Published in final edited form as:

Ethn Health. 2010 April ; 15(2): 181–197. doi:10.1080/13557851003615552.

Youth health risk behavior assessment in Fiji: The reliability of Global School-based Health Survey content adapted for ethnic Fijian girls

Anne E. Becker, M.D., Ph.D., Sc.M,

Department of Global Health and Social Medicine, Harvard Medical School, Boston, United States

Andrea L. Roberts, Ph.D.,

Department of Society, Human Development, & Health, Harvard School of Public Health, Boston, United States

Alexandra Perloe, B.A.,

Department of Psychiatry, Massachusetts General Hospital, Boston, United States

Asenaca Bainivualiku, B.A.,

School of Child & Youth Care, University of Victoria, Victoria, BC, Canada

Lauren K. Richards, B.A.,

Department of Psychology, Center for Anxiety and Related Disorders, Boston University, Boston, United States

Stephen E. Gilman, Sc.D., and

Department of Epidemiology and Department of Society, Human Development, & Health, Harvard School of Public Health, Boston, United States

Ruth H. Striegel-Moore, Ph.D.

Department of Psychology, Wesleyan University, Middletown, CT, United States

Abstract

Objective—The Global School-based Student Health Survey (GSHS) is an assessment for adolescent health risk behaviors and exposures, supported by the World Health Organization. Although already widely implemented—and intended for youth assessment across diverse ethnic and national contexts—no reliability data have yet been reported for GSHS-based assessment in any ethnicity or country-specific population. This study reports test-retest reliability for GSHS content adapted for a female adolescent ethnic Fijian study sample in Fiji.

Design—We adapted and translated GSHS content to assess health risk behaviors as part of a larger study investigating the impact of social transition on ethnic Fijian secondary schoolgirls in Fiji. In order to evaluate the performance of this measure for our ethnic Fijian study sample (n=523), we examined its test-retest reliability with kappa coefficients, % agreement, and prevalence estimates in a sub-sample (n=81). Reliability among strata defined by topic, age, and language was also examined.

Results—Average agreement between test and retest was 77%, and average Cohen's kappa was 0.47. Mean kappas for questions from core modules about alcohol use, tobacco use, and sexual behavior were substantial, and higher than those for modules relating to other risk behaviors.

Conclusions—Although test-retest reliability of responses within this country-specific version of GSHS content was substantial in several topical domains for this ethnic Fijian sample, only fair reliability for the module assessing dietary behaviors and other individual items suggests that population-specific psychometric evaluation is essential to interpreting language and country-specific GSHS data.

Keywords

Global School-based Student Health Survey; GSHS; reliability; ethnic Fijians; Fiji

Introduction

Health-risk behaviors are major contributors to adolescent mortality as well as to the global burden of disease (Murray and Lopez 1997, World Health Organization 2003, Blum & Nelson-Mmari 2005). Reliable and valid assessment of the prevalence and severity of these health-risk behaviors in adolescents is critical to developing health policy and interventions that reduce their impact on population health. The World Health Organization (WHO) supported the development of the Global School-based Student Health Survey (GSHS) for the systematic evaluation of health-risk behaviors in adolescents across diverse populations. Information obtained from the GSHS is intended to facilitate both in-country identification of health care priorities for adolescents and between-country comparison of prevalence and efficacy of interventions for health-risk behaviors.

The Global School-based Student Health Survey

The GSHS was developed to assess key behavioral risk and protective factors. A self-report questionnaire, its items are organized in 10 core modules that address demographics, alcohol and other substance use, tobacco use, sexual behaviors contributing to unintended pregnancy and/or sexually transmitted infections (STIs), violence and unintentional injury, mental health, dietary behaviors, physical activity, hygiene, and protective factors (WHO 2009). The survey draws content from the CDC's Youth Risk Behavior Survey (YRBS), which utilizes a self-report questionnaire to evaluate trends in youth behavioral risk in the United States (National Center for Chronic Disease Prevention and Health Promotion 2008a). In order to promote flexibility for adaptation across diverse global populations, GSHS core module content can be supplemented by selection from additional related questions; moreover, some of these questions include prompts suggesting adaptation with country-specific examples, options, or phrasing. Suggested implementation utilizes standard methodology across participating sites to generate population-specific data that can be used regionally and globally in identifying and responding to health priorities and in evaluating youth health promotion (WHO 2009).

At least 89 countries have initiated or completed implementation of the GSHS. Of these, there are 43 fact sheets, 20 datasets, and 13 full reports, along with questionnaires translated into 12 languages in addition to English, posted on the WHO website (WHO 2009). We identified eight additional studies based on data generated by the GSHS in country-specific settings published in the scientific literature (Singh *et al.* 2006, Granero *et al.* 2007, Rudatsikira *et al.* 2007a, Rudatsikira *et al.* 2007b, Rudatsikira *et al.* 2007c, Muula *et al.* 2007, Turagabeci *et al.* 2008, Due and Holstein 2008).

Reliability of youth risk behavioral assessment with the YRBS and GSHS

We could not identify published reliability data for any country-specific version of the GSHS. Although many local contextual factors potentially influence reliability across diverse settings, many GSHS questions are identical or similar to items in the YRBS for which reliability data are available on U.S. adolescents. The reliability of the YRBS has been formally evaluated in two large studies that examined the temporal stability of responses to questions about risk behaviors (i.e., test-retest reliability) (Brener *et al.* 1995, Brener *et al.* 2002). The test-retest reliability of the related Middle School Risk Behavior Survey (MSYRBS) has also been assessed (Zullig *et al.* 2006). Both Cohen's kappa coefficients (indicating chance-corrected agreement) and prevalence estimates at Time 1 and Time 2 (14 days apart) were calculated as a measure of reliability. Prevalence estimates at Time 1 and Time 2 did not differ substantially for either the 1992 YRBS or 2005 MSYRBS, and the majority of the 72 items evaluated in the 1999 YRBS had a Cohen's kappa coefficient of at least 41% (characterized as 'moderate' by the authors citing cut points suggested by Landis and Koch (1977)). However in this same study, 10 of the 72 items evaluated in the 1999 YRBS had kappa coefficients below a threshold of 61% (characterized as 'substantial reliability') as well as significant differences in prevalence estimates for Time 1 vs. Time 2. The authors also reported substantial variability in reliability across health-risk behaviors, with responses to items assessing tobacco use, alcohol use, and sexual behaviors exhibiting a higher degree of reliability than responses relating to dietary behaviors and physical activity (Brener *et al.* 2002).

Notwithstanding the similarity of these assessments, reliability of the GSHS in country-specific settings cannot be extrapolated from YRBS data with confidence. Although a key advantage of the GSHS is its flexibility for adaptation to country-specific settings to enhance its cultural and local public health relevance, its reliability for assessment across diverse ethnic groups—an essential component of its validity—remains unestablished.

The primary aim of the present study was to examine the test-retest reliability of a version of GSHS item content adapted for an ethnic Fijian study population in order to provide guidance for interpretation of prevalence estimates for risk behaviors assessed in this study sample. As such, this study reports the first evaluation of reliability of a country and ethnicity specific version of GSHS item content. Because some potential challenges to reliability in this study population are very likely shared with other ethnic groups across global settings, a secondary aim was to present data that could suggest cautions for interpretation of GSHS data across other ethnicity-specific, country-specific, and translated versions.

Methods

Study site and study population

Fiji is a nation located in the Western Pacific. Ethnic Fijians, a small-scale, indigenous population with Melanesian and Polynesian cultural roots, comprise approximately one half of Fiji's ethnically diverse general population. English is the formal language of instruction in Fiji, and is also widely used in commercial and communications sectors. However, most ethnic Fijians speak Standard Fijian, or one of its many communalects, as their primary language. The present study is a component of a larger cross-sectional epidemiologic and narrative data based study examining the impact of rapid social and economic transition on ethnic Fijian adolescent girls. Previous pilot data in Fiji had supported the emergence of disordered eating among ethnic Fijian adolescent girls in the setting of rapid social change there (Becker 1995, Becker *et al.* 2002, Becker 2004). This follow-up main (overarching) study sought to identify social determinants of the increased prevalence of disordered eating

among ethnic Fijians, given their previous resilience. Although disordered eating was the primary health outcome for this overarching study, additional health-risk behaviors and exposures—assessed by modules drawn from the GSHS-- were later included as secondary outcomes in response to focus group discussion data from local community leaders and parents. Some self-report data and household survey data on tobacco and alcohol use are available for adolescent Fijians (National Center for Chronic Disease Prevention and Health Promotion 2008b), but health risk data based on either the GSHS (WHO 2009) or YRBS are unavailable.

Sample selection

Eligibility criteria for the main study restricted entry to age and gender strata with the highest risk for eating disorders (that is, post-menarcheal adolescent girls) to facilitate culture-specific hypothesis testing. Therefore, inclusion criteria limited study participation to secondary school age ethnic Fijian adolescent girls (Forms 3-6 in Fiji, ages 15-20). Sampling for the main study was executed with the assistance of the Fiji Ministry of Education. We selected an administrative region characterized by relative cultural and linguistic homogeneity, but also having heterogeneity in communications and transportation infrastructure and urban proximity. All ethnic Fijian schoolgirls meeting eligibility requirements for age, grade, and school attendance within this administrative region were invited to participate in the main study. All 12 schools identified by the Ministry of Education in this region participated and 523 eligible study participants enrolled and completed the self-report survey (71% response rate) in 2007.

Study participants for the present study comprised a purposive sub-sample ($n=81$) from the main study sample. Three of the 12 participating schools were selected in order to represent broad geographic and linguistic diversity within the region. All study participants attending one of these three schools who had completed a self-report assessment at their school's first scheduled assessment visit (designated as *Time 1* for the present study) were recruited for participation in this retest reliability study. An additional eligibility criterion for the retest study participants was presence in school at the time of their school's retest survey visit, approximately one week later (referred to as *Time 2* for the present study's participants). The response rate was 100% for the retest study participation.

Adaptation and translation of GSHS item content for a study on the impact of social transition on ethnic Fijian adolescent girls

A 71-item version of the GSHS was developed by the investigators for the primary goal of measuring key outcomes for the main study. This version was adapted from both core and optional items available on the WHO GSHS website (WHO 2009) relating to modules addressing: tobacco use; alcohol use; dietary behaviors; physical activity; sexual behaviors that contribute to HIV, other STIs, and unintended pregnancy; mental health; violence and unintentional injury; and protective factors. Items that may be modified to allow country-specific flexibility according to the GSHS instructions were adapted for relevance to the local cultural context of ethnic Fijians based on the study team's ethnographic expertise. Selection from among these GSHS items was guided by goals of the main study as well as feedback from the Fiji National Research Ethics Review Committee (FN-RERC) to eliminate unnecessary redundancy with our other assessments and to avoid items probing illegal substance use. As a result, we expanded the GSHS dietary behaviours module substantially to address the primary health outcome, disordered eating. Two modules ("demographic" and "hygiene"), in addition to two items from the "alcohol and other drug use" core module, were not included in this version.

The selected GSHS items were translated from English into the local vernacular Fijian language by a bilingual native speaker, pilot-tested with 21 adolescent and adult ethnic Fijian respondents, modified and edited for clarity, and then back-translated into English by a bilingual scholar of Fijian languages. The back-translated and original versions were compared and the Fijian version was then edited to achieve consistency across the two versions. The penultimate Fijian language translation was reviewed and edited by a native speaker for grammatical and idiomatic accuracy and resulted in the final version.

Study procedures and administration of the GSHS questions

The self-report survey was administered and proctored by study staff in classrooms at the respondents' respective schools and completed within a single day by each respondent at both Time 1 and Time 2. The Time 2 assessment was completed approximately one week after the school's scheduled Time 1 visit. Study participants selected their preferred language (either English or the local Fijian language) at Time 1 and responded to this (same) preferred language version at Time 2. Study participants received a unique study identification code, which was used for labeling study related documents and preserved capacity to link their responses at Time 1 and Time 2.

Study procedures for assessment with items adapted from GSHS content were developed to facilitate the aims of the main study as well as to follow local (FN-RERC) guidance in responding to at-risk students. Specifically, the portion of the assessment adapted from GSHS content was included in a spiral bound packet alongside additional assessments for psychological and social risk factors relating to study hypotheses. After orienting the study participants to the questionnaire and procedures, study staff responded to queries about the meaning of terms throughout the assessment (e.g., words such as, "calories" and "laxatives"). Participants returned assessments to study staff and waited while they were checked for completeness if time permitted; when missing or duplicate responses were identified, study staff invited participants to complete or clarify their intended response prior to their departure. After we ascertained that a response option to an item about suicide-attempt related injury was frequently misinterpreted, study staff sought to clarify the intended response when relevant and possible as well.

As part of the related, overarching study protocol, we collected additional self-report and anthropomorphic (measured height and weight) data from all study participants as well as interview data from an independently selected sub-sample at follow-up school site visits. These anthropomorphic and interview data overlap with topic content of the GSHS, but are not presented in this paper. Additional relevant details of data collection procedures are reported elsewhere (Becker et al., 2009) and are also available upon request from the corresponding author. Written parental (or guardian) informed consent and youth assent was obtained for each study participant. Parents and guardians were informed about the study in a letter distributed with support from the school in advance of the study; youth assent was obtained in person. Consent documents were available in both English and in the local vernacular language. This study was part of a protocol carried out in accordance with universal ethical principles (Emanuel *et al* 2000) and approved by the Partners Healthcare Human Subjects Committee, the Harvard Medical School Committee on Human Studies, and the Fiji National Research Ethical Review Committee.

Data analysis

Transformation of responses to GSHS items—We examined the 43 country fact sheets and 8 (of 13) country-specific reports posted on the WHO website (WHO 2009) to determine the most frequent way in which GSHS data are presented. Because data extracted from the multiple-choice GSHS items are presented exclusively in dichotomous categories

—and moreover, dichotomized with the same cut points across almost all country reports that we examined – we transformed the majority of the items from our version into dichotomous variables with cut points that were either the same or, in some cases, that had particular relevance to this study population to evaluate reliability.

Measuring reliability—We measured test-retest reliability with both percent agreement, which is the percent of respondents whose answers were the same on both questionnaires, and Cohen's kappa coefficient, which adjusts for agreement expected by chance (Cohen 1960). We used simple kappas for dichotomous variables and weighted kappas for ordinal variables (Landis and Koch 1977). We examined 69 of 71 items, but discarded two items requesting height and weight data since the majority of participants did not self-report a height or weight and were also weighed and measured at Time 1. To explore possible differences in reliability according to participant demographic factors, mean kappa coefficients across items were calculated by respondent's age group, topic, and language of questionnaire. Additionally, mean kappa coefficients were calculated for the subset of items from our version corresponding to GSHS core module topics.

Results

The sub-sample of girls administered the retest was not significantly different from the main study sample with respect to key demographic characteristics and risk behaviors that we examined (Table 1). The majority of respondents in both the overall sample and sub-sample were 15 to 16 years of age and also selected the local vernacular Fijian language version of the GSHS. With the exception of frequently missing data for self-reported height and weight as noted above, and for an item accessing age of coitarche, the rates of missing responses were low (see Table 2).

Test-retest reliability varied considerably across items. Percent agreement ranged from 33% to 98%, with a mean of 77%. Cohen's kappa coefficients (i.e., chance-corrected agreement) ranged from 0.11 to 0.92, with a mean of 0.47. Using Landis and Koch's (1977) interpretation of kappa coefficients, 24% of items (15 of 63) had “substantial” agreement, with kappa values greater than 0.60, 43% (27 of 63) had “moderate” agreement, with kappa coefficients between 0.40 and 0.60, and 27% (17 of 63) had “fair” agreement, with kappa coefficients between 0.20 and 0.40. The remaining 6% (4 of 63) had “poor” agreement, with kappa coefficients less than 0.20 (Table 2). Six items measured rarely occurring risk factors or behaviors, for which it was not feasible to assess reliability (Byrt *et al.* 1993, Kraemer *et al.* 2002). For these items, we omitted the kappa coefficient and percent agreement from Table 2, but reported prevalence at the two time points. Three additional items were not included in mean kappa coefficients for language subgroups, because prevalence was too sparse in one or both of the groups.

For descriptive purposes, we grouped items according to content area, and calculated the average values of kappa coefficients for items within each domain (Table 3). The reliability of the questions about alcohol use (4 items) was very high, with all items having agreement at or above 90% and kappa coefficients above 0.65. Similarly, the reliability of the questions about sexual behavior (assessed in 4 of 5 items in that module) was also high, with the exception of the question about condom use during last intercourse. In contrast, the reliability of the questions about dietary behaviors (assessed in 21 of 23 items) was somewhat lower, with only a single item having a kappa coefficient greater than 0.60. Mean kappa coefficients for both core module and core plus expandable questions for dietary behaviors were only fair at 0.33 and 0.38, respectively. Among the tobacco-related questions, items about girls' personal tobacco use performed better than items about their exposure to tobacco in the media or their thoughts about tobacco harm and social

acceptance. Mean kappa coefficients for core module questions for alcohol use, sexual behaviors, and tobacco use were all substantial at 0.72, 0.69, and 0.62, respectively. Stratifying the sample by demographic factors, the average values of kappa coefficients for English speakers ($n=21$, $\text{kappa}=0.53$) and for speakers of the local Fijian vernacular ($n=60$, $\text{kappa}=0.41$) were both in the ‘moderate’ category, as were the average values of kappa coefficients for girls age 16 or older ($n=51$, $\text{kappa}=0.49$) and girls under age 16 ($n=30$, $\text{kappa}=0.42$).

Finally, prevalence at Time 1 was within 10 percentage points of prevalence at Time 2 for 52 of 64 (81%) items measuring prevalence. Items exceeding 10% difference in prevalence between Time 1 and Time 2 were most commonly items relating to dietary behaviors (4 items) and tobacco attitudes and behaviors (also 4 items, but 3 of these did not refer to smoking behavior). Other items with a prevalence discrepancy greater than 10% included one probing for an episode of depression and one for having been the target of a physical attack, each over the past year.

Discussion

The Global School-based Health Survey program provides key materials and support for the assessment of youth health risk behaviors across diverse global settings. Utility of these data, however, are limited by their unknown reliability and validity in country-specific settings or in assessing particular ethnic groups within these settings. In contrast to several studies on the reliability of the U.S. analog to the GSHS—the CDC's YRBS—we could identify no published studies evaluating the reliability of the GSHS. Therefore, this study presents the first evaluation of reliability of a population-specific version of GSHS content.

These study data support the acceptable reliability of the majority of questions adapted from GSHS content for a female adolescent ethnic Fijian study population. The percentage of items with unacceptable reliability was relatively small. Notably, each of these 4 items related to dietary behaviors, the topic that also had the lowest mean kappa coefficient and was the only one rated as having only “fair” agreement.

Despite the reliable performance of many items assessing key health risk behaviors in this study sample, items comprising this Fijian version of GSHS content had generally poorer reliability than has been reported with the YRBS in a U.S. adolescent population. In particular, the mean kappa coefficient of 0.47 (as well as the mean kappa coefficient of 0.52 for core items only) in this study and the percentage of items with substantial reliability (24%) were much lower than reported in the most recent study of reliability of the YRBS in a U.S. population (0.61 and 47.2%, respectively) (Brener *et al.* 2002). Similarities between findings from that study and the present one include higher reliability of items relating to tobacco use, alcohol use, and sexual risk behaviors compared with behaviors relating to diet and physical activity.

There are several limitations to interpretation of our study results. First, reliability data provide only one necessary condition for validity of the GSHS items in this study population. Threats to the validity of self-reported health risk behavior by adolescents include cognitive and situational factors such as social desirability, concern about stigma, and perceived confidentiality (Brener *et al.* 2003a). Collection of self-report data in culturally and linguistically diverse populations may be especially vulnerable to challenges to validity relating to facility with the survey language as well as with specific terms and concepts presented (such as bullying or purging). Variation in social norms with respect to expectations for privacy, candid disclosure, and concerns about reprisal may impact validity of self-report across diverse populations differentially. We could not identify any study

evaluating validity of GSHS items against a gold standard. Likewise, there are no studies evaluating the validity of YRBS items—with the exception of self-reported height and weight (Brener *et al.* 2003b)—against a gold standard.

Second, reliability estimates may be differentially affected by variation in both the temporal stability of health behaviors, knowledge, and attitudes and the different time intervals assessed (Brener *et al.* 1995; Brener *et al.* 2002). Moreover, reliability estimates for behaviors or attitudes that may vary over time or with situational context are likely to be conservative.

Third, our all female, 15 to 20 year old, ethnic Fijian study sample is not representative of school-going Fijian youth but rather reflects the socio-demographic strata of key interest in addressing aims of the main study designed to examine the impact of rapid economic change on disordered eating in ethnic Fijian adolescent girls. Moreover, the age range—15 to 20 years old—differs from the 13 to 15 year old range for which the GSHS was developed. However, the majority of countries (11 of 13) issuing full reports on their GSHS have also included respondents greater than 16 years of age. Likewise, the 2003 implementation of the YRBS in three selected Pacific territories included respondents through grade 12 (Balling *et al.* 2003). Although the impact of age on reliability is unknown in other populations, we suggest that older age and more education would be likely to improve comprehension of the items, and thereby improve reliability. Because these and other demographic characteristics each may influence reliability in unknown ways, our study findings cannot be generalized to other Fijian youth. Indeed, we emphasize the general desirability of population-specific evaluation of GSHS reliability.

Fourth, we were unable to assess the impact of our translation on reliability directly since subjects responded in only one language. Apparent language differences in the reliability may have been confounded by other characteristics of study participants who chose English over the vernacular. We identified only a few sources addressing translation of the GSHS; one study reports a protocol for validating a Spanish translation of the GSHS (Granero *et al.* 2007) and we found a reference to translation methods in only 4 country reports on the WHO website (for Kiswahili and Arabic languages only) (WHO 2009).

Fifth, the reliability of risk behavior assessment with this adapted version of GSHS content may have limited comparability with the reliability of GSHS-based assessment elsewhere, since study procedures were developed to address study-specific objectives and without the benefit of implementation workshop training provided by WHO. Specifically, several features of our overarching study protocol may have had unique impact on reliability of risk behavior assessment with this adapted version of GSHS content. In particular, exposure to other study assessments addressing overlapping topic content may have differentially impacted responses at either Time 1 or Time 2. If so, this was especially likely to have an impact on items relating to dietary behaviors and may be consistent with our finding that reliability for items relating to dietary behaviors was lower than for other topics in this study. Alternatively, it is possible that temporal instability of these behaviors resulted in actual differences in prevalence at Time 1 and Time 2. Because items relating to nutritional behaviors had comparatively low reliability in both our study and those reported for the YRBS, we suggest they warrant special caution for interpretation until reliability and validity are locally established.

Finally, the version of GSHS content developed for assessment in this study preceded an English version of the GSHS that has subsequently been developed for Fiji and made available on the WHO website (WHO 2009). Although core module questions in common

between these two versions were identically worded, some optional items and wording for some country-specific examples (e.g. vegetable names) differed.

This study presents test-retest reliability data for a version of GSHS content adapted for an ethnic Fijian female study population. Although risk behavior assessment utilizing adapted GSHS content for this study was implemented in a study-specific protocol with unique features, these reliability data contribute to the growing body of literature on the GSHS. Most notably, whereas reliability for items assessing health risk behaviors was generally lower in this study than has been reported for health risk behavior assessment utilizing the YRBS, the majority of the items showed at least fair reliability over a one-week interval. Moreover, core module items relating to tobacco use, alcohol use, and sexual risk behaviors demonstrated substantial reliability, suggesting that the GSHS may be a useful tool for population health surveillance of these behaviors in this study population. We suggest that caution is warranted in interpreting data from GSHS questions related to dietary behaviors among ethnic Fijian adolescent girls. Comparison of health risk survey data on alcohol and tobacco use among Fijian youth suggests that ethnicity, age, and/or mode of assessment may contribute to variation in prevalence estimates (Becker *et al.* in press). Therefore, psychometric evaluation of assessment performance within strata defined by ethnicity and age may be an essential first step in valid health risk surveillance among Fiji's ethnically diverse youth.

These study results may also have implications for the reliability of GSHS-based health risk behavior assessment data in populations outside of Fiji. In particular, the lower reliability of responses to GSHS content-based items in this ethnic Fijian population as compared with the reported reliability of assessment with the YRBS in a U.S. population supports the need for caution in interpreting data generated by country-specific versions of the GSHS with unknown reliability. The wide variation in the social context of assessment—including the cultural variability of sensitivity of risk behavior and comfort with self-disclosure—as well as conceptual and linguistic challenges to reporting in a non-native language or in translation may impact reliability and validity of assessment with the GSHS.

Notwithstanding the necessity for caution suggested by our findings, this study also provides important positive data relevant to the GSHS. In particular, this study confirms adequate test-retest reliability of a version of GSHS content in at least one study sample. Second, this study demonstrates the reliability of a non-English version of GSHS item content. Third, this study demonstrates a simple and feasible protocol for evaluating population-specific reliability of GSHS content. The minimal burden of evaluating reliability across culturally diverse populations appears justifiable in light of the potential benefits of understanding the relative strengths and limitations of health data yielded by the GSHS within these respective populations. School-based self-report surveys hold promise as cost-effective and systematic means of obtaining behavioral health-risk data on adolescents. Further studies on country, ethnicity, and language-specific reliability of GSHS content can provide critical support for interpreting these data.

Key messages

1. A GSHS strength is adaptability for implementation across culturally diverse populations. However, GSHS reliability has not previously been evaluated, thereby limiting the utility of potentially invaluable health risk data it yields.
2. Study findings support adequate retest reliability of GSHS content adapted for ethnic Fijian girls for assessing several risk behaviors, especially alcohol use

and sexual risk. Reliability for dietary behaviors was comparatively low and warrants caution in interpreting their relevance.

3. Local psychometric evaluation of the GSHS may be advisable when implementing it across ethnically distinct populations. A simple protocol for assessing retest reliability was feasible in Fiji.

Acknowledgments

Supported by K23 MH068575, a Harvard University Research Enabling Grant, and fellowship support from the Radcliffe Institute for Advanced Study (AEB). We gratefully acknowledge the assistance of Dr Lepani Waqatakiwewa, CEO - Fiji Ministry of Health and his team; the Fiji Ministry of Education; Joana Rokomatu, the *Tui Sigatoka*; Dr. Jan Pryor, Chair of the FN-RERC; Professor Paul Geraghty; and Dr. Tevita Qorimasi. We thank Professor Jane Murphy, Dr. Deborah Blacker, Dr. Gene Beresin, Dr. Jennifer Derenne, Kesaia Navara, and Aliyah Shivji. We are grateful to members of the Senior Advisory Group for the HEALTHY Fiji Study (Health-risk and Eating attitudes and behaviors in Adolescents Living through Transition for Healthy Youth in Fiji Study), including Professor Bill Aalbersberg (Chair), Nisha Khan, Alumita Taganesia, Livinai Masei, Asenaca Bainivualiku, Pushpa Wati Khan, and Fulori Sarai. Finally, we thank all the Fiji-based principals and teachers who facilitated this study.

References

- Balling, A.; Grunbaum, JA.; Speicher, N.; McManus, T.; Kann, L. Youth Risk Behavior Survey 2003: Commonwealth of the Northern Mariana Islands, Republic of the Marshall Islands, Republic of Palau. 2003 [14 April, 2009]. [online]. Available from: <http://www.cdc.gov/HealthyYouth/YRBS/pdf/pacific-islands.pdf>
- Becker, AE. Body, Self, and Society: The View from Fiji. Philadelphia: University of Pennsylvania Press; 1995.
- Becker AE, Burwell RA, Gilman SE, Herzog DB, Hamburg P. Eating behaviours and attitudes following prolonged television exposure among ethnic Fijian adolescent girls. *The British Journal of Psychiatry* 2002;180:509–14. [PubMed: 12042229]
- Becker AE. Television, disordered eating, and young women in Fiji: Negotiating body image and identity during rapid social change. *Culture, Medicine and Psychiatry* 2004;28:533–59.
- Becker AE, Perloe A, Richards L, Roberts AL, Bainivualiku A, Khan AN, Navara K, Gilman SE, Aalbersberg W, Striegel-Moore RH, HEALTHY Fiji Study Group. Prevalence and Socio-demographic Correlates of Cigarette Smoking, Alcohol Use, and Unsafe Sexual Behavior among Ethnic Fijian Secondary Schoolgirls. *Fiji Medical Journal*. in press.
- Becker AE, Thomas JJ, Bainivualiku A, Richards L, Navara K, Roberts AL, Gilman SE, Striegel-Moore RH. Adaptation and evaluation of the Clinical Impairment Assessment to assess disordered eating related distress in an adolescent female ethnic Fijian population. *International Journal of Eating Disorders*. 2009 online. 10.1002/eat.20665
- Blum RW, Nelson-Mmari K. The health of young people in a global context. *Journal of Adolescent Health* 2004;35:402–418. [PubMed: 15488435]
- Brener ND, Collins JL, Kann L, Warren CW, Williams BI. Reliability of the Youth Risk Behavior Survey Questionnaire. *American Journal of Epidemiology* 1995;141:575–80. [PubMed: 7900725]
- Brener ND, Kann L, McManus T, Kinchen SA, Sundberg EX, Ross JG. Reliability of the 1999 Youth Risk Behavior Survey Questionnaire. *Journal of Adolescent Health* 2002;31:336–342. [PubMed: 12359379]
- Brener ND, Billy JOG, Grady WR. Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature. *Journal of Adolescent Health* 2003a;33:436–457. [PubMed: 14642706]
- Brener ND, McManus D, Galuska R, Lowry R, Wechsler H. Reliability and validity of self-reported height and weight among high school students. *Journal of Adolescent Health* 2003b;32:281–287. [PubMed: 12667732]

- Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *Journal of Clinical Epidemiology* 1993;46:423–9. [PubMed: 8501467]
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960;20:37–46.
- Due P, Holstein BE. Bullying victimization among 13 to 15 year old school children: Results from two comparative studies in 66 countries and regions. *International Journal of Adolescent Medicine and Health* 2008;20:209–221. [PubMed: 18714557]
- Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *Journal of the American Medical Association* 2000;283:2701–2711. [PubMed: 10819955]
- Granero R, Poni ES, Sanchez Z. Sexuality among 7th, 8th and 9th grade students in the state of Lara, Venezuela. *The Global School Health Survey, 2003-2004. Puerto Rican Health Sciences Journal* 2007;26:213–219.
- Kraemer HC, Periyakoil VS, Noda A. Tutorial in biostatistics: Kappa coefficients in medical research. *Statistics in Medicine* 2002;21:2109–29. [PubMed: 12111890]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74. [PubMed: 843571]
- Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997;349:1436–42. [PubMed: 9164317]
- Muula AS, Kazembie LN, Rudatsikira E, Siziya S. Suicidal ideation and associated factors among in-school adolescents in Zambia. *Tanzania Health Research Bulletin* 2007;9:202–06. [PubMed: 18087900]
- National Centers for Chronic Disease Prevention and Health Promotion. YRBSS: Youth Risk Behavior Surveillance System. 2008a [6 October, 2008]. Available at: <http://www.cdc.gov/HealthyYouth/yrbs/>
- National Centers for Disease Control and Prevention. Global Youth Tobacco Surveillance, 2000-2007. *Morbidity and Mortality Weekly Report* 2008b;57:1–21.
- Rudatsikira E, Seter S, Muula AS. Suicidal ideation and associated factors among school-going adolescents in Harare, Zimbabwe. *Journal of Psychology in Africa* 2007a;17:93–98.
- Rudatsikira E, Ogwel AEO, Siziya S, Muula AS. Prevalence of sexual intercourse among school-going adolescents in Coast Province, Kenya. *Tanzania Health Research Bulletin* 2007b;9:159–63. [PubMed: 18087892]
- Rudatsikira E, Muula AS, Siziya S, Twa-Twa J. Suicidal ideation and associated factors among school-going adolescents in rural Uganda. *BMC Psychiatry* 2007c;7:67. [PubMed: 18034906]
- Singh AK, Maheshwari A, Sharma N, Anand K. Lifestyle associated risk factors in adolescents. *Indian Journal of Pediatrics* 2006;73:901–906. [PubMed: 17090902]
- Turagabeci AR, Nakamura K, Takano T. Healthy lifestyle behaviour decreasing risks of being bullied, violence and injury. *PloS One* 2008;3:e1585. [PubMed: 18297133]
- World Health Organization. Geneva, Switzerland: 2003 [7 July 2009]. Strategic Directions for Improving the Health and Development of Children and Adolescents. online Available from: <http://whqlibdoc.who.int/publications/2003/9241591064.pdf>
- World Health Organization. Global School-based student health survey (GSHS). 2009 [17 April through 12 August, 2009]. [online]. Available from: <http://www.who.int/chp/gshs/en/>
- Zullig KJ, Pun S, Patton JM, Ubbes VA. Reliability of the 2005 Middle School Youth Risk Behavioral Survey. *Journal of Adolescent Health* 2006;39:856–60. [PubMed: 17116516]

Table 1

Demographic characteristics and selected risk behaviors of the HEALTHY Fiji sample and the retest reliability sub-sample

	HEALTHY Fiji sample (N=523)		Retest sub-sample (N=81)*	
	N	%	N	%
Age				
15-16	339	64.8	54	66.7
17-18	170	32.5	23	28.4
19-20	14	2.7	4	4.9
Language of questionnaire				
English	147	28.11	21	25.9
Fijian (local vernacular)	376	71.9	60	74.1
School location				
Peri-urban	262	50.1	35	43.2
Rural	261	49.9	46	56.8
Sexual intercourse ever				
No	408	79.2	62	80.5
Yes	107	20.8	15	19.5
Alcohol use, past 30 days				
None	417	79.9	70	87.5
Any	105	20.1	10	12.5
Cigarette use, past 30 days				
None	430	82.4	61	75.3
Any	92	17.6	20	24.7

* Retest sub-sample is not statistically different from the HEALTHY Fiji sample for any listed characteristic, t-test at $p < .05$ level of significance.

Table 2

Reliability of adapted GSHS items¹ in a sub-sample of the HEALTHY Fiji sample of adolescent Fijian girls (N=81)

Item	Prevalence, Time 1 (n)	Prevalence, Time 2 (n)	Kappa	Percent Agreement	N
Alcohol use					
Drank alcohol, past 30 days*	12.5 (80)	12.4 (81)	.66	93	80
≥1 drink per day on drinking days, past 30 days*	8.8 (80)	7.4 (81)	.75	96	80
Method of procuring alcohol*	Friends, 10.1 ² (80)	Friends, 6.2 ² (81)	.76	95	81
Ever been drunk*	22.5 (80)	19.8 (81)	.70	90	80
Dietary behaviors					
Hungry sometimes or mostly*	55.6 (81)	25.9 (81)	.34	65	81
Times/day ate fruit, past 30 days*	At least 1/day: 66.7 ² (81)	At least 1×/day: 59.4 ² (81)	.31	33	81
Times/day ate vegetables, past 30 days*	At least 1/day: 88.9 ² (81)	At least 1×/day: 82.9 ² (81)	.33	38	81
Ate breakfast rarely or never, past 30 days	15.0 (80)	11.1 (81)	.18	81	80
Perceive self as overweight	68.0 (78)	62.5 (80)	.49	77	77
Trying to lose weight	51.9 (79)	61.7 (81)	.52	76	79
Drank soda, past 30 days	62.5 (80)	45.6 (79)	.50	74	78
Ate fast food, past 7 days	43.8 (80)	37.0 (81)	.51	76	80
Been weighed and measured, past 12 months	19.8 (81)	36.3 (80)	.13	64	80
Exercised to lose or keep from gaining weight, past 30 days	45.7 (81)	32.5 (80)	.51	76	80
Ate less food, fewer calories, or less fat to lose or keep from gaining weight, past 30 days	47.5 (80)	42.0 (81)	.75	88	80
Fasted to lose or keep from gaining weight, past 30 days	13.8 (80)	14.8 (81)	.44	86	80
Vomited or took laxatives to lose or keep from gaining weight, past 30 days	3.7 (81)	5.0 (80)	.25	94	80
Took diet pills, powders or liquids to lose or keep from gaining weight, past 30 days	3.8 (79)	1.3 (80)			79
Weight change, past 30 days	DK, 62.5 ² (80)	DK, 63.8 ² (80)	.32	63	80
Exercised to gain weight, past 30 days	20.0 (80)	11.1 (81)	.11	76	80
Ate more food, calories or fat to gain weight, past 30 days	27.9 (79)	18.5 (81)	.48	81	79
Took diet pills, powders or liquids to gain weight, past 30 days	0.0 (80)	2.5 (79)			79
Taught benefits of healthy eating, this school year	71.6 (81)	75.3 (81)	.19	65	81
Taught benefits of eating fruits and vegetables, this school year	80.3 (81)	76.5 (81)	.35	77	81
Taught healthy ways to gain weight, this school year	27.5 (80)	29.6 (81)	.34	58	80

Item	Prevalence, Time 1 (n)	Prevalence, Time 2 (n)	Kappa	Percent Agreement	N
Taught healthy ways to lose weight, this school year	39.2 (79)	42.0 (81)	.57	72	79
Taught how to make healthy meals and snacks, this school year	58.2 (79)	62.5 (80)	.41	68	78
Mental health					
Felt lonely, past 12 months*	10.0 (80)	6.2 (81)	.42	91	80
So worried that couldn't sleep, past 12 months*	7.6 (79)	3.7 (81)	.41	94	79
Sad/hopeless, stopped doing usual activities for ≥ 2 weeks, past 12 months*	46.3 (80)	29.6 (81)	.41	71	80
Seriously considered attempting suicide, past 12 months*	9.9 (81)	7.4 (81)	.69	95	81
Made a plan about attempting suicide, past 12 months*	12.4 (81)	9.9 (81)	.63	93	81
No close friends*	1.2 (81)	1.2 (81)			81
Attempted suicide, past 12 months	1.2 (81)	3.7 (81)			81
Made injurious suicide attempt, past 12 months	0.0 (79)	2.5 (81)			79
Physical activity					
Days physically active, past 7 days*	1.6 ³ (80)	1.3 ³ (81)	.42	56	80
Days physically active, typical week*	1.8 ³ (80)	1.5 ³ (81)	.49	56	80
Hours/day sitting activities on typical day >2 *	18.2 ⁴ (77)	17.5 ⁴ (80)	.30	79	77
Days walked or biked to school, past 7 days*	2.7 ³ (78)	2.6 ³ (79)	.50	73	78
Time to/from school, past 7 days*	≥ 60 min: 34.7 ² (75)	≥ 60 min: 22.7 ² (79)	.45	41	75
Absence of Protective Factors					
Miss class or school without permission, past 30 days*	60.0 (80)	53.1 (81)	.75	88	80
Most students in school were never or rarely kind and helpful, past 30 days*	17.5 (80)	18.8 (80)	.45	84	80
Parents never or rarely checked if homework done, past 30 days*	17.7 (79)	20.0 (80)	.42	82	79
Parents never or rarely understood your problems or worries, past 30 days*	22.8 (79)	40.3 (77)	.39	73	77
Parents never or rarely knew what you were doing in your free time, past 30 days*	30.8 (78)	38.8 (80)	.52	78	78
Sexual behaviors					
Ever had sexual intercourse*	19.5 (77)	20.0 (80)	.92	97	77
Age, first intercourse, in those reporting intercourse*	Mean age: 15.4 (17)	Mean age: 15.2 (16)	.72	79	14
Two or more lifetime sexual partners*	1.3 (77)	2.5 (80)			77
Sexual intercourse, past 12 months*	10.3 (78)	7.5 (80)	.69	95	78

Item	Prevalence, Time 1 (n)	Prevalence, Time 2 (n)	Kappa	Percent Agreement	N
Condom use during last intercourse [*]	5.19 (78)	6.25 (80)	.41	94	78
Tobacco					
Smoked first cigarette at age < 13 years [*]	3.7 (81)	3.7 (81)	.65	98	81
Smoked cigarettes, past 30 days [*]	24.7 (81)	19.8 (81)	.72	90	81
Used other forms of tobacco, past 30 days [*]	10.0 (80)	6.2 (81)	.50	90	80
Tried to quit smoking, past 12 months [*]	25.0 (80)	58.0 (81)	.59	84	80
Days people smoked in your presence, past 7 days [*]	Some or more: 71.6 ² (81)	Some or more: 58.0 ² (81)	.37	46	81
At least one parent or guardian uses any form of tobacco [*]	50.0 (78)	50.6 (80)	.87	94	78
See actors smoking most of the time or always	25.3 (80)	18.5 (81)	.29	75	80
See cigarette brand names in advertisements, past 30 days	>=sometimes: 42.5 ² (80)	>=sometimes: 40.0 ² (81)	.27	43	80
See cigarette billboard ads, past 30 days	Any: 58.8 ² (80)	Any: 48.1 ² (81)	.46	67	80
Think girls who smoke have more friends	71.3 (80)	56.3 (80)	.38	65	79
Think smoking is harmful to health	Definitely yes: 74.1 ² (80)	Definitely yes: 80.3 ² (80)	.23	74	80
Most or all of closest friends smoke cigarettes	25.0 (80)	21.3 (80)	.52	68	80
Violence and unintentional injury					
Physically attacked, past 12 months [*]	17.3 (81)	4.9 (81)	.28	85	81
In physical fight, past 12 months [*]	28.4 (81)	21.0 (81)	.21	70	81
Activity during most serious injury, past 12 months [*]	Sports: 7.4 ² (81)	Sports: 4.9 ² (81)	.43	78	81
Cause of most serious injury, past 12 months [*]	Fell: 11.1 ² (81)	Fell: 7.4 ² (81)	.61	88	81
Perpetrator and deliberateness of most serious injury, past 12 months [*]	Self accident: 3.8 ² (80)	Self accident: 9.9 ² (81)	.52	88	80
Type of most serious injury in past 12 months [*]	Broken bone: 6.3 ² (80)	Broken bone: 7.4 ² (81)	.47	86	80
Bullied, past 30 days [*]	15.2 (79)	7.4 (81)	.38	87	79

¹ Approximate phrasing, behavior, or attitude assessed by the item;

^{*} denotes items drawn from GSHS core module content.

² Prevalence is given for most common response.

³ Mean number of days

⁴ Mean number of hours

Table 3

Mean kappa coefficients for GSHS content by topic, age, and language of survey

Category	Mean Kappa	Mean kappa, core items only
All items	.47	.52
By Topic		
Alcohol use	.72	.72
Dietary behaviors	.38	.33
Mental Health	.51	.51
Physical Activity	.43	.43
Protective Factors	.51	.51
Sexual Behaviors	.69	.69
Tobacco	.49	.62
Violence and injury	.41	.41
By Age		
15 years (n = 30)	0.42	.45
16-20 years (n = 51)	0.49	.53
By Language of Survey		
English (n = 21)	0.53	.58
Local dialect (n=60)	0.41	.46