy, AIDS, and STDs, relationship to parents, perceptions of neighborhood, romantic aspirations, and current behaviors. No psychometric information about the questionnaire and interview could be located on the Add-Health website (Carolina Population Center, 1999).

Teen Activities and Attitudes Questionnaire (TAAQ). The TAAQ consisted of 6 scales: (a) daily activities and schools, (b) attitudes and experiences, (c) health facts and education, (d) some things you might think or do, (e) the future, and (f) experiences. These scales measured current behaviors, perceptions of the consequences of sexual activity, knowledge of STDs, perceptions of the youth's own current behavior, future plans and future sexual behavior, and past sexual behavior. No psychometric properties of the TAAQ were reported (Trenholm et al., 2007).

Parent and teen surveys used to evaluate the Facts & Feelings. The teen survey used by Miller et al. (1993) assessed sexual values, communication, knowledge, and demographics as well as physical maturation, avoidance skills, peer sexual environment, dating, and sexual intentions and behavior. All scales except for the knowledge scale were tested for internal reliability and had alpha values ranging from .55 to .93 (Miller et al., 1993).

Structured questionnaire for parent-adolescent communication. The questionnaire used by Pistella and Bonati (1999) to assess communication between parents and daughters was administered to female youth 19 years and younger. Its quantitative portion measured youth's strategies to improve communication with parents based on items derived from the literature, while its qualitative portion measured youth's recommendations for improving communication with parents. No psychometric properties of the questionnaire were reported (Pistella & Bonati, 1999).

National Survey of Adolescents and Young Adults (NSAYA). The NSAYA was conducted with 3 questionnaires: one for parents of 13-17 year olds, one for adolescents aged 13-14 years, and one for adolescents aged 15-24 years. The adolescent questionnaires measured concern about sexual and general health threats, sources and depth of knowledge about potential consequences of sexual activity, peer pressure, perceived effectiveness of birth control methods, opinions about using birth control, sexual experience, pressure to become sexually active, and concerns about consequences of own sexual behavior. No psychometric properties of the questionnaires were reported (Hoff, Greene, & Davis, 2003).

Assessment used to evaluate the efficacy of Keepin' it R.E.A.L.! DiIorio et al. (1999) used a youth assessment that focused on self-esteem, stress, perspectives of the future, ethnic identity, abstinence self-efficacy, abstinence outcome expectations, parenting, and communication with parents. Although no psychometric properties of the assessment were reported, the measured constructs showed associations with sexual initiation that were consistent with the literature (DiIorio et al., 1999).

Instruments used to assess the impact of older boyfriends and girlfriends on 6th graders sexual initiation. Marin et al. (2000) used instruments that measured the age of the oldest boy- or girlfriend, unwanted sexual advances, peer norms, and sexual behavior. No psychometric properties of the instruments were reported; however the expected association between having an older boy- or girlfriend and greater exposure to unwanted sexual advances as well as greater sexual activity lent the instruments some construct validity (Marin, Coyle, Gomez, Carvajal, & Kirby, 2000).

Survey used to assess the impact of "For Keeps." Borawski, Trapl, Lovegreen, Colabianchi, & Block, (2005) administered a youth survey that measured knowledge of

HIV/STDs, abstinence values, self-efficacy, behavioral intentions, and behavioral outcomes. Outcomes were in the hypothesized direction and supported by extant literature; therefore it might be reasonable to assume that the survey had construct validity, even though no psychometric properties were reported Borawski et al. (2005).

Youth Risk Behavior Survey (YRBS). The YRBS was developed by the Centers for Disease Control (CDC) to monitor youth's health risk behaviors. Two versions exist; one for middle school students (grades 7-8), and one for high school students (grades 9-12). In addition to demographic information, body image, exercise, and dietary habits, the survey assesses the extent to which students take safety precautions in everyday activities such as bicycling or riding in cars, engage in violence related behaviors, have been bullied, have attempted suicide, use illegal substances, and have access to medical care. The high school version of the survey adds items about sexual behavior, intentions, and knowledge of sexually transmitted diseases. The survey has been found to have good test-retest reliability with kappa for individual items ranging from 61-100%. Responses from students in 8th grade and above had greater reliability than responses from younger students. A literature review conducted by CDC in 2003 indicated that self-report of sexual behaviors depended on cognitive and situational factors for which no commonly accepted validity standards exist. Instead, CDC recommended that researchers interpret YRBS data in the context of their knowledge of cognitive and situational factors (Brener et al., 2004).

Of the 13 published source instruments, 8 (62%) did not have documented psychometric properties, 3 (23%) had some form of documented reliability (internal consistency or test-retest reliability), and 2 (15%) had both documented reliability and validity. Of the 43 items on the CQ-B, 36 were derived from the source instruments reviewed above. Of those 36 items, 23

(64%) were derived from instruments with no published psychometric properties, 7 (19%) from

instruments with published reliability data, and 6 (17%) from instruments with published

reliability and validity data.

Scale Reliability Analyses

Table 1 provides an overview of the outcomes of our reliability analyses for the CQ-B by

subscale. Coefficients alpha ranged from .550 (subscale 2) to .846 (subscale 6). The low alpha

value for Subscale 2 might be due to the relatively small number of items (k = 5); the Spearman-

Brown prophecy formula yielded an adjusted alpha value of .710. Inter-item correlations varied

widely, from a minimum value of r = .048 to a maximum value of r = .901. Despite this

variability in inter-item correlations, 95% confidence intervals calculated with the standard error

around subscale means were relatively narrow, ranging from .06 (subscale 2) to .23 (subscale 7-

boys) measurement units. Figure 1 illustrates the width of 95% confidence intervals around each

subscale mean.

Construct Validity Analysis

Compared to the principal components Tobin et al. (in review) identified, our analyses

yielded somewhat comparable outcomes. Items 13a-o measuring respondents' frequency of

talking with their parent(s) about teen sexual activity clearly emerged as one cohesive factor.

Item 15 (frequency of dating a person at least 3 years older) aligned with items 19a-c (use of

tobacco, alcohol, and marijuana), items 30-31 (likelihood of engaging in sexual activity within

the next year), item 35 (experience with tongue kissing), and items 38-40 (experience with

touching others or having been touched in a sexual manner and having sex). Taken together,

these items seemed to measure risky behaviors. Item 5d (family rules for watching movies and

TV), item 18 (experience with avoiding people who might spell trouble), item 24 (importance of

delaying sex until marriage), items 25a-c (consequences of teen sex), items 26a-b (appropriateness of teen sexual activity), item 32 (intention to abstain until marriage and item 34 (experience with lip kissing) converged into one factor that appeared to measure the extent to which respondents adhered to adult rules and expectations. Items 2a-b (presence of positive adults), item 8 (closeness to mother), items 9a-b (ease of talking with mother), item 11 (closeness to father), items 12a-b (ease of talking with father), and item 14 (frequency of communication with parents about teen sex) aligned into one factor that clearly measured teens' access to parental or adult support. Item 17 (ability to resist wrong activities) aligned with items 29a-f (ability to resist sexual activity and advances) and appear to represent a factor measuring refusal skills. Finally, items 5a-c and 5e (family rules concerning dating and hanging out), and items 6a-d (parents' efforts to monitor dating activities and free time) aligned into one factor that appeared to measure awareness of adult rules and monitoring.

The following items seemed unaligned with any of the six factors: item 28 (sexual activity of closest friends), item 33 (likelihood of using birth control), item 31-c (clarity of future goals), item 1 (after school activities), item 16 (exposure to peer pressure), item 26c (only abstinence avoids pregnancy and diseases) and item 27b (good marriage seems realistic). With the exception of items 1 and 16, these items appear to measure concepts that either lie far in the future or with which teens might simply be unfamiliar, e.g. birth control, the value of abstinence, and transmission of disease.

Similar to Tobin et al., our analysis yielded a factor focused on refusal skills and a factor focused on past sexual experience. The remaining factors we identified appeared to focus on communication with parents, parental and adult support, and awareness of and adherence to adult rules and expectations.

Table 2 provides an overview of the variance explained by the identified factors. Taken together, the 6 identified factors explained 38.91 % of the variance in the data. The rotated model indicated that factor 1 (communication with parents) explained approximately a third of the total explained variance, while factors 5 (refusal skills) and 6 (awareness of adult rules and monitoring) each explained approximately a tenth of the total explained variance. Because we used an orthogonal factor structure, the explained variance reported in Table 2 does not overlap.

<div align="center">Discussion</div>

Interpretation of Results

Overall, the results of our examination of the psychometric properties of the source instruments, reliability analyses, and construct validity analysis suggested that the CQ-B represents a reliable and valid instrument. Our close review of the measures on which the CQ-B was based indicated that of the total 43 items contained in the CQ-B, 23 (54%) were derived from measures with undocumented psychometric properties, 7 (16%) were derived from measures with some form of known reliability, and 6 (14%) were derived from measures with known reliability and validity. It is important to consider that adopting items from reliable and validated source instruments does not automatically guarantee that those items function in an equally reliable and valid way in the newly created instrument. Additional psychometric testing of the CQ-B might be needed.

Given that the CQ-B is a self-report measure, it is important to view item relevance through the lenses of respondent and measure administration characteristics. Brener and colleagues (2003) found that cognitive and situational factors impact the validity of self reports. This suggests that item relevance might not be exclusively a property of an instrument, but rather a property of the use of the instrument (Thompson et al., 2005). Our additional analyses focused

precisely on the use of the CQ-B within a given sample and context.

The outcomes of our reliability analyses suggested that the CQ-B had very good internal consistency within our sample and context. Completed by 8th graders during their health classes in middle schools located in the Pacific Northwest, all CQ-B subscales had alpha values within the range recommended in the literature. For research purposes, Nunnally (1967, 1978) identified alpha values between .70 and .80 as adequate, and values of .80 and higher as very good. For clinical testing and consequential placement decisions, Rosenthal & Rosnow (1991) recommended alpha values of .85 or higher, while Nunnally and Bernstein (1994) recommended alpha values of .95 as a "desirable standard" (p. 265). With alpha values ranging from .710 (subscale 2) to .846 (subscale 6), it appears that the subscales of the CQ-B functioned as cohesive measures assessing well-defined constructs.

The width of the 95% confidence intervals constructed around subscale means indicated that the amount of error contained within each subscale varied but appeared overall small. Subscale 2 ("adult support and goal directedness"), subscale 3 ("ease of communication about sexual activities"), and subscale 4 ("ability to resist peer pressure") contained the smallest amount of error, while subscale 6 ("consequences of teen sexual activity") and subscale 7 ("past sexual behavior and current intentions") for both boys and girls contained the greatest amount of error. Subscale 1 ("current activities") also contained a fair amount of error; however, given that the subscale consisted of one item only, calculations of its stability might be compromised. These outcomes suggest that items related to perceptions of social environments and how to fit into them might have been more easily interpretable to our respondents than items related to sexual activity. Eighth graders might be keen observers of the dynamics of their immediate social environments; they might be less familiar with their own or their peers' sexual activity and

intentions and their consequences.

Strong cohesion of subscales, however, does not automatically attest to the subscales' ability to measure clearly defined constructs. Our exploratory factor analysis provided evidence for the CQ-B's ability to measure clearly defined constructs. The factors that emerged suggested communication with parents, access to adult support, and adherence to adult rules and expectations as clearly defined aspect of teens' sexual maturation. Similar to the outcomes of our reliability analyses, the most clearly defined factors seemed to focus on teens' efforts to fit into their immediate social environments defined by adult rules and expectations through communication with parents and access to adult support. Items relating to less familiar territory, e.g. the value of abstinence, transmission of disease, use of birth control, or likelihood of future marriage seemed less clearly aligned with identifiable factors.

Taken together, this evidence suggests that the CQ-B consists of relevant items and has internal consistency and construct validity to yield interpretable data. Data collected with the CQ-B might provide important insights into the effectiveness of pregnancy prevention programs administered in middle school settings.

Limitations

The results of our analyses need to be interpreted in the context of the following limitations. Several respondents omitted various items. It seems reasonable to assume that some items—perhaps especially items measuring past sexual activity or current and future sexual intentions—might have been purposefully omitted because the respondent perceived the item as too intrusive or inappropriate. While the omission of an item constitutes missing data whose frequency can be counted, the reasons why missing data occurred are unknown. The loss of data alone did not substantively impact our analyses; we always exceeded recommended minimum

sample sizes. Potential reasons for missing data, however, are apt to impact interpretability of our findings, because they might have subjected our sample to unknown selection bias.

Because of the great variability in scaling and directionality of item scoring, a fair amount of rescaling and reverse coding was necessary to prepare the data for analysis. While great care was taken in the completion of these preparatory procedures, our rescaling and reverse-coding might have introduced slight changes in the variance available for analysis.

Importance and Implications of Findings

Knowledge of the reliability and dimensionality of a measure is critical to advance the theory of the construct being measured (Gorsuch, 1983; Nunnally & Bernstein, 1994). Our findings might be useful for further development of the theory of teenage sexual initiation, especially when it tends to occur, what individual and contextual factors might delay it, and what impacts teens' knowledge of its consequences. Since our findings suggested that 8th graders might be less familiar with sexual activities and future plans for marriage, it might be useful to investigate the value of multiple versions of the measure geared towards specific age groups (e.g. middle school students, high school students), similar to the age-specific versions of the YRBS . It might be that younger teenagers are much more influenced by their immediate social environment, while older teenagers at the cusp of adulthood are better able to assess their immediate and future sexual intentions and marriage plans. More research in the area of measuring adolescent sexual intentions and behaviors is needed.

References

Allen, J.P., Philliber, S., Herrling, S. & Kupermine, G.P. (1997). Preventing teen pregnancy and academic failure: Experimental evaluation of a developmentally based approach. *Child Development, 64*, 729-742.

American Educational Research Association (1999). *Standards for Educational and Psychological Testing.* Washington, DC: Author.

Borawski, E.A., Trapl, E.S., Lovegreen, L.D., Colabianchi, N., & Block, T. (2005). Effectiveness of abstinence-only intervention in middle-school teens. *American Journal of Health Behavior, 29,* 423-434.

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. *Journal of Adolescent Health, 33*, 436-457.

Brener, N.D., Kann, L., Kinchen, S.A., Grunbaum, J.A., Whalen, L., Eaton, D. et al. (2004). Methodology of the youth risk behavior surveillance system. *Morbidity and Mortality Weekly Report 53* (RR12), 1-13. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5312a1.htm

Card, J.J., Lessard, L., & Benner, T. (2007). PASHA: Facilitating the replication and use of effective adolescent pregnancy and STI/HIV prevention programs. *Journal of Adolescent Health, 40*, 275.e1-275.314.

Carolina Population Center (1999). *National longitudinal study of adolescent health: Parental questionnaire code book*. Chapel Hill, NC: University of North Carolina.

Carter-Jessop, L., Franklin, L.N., Heath, J.W., Jimenez-Irizarry, G. & Peace, M.D. (2000). Abstinence education for urban youth. *Journal of Community Health, 25*, 293-304.

Centers for Disease Control and Prevention. (2006). Youth Risk Behavior Surveillance-United States, 2005. *Morbidity and Mortality Weekly Report;* 55 (SS-5):1-108.

Coffman, J.K., Guerin, D.W., & Gottfried, A.W. (2006). Reliability and validity of the parent-child relationship inventory (PCRI): Evidence from a longitudinal cross-informant

investigation. *Psychological Assessment, 18*, 209-214.

Coley, R.L. Votruba-Drzal, E., & Schindler, H.S. (2009). Fathers' and mothers' parenting

predicting and responding to adolescent sexual risk behaviors. *Child Development, 80*,

808-827.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*,

297-334.

Cronbach, L.J., & Shavelson, R.J. (2004). My current thoughts on coefficient alpha and

successor procedures. *Educational and Psychological Measurement, 64*, 391-418.

Denny, G., Young, M., Rausch, S., & Spear, C. (2002). An evaluation of an abstinence

education curriculum series: *Sex Can Wait. American Journal of Health Behavior, 26*,

366-377.

DiIorio, C., Resnicow, K., Manteuffel, B., Dudley, W., Denzmore, P., Rodgers Tillman, G., et

al., (1999). *Keepin' it R.E.A.L.!: Results of baseline assessments for participants enrolled*

in a mother-adolescent HIV prevention program. Paper presented at the National HIV

Prevention Conference. Atlanta, GA.

DuRant, R.H., Wolfson, M., LaFrance, B., Balkrishnan, R., Pharm, M.S., & Altman, D. (2008).

An evaluation of a mass media campaign to encourage parents of adolescents to talk to

their children about sex. *Journal of Adolescent Health, 38*, 298e1-298c9

Genuis, S.J., & Genuis, S.K. (2004). Managing the sexually transmitted disease pandemic: A

time for reevaluation. *American Journal of Obstetrics and Gynecology, 191*, 1103-1112.

Gorsuch, R.L. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Greenberg, M.T., Weisberg, R.P., O'Brien, M.U., Zins, J.E., Fredericks, L., Resnik, H., et al.

(2003). Enhancing school-based prevention and youth development through coordinated

social, emotional, and academic learning. *American Psychologist, 58*, 466-474.

Jackson, C., Henriksen, L., & Foshee, V.A. (1998). The Authoritative Parenting Index:

Predicting health risk behaviors among children and adolescents. *Health Education and Behavior, 25,* 319-337.

Hoff, T., Greene, L., & David, J. (2003). *National survey of adolescents and young adults: Sexual health knowledge, attitudes and experiences*. Menlo Park, CA: Henry J. Kaiser Family Foundation.

Kirby, D., & Laris, B.A. (2009). Effective curriculum-based sex and STD/HIV education programs for adolescents. *Child Development Perspectives, 3*, 21-29.

Lederman, R.P., Chann, W., & Roberts-Gray, C. (2008). Parent-adolescent relationship education (PARE): Program delivery to reduce risks for adolescent pregnancy and STDs. *Behavioral Medicine, 33*, 137-143.

Marín, B.V., Coyle, K., Gomez, C., Carvajal, S., & Kirby, D. (2000). Older boyfriends and girlfriends increase risk of sexual initiation in young adolescents. *Journal of Adolescent Health, 27*, 409-418.

Miller, B.C., Norton, M.C., Jenson, G.O., Lee, T.R., Christopherson, C., & King, P.K. (1993). Impact evaluation of Facts & Feelings: A home-based video sex education curriculum. *Family Relations, 42*, 392-400.

Moore, K.A., & Sugland, B.W. (1997). Using behavioral theories to design abstinence programs. *Children and Youth Services Review,19*, 485-500

Nunnally, J. C. (1967) *Psychometric theory*. New York: McGraw-Hill.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Donnell, L., Stueve, A., Agronick, G., Wilson-Simmons, R., Duran, R., & Jeanbaptiste, V. (2005). Saving sex for later: An evaluation of a parent education intervention. *Perspectives on Sexual and Reproductive Health, 37*, 166-173.

Oman, R.F., Vesely, S.K., McLeroy, K.R., Harris-Wyatt, V., Aspy, C.B., Rodine, S., & Marshall, L. (2002). Reliability and validity of the youth asset survey (YAS). *Journal of Adolescent Health, 31*, 247-255.

Patterson, G. R., & Stouthamer-Loeber, M. (1984). The correlation of family management practices and delinquency. *Child Development, 55*, 1299-1307.

Pistella, C.L., & Bonati, F.A. (1999). Adolescent women's recommendations for enhanced parent-adolescent communication about sexual behavior. *Child and Adolescent Social Work Journal, 16*, 305-315.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Russell, D.W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin. Personality and Social Psychology Bulletin, 28*, 1629-1646.

Salvia, J., & Ysseldyke, J. (2001). *Assessment.* Boston: Houghton Mifflin.

Sipe, C. (2002). Mentoring programs for adolescents. *Journal of Adolescent Health, 31*, 251-260.

Spearman, C. (1910). Correlation calculated with faulty data. *British Journal of Psychology, 3*, 271-295.

Thompson, B., Diamond, K., McWilliam, R., Snyder, P., & Snyder, S. (2005). Evaluating the

quality of evidence from correlational research for evidence-based practice. *Exceptional*

Children, 71, 181-194.

Tobin, T.J., Vincent, C.G., Xue, L., & Fuller, R. (in review). *Pre-coital sexual behaviors and*

intentions of young adolescents: Results of a pilot study of a health class that included a

positive youth development project.

Trenholm, C., Devaney, B., Fortson, K., Quay, L., Wheeler, J., & Clark, M. (2007). *Impacts of*

four Title V, Section 510 abstinence education programs. Princeton, NJ: Mathematical

Policy Research, Inc.

Unruh, D., Bullis, M., & Yovanoff, P. (2004). Adolescent fathers who are incarcerated juvenile

offenders: Explanatory study of the co-occurrence of two problem behaviors. *Journal of*

Child and Family Studies, 13, 405-419

Weed, S.E., Ericksen, I.H., Lewis, A., Grant, G.E., & Wibberly, K.H. (2008). An abstinence

program's impact on cognitive mediators and sexual initiation. *American Journal of*

Health Behavior, 32, 60-73.

White, F.A. (1996). Parent-adolescent communication and adolescent decision-making. *Journal*

of Family Studies, 2, 41-56.

Table 1
Core Baseline Questionnaire Subscales, items rescaled and recoded, sample sizes, reliability coefficients, and inter-item correlations

Scale (number of items) and tentative description of construct being measured	Rescaled and reverse-coded items		Number of Respondents completing all items	Cronbach's coefficient α	Inter-item correlation		
	rescaled	reverse coded			Mean	Minimum	Maximum
1. "Activities" (1) – current afterschool activities			Omitted from the analyses				
2. "What You Think" (5) – adult support and goal directedness	3a, 3b, 3c	3a, 3c	567	.550 (.710)*	.209	.063	.481
3. "Your Family" (34) – ease of communication about sexual activities of self and others	5a-5e, 6a-6d, 9a-b, 12a-b	9a, 9b 12a, 12b	480	.828	.141	-.321	.781
4. "What Young People Think and Do" (7) – Exposure to and ability to resist peer pressure	19a-c	16	517	.751	.309	.070	.573
5. "About You" (4) – demographic information			Omitted from the analyses				
6. "More Questions" (16) – Consequences of teen sexual activity	24, 25a-c, 26a-c, 27a-b, 28	26a, 27a, 29a, 29b, 29e	478	.846	.253	.048	.753
7. Girls: "What You Think and Do" (15) – Past sexual behavior and current intentions	34, 35, 37, 38, 39, 40	31, 33	271	.838	.351	-.216	.901
7. Boys: "What You Think and Do" (15) – Past sexual behavior and current intentions	34, 35, 36, 38, 39, 40		253	.810	.304	-.228	.694

* Low α might be due largely to small number of items (k = 5). Applying the Spearman-Brown Prophecy Formula demonstrates that doubling the scale length to k=10 would produce α in parentheses.

Table 2: Variance explained by unrotated and rotated factors

Factor	Principal Axis Factors	
	Unrotated	Rotated
1	15.65	10.66
2	10.82	7.46
3	4.48	7.39
4	3.36	5.73
5	2.44	3.92
6	2.16	3.75
Total	38.91	38.91

Figure 1: 95% confidence intervals around subscale means