

Reducing Response Bias in Reports of Trauma and PTSD: An Application of the Nonverbal Response Card in a Survey of Youth in Burkina Faso

David P. Lindstrom¹, Guy Harling²

Abstract

We present results on response bias in reports of life time trauma and PTSD from an experimental trial of the nonverbal response card (NVRC) in a survey of 12-20 year olds in Burkina Faso. The NVRC is a laminated two-sided card that allows respondents to nonverbally respond to questions without the interviewer knowing the actual response. It is easy to use, highly portable and does not require literacy. The NVRC was randomly assigned to 50% of the sample and used for sensitive questions on self-harm, depression, trauma, PTSD, first sexual intercourse and unwanted sex. Respondents who used the NVRC reported 18% more life time trauma types than verbal respondents, and they were 2.8 times as likely to report 3-4 PTSD symptoms. The reports of trauma and PTSD by respondents who used the NVRC compared to the verbal method had higher internal consistency reliability and several types of PTSD correlated more strongly with reports of self-harm and depression. The question-specific error rates for the NVRC were on average higher than the rates for the verbal method, but declined across questions as respondents became more comfortable with the method, and the mean error rate for the NVRC method was well below 1 percent.

Key Words: Trauma, PTSD, response mode, social desirability bias

¹ Population Studies and Training Center, Brown University.

² Institute for Global Health, University College London

1. Introduction

Trauma is broadly defined as exposure to actual or threatened death, serious injury, or sexual violence, which can occur witnessing these events, learning that the event has happened to someone close, or directly experiencing the event (American Psychiatric Association 2022, also see Weathers and Keane 2007). Survey questions and diagnostic check lists of trauma vary with respect to the types of events, the period covered and the frequency of events (Amaya-Jackson et al. 2000, McDonald et al. 2014). Nevertheless, studies reveal that life time experience of some form of trauma is very common (Gray et al. 2004; Kessler et al. 1995; Kilpatrick et al. 2013). PTSD is marked by the presence of one or more of the following symptoms brought about by a traumatic event or events: recurrent distressing memories or dreams, persistent avoidance reminders, negative changes in thoughts and moods, and alternations in arousal and reactivity (American Psychiatric Association 2022). Trauma is a leading cause of PTSD, although most people who experience trauma do not experience PTSD (Breslau et al. 1998; Winders et al. 2019). However, for populations exposed to natural disasters, personal insecurity, criminal violence, and military conflicts, the prevalence of PTSD can be considerably higher. Trauma and PTSD have been linked to a number of negative psychological and health outcomes that can extend across the lifespan (Beck et al. 2009; Olatunji et al. 2007; Steel et al. 2011), as well as high use of health services (Solomon and Davidson 1997).

Studies find significant gender differences in reports of trauma and PTSD (Ditlevsen and Elklit 2012; Tolin and Foa 2006). The most common findings are that men are more likely than women to report trauma due to non-sexual forms of violence and accidents, whereas women are more likely than men to report trauma related to sexual violence (Haldane and Nickerson 2016; Kilpatrick et al. 2013; Norris et al. 2003; Tjaden and Thoennes 1998). Among males and females who report trauma, women are significantly more likely than men to report PTSD (Haldane and Nickerson 2016; Karatzias et al. 2017; Kilpatrick et al. 2013 McDonald et al. 2014; Steel et al. 2011; Tolin and Foa 2006). One possible explanation that has been put forward for this difference in reported PTSD symptoms are gender norms that depict men as less vulnerable than women to certain types of traumatic events and emphasize men's resiliency and ability to psychologically withstand trauma (Saxe and Wolfe 1999).

These same gender-typed behaviors and attributes that differentiate men and women in terms of how trauma and PTSD is processed and expressed, may also contribute to differences in the willingness to report certain forms of trauma and PTSD in the context of an interviewer administered survey. The willingness to provide accurate reports of trauma and PTSD might especially be problematic for experiences and feelings that evoke shame and embarrassment, which are important facets of trauma and PTSD (Beck et al. 2011; Leskela et al. 2002; Øktedalen et al. 2014). Studies of trauma and PTSD have identified shame, embarrassment and concerns about the judgement of others as reasons why trauma and PTSD may be underreported (MacDonald and Morley 2001; Rasmussen et al. 2007). Survey researchers refer to the inaccurate reporting of certain behaviors, experiences and attitudes as social desirability bias (Tourangeau et al. 2000). Social desirability bias can take the form of both underreporting and overreporting depending on the behavior or attitude, the characteristics of the respondent, and the interview context. Researchers have found that respondents tend to underreport embarrassing or stigmatized behaviors such as nonmarital sex, multiple sexual partners, substance abuse, abortion, and sexual violence (Jones and Forest 1992, Lindstrom et al. 2010; Lindstrom et al. 2012; Mensch et al. 2008) and over-report normative behaviors or attitudes such as contraceptive knowledge, healthy

behaviors and non-racist attitudes (Krysan 1998; Lindstrom et al. 2010). In addition to social desirability, respondents' desire to protect their privacy from intrusive questions and concerns about the risk of disclosure are other reasons for response bias with sensitive questions (Amaya-Jackson et al. 2000; Tourangeau et al. 2000).

Research on measuring trauma and PTSD has recognized social desirability as a potential source of bias (Hecker et al. 2015; Nader 2008; Weathers and Keane 2007). Studies of children for example have identified peer and group affects as resulting in overreporting of certain forms of trauma (Nader 2008, 91). Similarly, investigators have pointed out the potential for refugees in particular settings (Hollifield et al. 2002) and veterans (Kubany et al. 2000; Weathers and Keane 2007) to overreport trauma and PTSD in expectation of advantages. Other investigators have suggested that especially vulnerable individuals fleeing unrest and conflict may be reluctant to report traumatic experiences (Kienzler 2008), and studies have pointed out the potential for underreporting of traumatic events and PTSD in clinical contexts due to shame and embarrassment (Bremner et al. 2000). In spite of this recognition, very little has been actually done in trauma and PTSD research to measure the magnitude of social desirability bias and to reduce it.

We present results from a randomized trial experiment of the use of a nonverbal response card (NVRC) in an interviewer administered survey of youth ages 12-20 conducted in Burkina Faso. The nonverbal response card provides an additional layer of privacy and confidentiality beyond what is available for the conventional verbal response method and its effectiveness in reducing social desirability bias in surveys of adolescents and young adults has been demonstrated in other contexts (Aichele et al. 2014; Harling et al. 2020; Lindstrom et al., 2010; Lindstrom et al., 2012).

The nonverbal response card was randomly assigned in advance of interviewing to 50% of the sample for questions about trauma, sexual violence, mental health and PTSD, and the verbal response method was used by the other half of the sample. We first examined reports of life time trauma and PTSD by response method and by gender. Second, we evaluated the trauma and PTSD reports for internal consistency reliability and convergent validity within response method. Third, we examined question-specific and interviewer error rates by response method.

2. Methods

2.1 The Nonverbal Response Card

The nonverbal response card (NVRC) was developed and first tested in an interviewer-administered survey questionnaire on sexual behavior and relationships in a sample of Ethiopian youth (Lindstrom et al. 2020; Lindstrom et al. 2012). The laminated, two-sided NVRC is held by the respondent so that one side is visible to the respondent and the other side is visible to the interviewer. Each side of the card is divided into rows and columns of cells with a small hole punched through the center of each cell. The cells on the respondent side of the card contain the set of valid responses, typically "yes/no" and a range of numeric responses, and the interviewer side of the card has unique three-digit numbers in each of the cells (Figure 1). Alternating cells on the interviewer side of the card are shaded in yellow to make it easier for the interviewer to visually distinguish between adjacent cells. The respondent indicates a response by inserting a stick through the hole in the appropriate response cell. The interviewer records the three-digit code corresponding to the cell through which the stick protrudes, without knowing the actual response.

The card is divided into two panels; a “yes/no” panel with multiple “yes/no” cells, and a numeric panel. The division of the card into two panels allows the interviewer to detect if a respondent is distracted and providing a numeric response to a “yes/no” question or vice versa. In addition to containing “yes/no” written in the study population language(s), the “yes” cells are shaded green and the “no” cells are shaded red to facilitate the use of the card by illiterate respondents. The numeric response cells contain hash marks along with Arabic numerals to facilitate the use of the card by innumerate respondents. Blue shading in the numeric cells can also be used for scale responses, such as 0 to 5 for “never” to “always” or for Likert scales.

Procedurally, the interviewer hands the respondent a large envelope containing multiple versions of the card and a response stick, and instructs the respondent to remove and inspect the cards. Each version of the card has a different arrangement of “yes/no” cells and different starting points for the sequence of numeric responses. The provision of multiple versions of the card is intended to assure the respondent that the interviewer will not know the actual response to any selected cell because each card is unique. The respondent is instructed to select any card, and the interviewer then uses a demonstration card to show how the card works. The three-digit codes are unique across all cells and all cards, and are coded with the appropriate response at the data processing stage.

The NVRC is inexpensive to produce, portable, easy to use, and it provides increased privacy. In contrast to self-administered questionnaires or computer-based response methods, the NVRC does not require literacy nor familiarity with a computer keyboard or touch screen. In two experimental tests of the NVRC in Ethiopia, respondents who used the card were more likely than verbal respondents to report sexual coercion, forced sex and nonmarital sex, and less likely to report condom knowledge (Lindstrom et al., 2010, Lindstrom et al., 2012). In a survey of Tanzanian youth, females who used the NVRC were more likely than female verbal respondents to report first sex, early ages at first sex, having been tested for HIV, and more lifetime sexual partners (Aichele et al. 2014). In a survey of youth in Burkina Faso that is the source of data for the present study, respondents who used the NVRC were more likely than verbal respondents to report lifetime sexual assault, other unwanted sexual experience, attempted forced sex, and forced sex (Harling et al. 2020).

2.2 Study Population

The NVRC was used in the first wave of the Africa Research, Implementation Science and Education (ARISE) adolescent survey in Burkina Faso. ARISE is a collaboration between nine sub-Saharan African institutions in seven countries, the Harvard University T.H. Chan School of Public Health and the University of Heidelberg. The ARISE Burkina Faso survey was conducted within the Nouna Health and Demographic Surveillance System (HDSS) overseen by the Centre de Recherche en Santé de Nouna (Sié et al., 2010). The Nouna HDSS comprises 59 villages and Nouna town. The ARISE study randomly sampled 2,544 youth aged 12-20 from the most recent HDSS census living in 10 villages and 7 sectors of Nouna town that were purposively selected to represent the ethnic make-up of the surveillance area. A total of 1,644 youth were located and agreed to participate in the ARISE survey. Youth aged 18 and above provided written consent and for youth under age 18, written consent was provided by a parent or guardian along with the assent of the youth. Interviews were conducted in 2017 at the youth’s home in French or a local language. The NVRC was randomly assigned to 50 percent of the sample in advance of interviewing, and the other half of the sample was assigned the verbal response method. Table 1 presents selected characteristics of the verbal and NVRC samples. The verbal sample had a larger percentage of female respondents than the NVRC sample (44.9% compared to 39.5%).

However, the two samples were not significantly different from one another with respect to age, education, marital status and religion.

The Boucle du Mouhoun region within Nouna abuts the Ségou and Mopti regions of Mali, areas which had seen substantial violence and displacement due to religious-political conflict and intercommunal violence. Warfare within Burkina Faso had been limited, but kidnappings and intercommunal violence rose in the period leading up to the ARISE survey, and armed violence in the Nouna district displaced some households.

2.3 Questionnaire

The ARISE questionnaire collected information on sociodemographics, family resources, diet, physical and mental health, body image, substance use and sexual and reproductive health. The NVRC was used for a section of sensitive questions that were placed at the end of the questionnaire on self-harm, depression, trauma, PTSD, first sexual intercourse and unwanted sex. Trauma was measured with 15 “yes/no” questions based on the commonly used Life Events Checklist (Weathers et al., 2013, Winders et al. 2019) , and PTSD was measured with the four-question Primary Care PTSD screen which included measures of the B, C, D, and E criteria of the American Psychiatric Association’s diagnostics for posttraumatic stress disorder (Cameron & Gusman, 2003; American Psychiatric Association 2022), and in the Harvard HTQ-5 revision of the Harvard Trauma Questionnaire (Berthold et al. 2018). The questionnaire was administered by interviewers using digital tablets.

2.4 Data Analysis and Statistical Methods

We first compared the distributions of responses to the trauma and PTSD questions for the verbal and NVRC methods. We next used Cronbach’s alpha to measure internal consistency reliability for the two response methods (Hollifield et al. 2002). Although some investigators have used Cronbach’s alpha to evaluate the internal consistency of trauma scales (Bae et al. 2008; Bremner et al. 2000), others have argued that the assessment of the internal consistency for questions on trauma is inappropriate because of the disparate nature of many trauma events (Gray et al. 2004; Netland 2001). Our objective in using Cronbach’s alpha was not to evaluate the internal consistency of the trauma questions used in Burkina Faso survey relative to other trauma scales or other populations, but rather to compare response patterns by response method. Our assumption was that individuals who were exposed to one type of trauma were likely to be exposed to related traumas (see Haldane and Nickerson 2016; Kilpatrick et al. 2013), and that respondents who used the nonverbal response card were more likely than respondents who used the verbal method to provide an accurate reporting of more sensitive or potentially stigmatized trauma types. To this end, we grouped related types of trauma into three domains: disaster/accident, assault, and war/violence, and conducted within domain comparisons of Cronbach’s alpha by response method as well as comparisons of Cronbach’s alpha for all 15 trauma types together. We used the chi-square test described in Feldt et al. (1987) to test for the statistical significance of differences in Cronbach’s alpha.

We also estimated the correlation between the number of reported trauma types and the number of reported PTSD symptoms as an additional measure of reliability. Prior research shows a positive correlation between the variety of trauma events and the risk and severity of PTSD (Gray et al. 2004; Kilpatrick et al. 2013; Kolassa et al. 2010; Rasmussen et al. 2020). We used the bivariate correlations of each of the four PTSD questions with reports of ever self-harm and felt life not worth living, both of which are associated with PTSD (American Psychiatric Association 2022), to check for convergent validity. Convergent

validity refers to the extent to which the responses on a measure are similar to the responses on another accepted measure of the same underlying construct (Steel et al. 2011). Our expectation was that the inter-item correlations of the trauma and PTSD measures, and the PTSD measures and self-harm and depression would be higher among the NVRC respondents compared to the verbal respondents if the NVRC method reduced social desirability bias and increased accuracy. Finally, we examined question- and interviewer-specific error rates by response method. We used a threshold *P* value of less than 0.05 to indicate statistical significance in our tests. We conducted the analyses using SPSS version 26.

3. Results

3.1 Reports of Trauma and PTSD

Table 2 presents the percentages of “yes” responses to the 15 trauma questions by response method. On 9 of the 15 questions, the percentage of “yes” responses was significantly higher for respondents who used the NVRC compared to the verbal response method (6 of the differences were statistically significant using the more conservative Bonferroni test). The mean number of reported life time trauma types was also significantly higher among respondents who used the NVRC. The difference in means by method was greater for female than male respondents, and only significant among female respondents. Among respondents who used the verbal response method, females reported significantly fewer trauma event types than males (1.36 vs. 1.61). Among respondents who used the NVRC, the means for females and males (1.86 vs. 1.71) were not significantly different.

The number of reported PTSD symptoms experienced in the last 30 days was also significantly higher among respondents who used the NVRC compared to the verbal method. On 3 of the 4 symptoms, the percentage of “yes” responses was significantly higher among NVRC respondents compared to verbal respondents. The percentages of female and male respondents who reported 3-4 symptoms were also significantly higher among NVRC respondents compared to verbal respondents. The percentage of respondents who reported 3-4 PTSD symptoms was higher among females compared to males for both response methods, but the difference was only statistically significant among verbal respondents.

In summary, females who used the NVRC method compared to the verbal method reported more trauma types and PTSD symptoms. Among male respondents, the response method made the most difference in reports of PTSD: 7% of male respondents who used the NVRC reported 3-4 PTSD symptoms compared to 1% of male respondents who used the verbal method. Significant gender differences in reported trauma and PTSD were present among verbal respondents, but not among respondents who used the NVRC method.

3.2 Reliability and Validity

On all three of the trauma domains and for all 15 trauma types, Cronbach’s alpha was significantly higher for NVRC respondents compared to verbal respondents (Table 3). Similarly, Cronbach’s alpha for the 4 PTSD questions was significantly higher among NVRC respondents (0.708) compared to verbal respondents (0.518). The results from Tables 2 and 3 indicate that not only were NVRC respondents more likely than verbal respondents to report life time experiences of trauma and recent symptoms of PTSD, the internal consistency reliability of their reports was considerably higher as well.

With respect to gender, Cronbach's alpha for the 15 trauma types was also significantly higher among females and males who used the NVRC method compared to the verbal method. However, Cronbach's alpha was not significantly different between females and males who used the verbal method, nor was it significantly different between females and males who used the cards. In the case of the PTSD symptoms, Cronbach's alpha was higher among females and males who used the NVRC method compared to the verbal method, but the difference was only significant for males. Cronbach's alpha was 0.735 for the four PTSD symptoms among males who used the cards compared to 0.410 among males who used the verbal response method. Similar to what we observed for the mean number of trauma types and the mean number of PTSD symptoms, the within method gender difference in Cronbach's alpha for PTSD symptoms was significant among verbal respondents but not among card respondents.

As a further check on reliability, we estimated for each of the two response methods, the correlation between the number of reported trauma types and the number of reported PTSD symptoms. Among respondents who used the verbal method, the number of reported traumas explained only 9% ($r^2=0.089$) of the variance in the number of PTSD symptoms, whereas among NVRC respondents the number of reported traumas explained 30% ($r^2=0.295$) of the variance. We suspect that an important source of the difference in the correlations is due to response bias for the assault questions. The largest difference in Cronbach's alpha between the two response methods was for the assault questions, which other studies of trauma and PTSD have found to be the strongest predictors of PTSD (Ozer et al. 2003; Resnick et al. 1993). With respect to gender differences, we found no significant within response method differences in the trauma and PTSD correlations for female and male respondents.

We also tested for convergent validity for the PTSD questions by response method (Table 4). Convergent validity refers to the extent to which the measures of PTSD are positively correlated with other related measures. Given the strong association between PTSD and self-destructive behavior and depression found in other studies (American Psychiatric Association 2022; Haldane and Nickerson 2016; Steel et al. 2011), we estimated the correlations between the four measures of PTSD and reports of ever self-harm and felt life not worth living. All eight of the bivariate correlations among NVRC respondents were larger than the corresponding correlations among verbal respondents, but the differences were only statistically significant for two comparisons. The correlations between reports of feeling numb or detached and ever engaged in self-harm, and reports of being constantly on guard and felt life not worth living were significantly larger among NVRC respondents than verbal respondents.

3.3 Error Rates

Table 5 presents the question-specific error rates for the 26 "yes/no" questions in the ARISE survey with which the NVRC method was used. The error rates are expressed as the percentage of questions that had an invalid or missing value in the data set. Sources of error for the NVRC method included a respondent provided a numeric response to a "yes/no" question, the interviewer entered a non-valid 3-digit number into the tablet, and a missing response (nonresponse). The only measurable source of error for the verbal respondents was a missing response because the survey tablet included range checks. At the start of the sensitive question section of the questionnaire, the NVRC respondents were asked if they listened to the radio in the last 7 days to provide a measure of whether the respondent understood how to use the card. Two patterns are immediately apparent in Table 5. First, the error rate for the NVRC method declined across questions. Second, by

the fifth “yes/no” question the error rates for the NVRC method were low and comparable to the error rates for the verbal response method. The NVRC error rate was significantly higher than the verbal error rate for 7 of the 26 “yes/no” questions, and with exception of the second “yes/no” question, the NVRC error rates for these questions were between 0.5 and 1.7 percent.

Table 6 presents the interviewer-specific error rates by response method. The error rates were calculated as the mean percentage of missing or invalid responses to the 26 “yes/no” questions that were asked of both NVRC and verbal respondents. Thirteen of the 15 interviewers had verbal error rates well below 1%, and 11 interviewers had NVRC error rates below 1%. The mean NVRC error rate was over twice that for the verbal response method (0.8% compared to 0.3%) but was small. Only one interviewer had a mean NVRC error rate significantly larger than the mean verbal error rate. The intra-interviewer correlation between the NVRC and verbal error rates was close to zero, indicating that interviewers who had comparatively higher error rates with one response method did not also have higher error rates with the other response method.

4. Discussion

Due to the shame and embarrassment that can be attached to some traumatic experiences and PTSD symptoms, both trauma and PTSD are subject to underreporting in face-to-face interviews. Scholars and clinicians of trauma and PTSD are well aware of the potential for both underreporting, as well as overreporting in contexts where respondents believe that trauma and PTSD will provide access to resources and protected status. The magnitude and direction of bias in respondent reports of trauma and PTSD are not only context specific, but can also vary by demographic characteristics such as gender and age depending upon cultural norms. For example, studies tend to find that males are less likely to report PTSD symptoms than females, but it not clear whether cultural norms regarding male resiliency are protective of PTSD or make males less likely to feel comfortable reporting PTSD. The nonverbal response card (NVRC) was developed to reduce the influence of social desirability bias on respondent reports of potentially sensitive topics. In addition to hiding the actual response from the interviewer, the NVRC eliminates the risk that other people hear the response. Face-to-face interviewing, considered the gold standard for survey questionnaires, creates a dynamic in which young respondents in particular are afraid or uncomfortable to report nonmarital sex, sexual coercion, and forced sex (Harling et al. 2020; Lindstrom et al., 2010; Lindstrom et al., 2012). Results from this study of youth in Burkina Faso add trauma and PTSD to the growing body of evidence of the presence of response bias for questions regarding sensitive and stigmatized behaviors, attitudes and experiences.

Respondents who used the NVRC reported an average of 18% more types of life time trauma than verbal respondents (a mean of 1.77 compared to 1.50), and they were 2.8 times as likely to report 3-4 PTSD symptoms (8.4% compared to 3.0%). In addition, the significant gender differences observed in the reports of trauma and PTSD among verbal respondents, were smaller and not statistically significant among the youth who used the NVRC. Female reports of trauma and male reports of PTSD were most affected by response method. In particular, close to 7% of males who used the NVRC reported 3-4 PTSD symptoms compared to less than 1% who gave verbal responses.

Youth who used the NVRC rather than the verbal response method were not only more likely to report trauma and PTSD, the internal consistency of their reports across three

domains was also higher as well as the extent to which their total number of reported trauma types correlated with the number of PTSD symptoms. Among the NVRC respondents, Cronbach's alpha, a measure of internal consistency, was 0.844 for the 15 trauma questions and 0.708 for the 4 PTSD questions. By contrast, among verbal respondents Cronbach's alpha was 0.599 for the trauma questions and 0.518 for the PTSD questions. The number of reported traumas also explained a substantially higher proportion of the variance in the number of reported PTSD symptoms among the NVRC respondents compared to the verbal respondents (29.5% compared to 8.9%).

In addition to checking for internal consistency reliability, we also checked for convergent validity. We found that the bivariate correlations between reports of the 4 PTSD symptoms with reports of self-harm and feeling life not worth living were higher among NVRC respondents than verbal respondents. However, the magnitudes of the differences were statistically significant in only two of the comparisons. Similar to PTSD, reports of ever self-harm and felt life not worth living were significantly higher among NVRC respondents compared to verbal respondents which resulted in within method bivariate correlations that were not significantly different. The correlations between numb and detached and ever self-harm, and constantly on guard and felt life was not worth living were both significantly greater among NVRC respondents than verbal respondents. Among the four PTSD symptoms measured by the survey, the differences by response method in reports of being constantly on guard and feeling numb and detached were the greatest.

Interviewer and respondent use of the NVRC method placed only a minor additional burden compared to the verbal response method. Although the question-specific nonresponse/error rate for the NVRC was significantly greater on seven of the 25 "yes/no" questions in the survey for which both response methods were used, the error rates were nevertheless comparatively small and well below 1% on most questions. We found large variability across interviewers in their error rates, with 3 of the 15 interviewers having NVRC error rates over twice as large as the verbal error rate. With careful monitoring of interviewers in the field, it is quite feasible that the error rate for the NVRC can be brought down to levels not significant different from the verbal error rate.

5. Strengths and Limitations

The NVRC method provides an inexpensive, portable and easy to understand and use response method that reduces social desirability bias in the context of interviewer administered survey questionnaires by hiding the actual response from the interviewer. It does not require literacy or familiarity with a computer keyboard or touch screen, which is particularly advantageous in low-income settings. Error rates for "yes/no" questions using the cards are slightly higher than error rates using the verbal response method, but the differences in response patterns between the two methods far outweigh the higher error rates. The presumption with the NVRC is that when given a more private response method, respondents will provide more accurate answers to sensitive questions. Without external sources of information to validate the responses, we cannot prove our presumption. However, it is implausible that youth who used the NVRC provided misleading responses in a way that generated higher levels of internal consistency reliability and convergent validity compared to verbal respondents.

Harling et al. (2020) in an analysis of the sexual experience questions from the ARISE survey found no significant difference by response method in reports of ever had sexual intercourse, someone joked about wanting to have sex and someone touched their

genitals/breasts. However, they did find that reports of someone tried to force sex with them but failed or someone forced sex were significantly higher among NVRC respondents compared to verbal respondents. The presence of significant differences on the most sensitive items but not all items provides further evidence that respondents using the NVRC were revealing actual experiences. Results from the ARISE survey, add Burkina Faso to the list of countries (Ethiopia and Tanzania) where the efficacy of the NVRC in reducing social desirability bias in samples of youth and young adults has been demonstrated. Further experimental trials are needed to test the value of the method in other age groups and populations, especially given the importance of cultural in determining what types of experiences and behaviors are considered shameful or stigmatized.

6. Public Health Implications

Field tests of the NVRC method in three different African countries (Burkina Faso, Tanzania, and Ethiopia) provide evidence of significant underreporting of many stigmatized behaviors and experiences. The extent of underreporting likely varies by culture and gender depending upon norms and expectations. For instance, nonmarital sexual experience in many contexts is frowned upon for women but not for men. In contrast, rape and sexual victimization is stigmatized for both men and women. In this study, we found that some youth were reluctant to report traumatic experiences and PTSD symptoms when not provided a more private response method. The NVRC method reduces social desirability bias and is useful in surveys designed to generate estimates of the prevalence of trauma and PTSD in a population, as well as to identify risk factors associated with both. The NVRC method does not provide a solution to social desirability bias in the context of clinical evaluations, where the clinician needs to know the nature and extent of an individual's trauma and psychological state. However, in situations where it is important to have accurate estimates of the prevalence of trauma and PTSD or to provide an initial screening of subjects for follow-up evaluation, the NVRC is a valuable tool.

Acknowledgements

We are grateful to the Population Studies and Training Center at Brown University, which receives funding from the NIH (P2C HD041020), for general support.

References

- Aichele SR, Mulder MB, James S, Grimm K. Attitudinal and Behavior Characteristics Predict High Risk Sexual Activity in Rural Tanzanian Youth. *PLoS ONE*. 2014;9(6), e99987. Doi:10.1371/journal.pone.0099987
- Amaya-Jackson L, Socolar RRS, Hunter W, Runyan DK, Colindres R. Directly Questioning Children and Adolescents about Maltreatment: A Review of Survey Measures Used. *Journal of Interpersonal Violence*. 2000;15(7):725-759.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision*. Washington, D.C.: American Psychiatric Association; 2022.
- Bae H, Kim D, Koh H, Kim S, Park JS. Psychometric Properties of the Life Events Checklist-Korean Version. *Psychiatry Investigation*. 2008;5:163-167.
- Beck JG, McNiff J, Clapp JD, Olsen SA, Avery ML, Hagedorn JH. Exploring Negative Emotion in Women Experiencing Intimate Partner Violence: Shame, Guilt, and PTSD. *Behavioral Therapy*. 2011;42:740-750.

- Beck JG, Grant DM, Clapp JD, Palyo SA. Understanding the Interpersonal Impact of Trauma: Contributions of PTSD and Depression. *Journal of Anxiety Disorders*. 2009;23:443-450.
- Berthold SM, Mollica RF, Silove D, Tay AK, Lavelle J, Lindert J. The HTQ-5: Revision of the Harvard Trauma Questionnaire for Measuring Torture, Trauma and DSM-5 PTSD Symptoms in Refugee Populations. *The European Journal of Public Health*. 2018;29(3):468-474.
- Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*. 2015;28(6):489-498.
- Bremner JD, Vermetten E, Mazure CM. Development and Preliminary Psychometric Properties of an Instrument for the Measurement of Childhood Trauma: The Early Trauma Inventory. *Depression and Anxiety*. 2000;12:1-12.
- Breslau N, Kessler R, Chilcoat H, Schultz L, Davis G, Andreski P. Trauma and Posttraumatic Stress Disorder in the Community: The 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry*. 1998;55:626-632.
- Cameron RP, Gusman D. The primary care PTSD screen (PC-PTSD): development and operating characteristics. *Primary Care Psychiatry*. 2003;9(1):9-14.
- Darling AM, Assefa N, Bärnighausen T, Berhane Y, Canavan CR, Guwatudde D, Killewo J, Oduola A, Sando MM, Sie A, Sudfeld C, Vuai S, Adanu R, Fawzi WW. Design and Field Methods of the ARISE Network Adolescent Health Study. *Tropical Medicine and International Health*. 2020;25(1):5-14.
- Ditlevsen DN, Elklit A. Gender, Trauma Type, and PTSD Prevalence: A Re-analysis of 18 Nordic Convenience Samples. *Annals of General Psychiatry*. 2012;11(26): <http://www.annals-general-psychiatry.com/content/11/1/26>
- Feldt LS, Woodruff DJ, Salih FA. Statistical Inference for Coefficient Alpha. *Applied Psychological Measurement*. 1987;11:93-103. doi:10.1177/014662168701100107
- Gray MJ, Litz BT, Hsu JL, Lombardo TW. Psychometric Properties of the Life Events Checklist. *Assessment*. 2004;11(4):330-341.
- Haldane J & Nickerson A. The Impact of Interpersonal and Noninterpersonal Trauma on Psychological Symptoms in Refugees: The Moderating Role of Gender and Trauma Type. *Journal of Traumatic Stress*. 2016;29:457-465.
- Hecker T, Fetz S, Ainamani H, Elbert T. The Cycle of Violence: Associations between Exposure to Violence, Trauma-related Symptoms and Aggression – Findings from Congolese Refugees in Uganda. *Journal of Traumatic Stress*. 2015;28:448-455.
- Hollifield M, Warner TD, Lian N, Krakow B, Jenkins JH, Kesler J, Stevenson J, Westermeyer J. Measuring Trauma and Health Status in Refugees. *JAMA*. 2002;288(5):611-621.
- Jones EF, Forrest JD. Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988. *Demography*. 1992;29(1):113-126.
- Karatzias T, Cloitre M, Maercker A, Kazlauskas E, Shevlin M, Hyland P, Bisson JI, Roberts NP, Brewin CR. PTSD and Complex PTSD: ICD-11 Updates on Concept and Measurement in the UK, USA, Germany and Lithuania. *European Journal of Psychotraumatology*. 2017;8:sup7, 1418103, DOI:10.1080/20008198.2017.1418103
- Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic Stress Disorder in the National Comorbidity Survey. *Archives of General Psychiatry*. 1995;52:1048-1060.
- Kienzler H. Debating War-trauma and Post-traumatic Stress Disorder (PTSD) in an Interdisciplinary Arena. *Social Science & Medicine*. 2008;67:218-227.

- Kilpatrick DG, Resnick HS, Milanak ME, Miller MW, Keyes KM, Friedman MJ. National Estimates of Exposure to Traumatic Events and PTSD Prevalence Using DSM-IV and DSM-5 Criteria. *Journal of Traumatic Stress*. 2013;26:537-547.
- Kolassa I-T, Ertl V, Eckart C, Kolassa S, Onyut LP, Elbert T. Spontaneous Remission from PTSD Depends on the Number of Traumatic Event Types Experienced. *Psychological Trauma: Theory, Research, Practice, and Policy*. 2010;2:169-174.
- Krysan M. Privacy and the Expression of White Racial Attitudes: A Comparison Across Three Contexts. *The Public Opinion Quarterly*. 1998;62(4):506-544.
- Kubany ES, Leisen MB, Kaplan AS, Kelly MP. Validation of a Brief Measure of Posttraumatic Stress Disorder: The Distressing Event Questionnaire (DEQ). *Psychological Assessment*. 2000;12(2):197-209.
- Leskela J, Dieperink M, Thuras P. Shame and Posttraumatic Stress Disorder. *Journal of Traumatic Stress*. 2002;15:223-226.
- Lindstrom DP, Belachew T, Hadley C, Klein-Hattori M, Hogan D, Tessema F. Nonmarital Sex and Condom Knowledge among Ethiopian Young People: Improved Estimates Using a Nonverbal Response Card. *Studies in Family Planning*. 2020;41(4):251-262.
- Lindstrom DP, Klein-Hattori M, Belachew T, Tessema F. Lifting the Curtain on the Conditions of Sexual Initiation among Youth in Ethiopia. *Journal of Adolescent Health*. 2012;50:614-620.
- Macdonald J, Morley I. Shame and Non-disclosure: A Study of the Emotional Isolation of People Referred for Psychotherapy. *British Journal of Medical Psychology*. 2001;74:1-21.
- McDonald MK, Borntrager CF, Rostad W. Measuring Trauma: Considerations for Assessing Complex and Non-PTSD Criterion A Childhood Trauma. *Journal of Trauma & Dissociation*. 2014;15:184-203.
- Mensch BS, Hewett PC, Gregory R, Helleringer S. Sexual Behavior and STI/HIV Status among Adolescents in Rural Malawi: An Evaluation of the Effect of Interview Mode on Reporting. *Studies in Family Planning*. 2008;39:321-334.
- Nader K. *Understanding and Assessing Trauma in Children and Adolescents: Measures, Methods, and Youth in Context*. New York: Routledge; 2008.
- Netland M. Assessment of Exposure to Political Violence and Other Potentially Traumatizing Events. *Journal of Traumatic Stress*. 2001;14:311-326.
- Norris FH, Murphy AD, Baker CK, & Perilla JL, Gutiérrez Rodríguez F, Gutiérrez Rodríguez JJ. Epidemiology of Trauma and Posttraumatic Stress Disorder in Mexico. *Journal of Abnormal Psychology*. 2003;112(4):646-656.
- Øktedalen T, Hagtvet KA, Hoffart A, Langkaas TF, Smucker M. The Trauma Related Shame Inventory: Measuring Trauma-Related Shame among Patients with PTSD. *Journal of Psychopathology and Behavior Assessment*. 2014;36:600-615.
- Olatunji BO, Cisler JM, Tolin DF. Quality of Life in the Anxiety Disorders: A Meta-analytic Review. *Clinical Psychology Review*. 2007;27:572-581.
- Ozer EJ, Best SR, Lipsey TL, Weiss DW. Predictors of Posttraumatic Stress Disorder and Symptoms in Adults: A Meta-analysis. *Psychological Bulletin*. 2003;129:52-73.
- Rasmussen A, Romero S, Leon M, Verkuilen J, Morales P, Martinez-Maganalles S, García-Sosa I. Measuring Trauma Exposure: Count Versus Variety of Potentially Traumatic Events in a Binational Sample. *Journal of Traumatic Stress*. 2020;233:973-983.
- Rasmussen A, Rosenfield B, Reeves K, Keller AS. The Subjective Experience of Trauma and Subsequent PTSD in a Sample of Undocumented Immigrants. *The Journal of Nervous and Mental Disease*. 2007;195(2):137-143.

- Resnick HS, Kilpatrick DG, Dansky BS, Saunders BE, Best CL. Prevalence of Civilian Trauma and Posttraumatic Stress Disorder in a Representative National Sample of Women. *Journal of Consulting and Clinical Psychology*. 1993;61:984-991.
- Saxe G, Wolfe J. Gender and Posttraumatic Stress Disorder. Pp. 160-179 in P Saigh & J Bremner (eds), *Posttraumatic Stress Disorder: A Comprehensive Text*. Boston: Allyn & Bacon; 1999.
- Sié A, Louis VR, Gbangou A, Müller O, Niamba L, Stieglbauer G, Yé M, Kouyaté B, Sauerborn R, Becher H. The Health and Demographic Surveillance System (HDSS) in Nouna, Burkina Faso, 1993–2007. *Global Health Action*. 2010;3(1):5284. DOI: 10.3402/gha.v3i0.5282
- Solomon S, Davidson J. Trauma: Prevalence, Impairment, Service Use and Cost. *Journal of Clinical Psychiatry*. 1997;58:5-11.
- Steel JL, Dunlavy AC, Stillman J, Pape HC. Measuring Depression and PTSD after Trauma: Common Scales and Checklists. *Injury, International Journal of the Care of the Injured*. 2011;42:288-300.
- Tjaden P, Thoennes N. Prevalence, Incidence, and Consequences of Violence Against Women. *Research in Brief* (NCJ172837). Washington, DC: U.S. Department of Justice, National Institute of Justice; 1998.
- Tolin DF, Foa EB. Sex Differences in Trauma and Posttraumatic Stress Disorder: A Quantitative Review of 25 Years of Research. *Psychological Bulletin*. 2006;132:959-992.
- Tourangeau R, Rips LJ, Rasinski K. *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press; 2000.
- Weathers F, Blake D, Schnurr P, Kaloupek D, Marx B, Keane T. The life events checklist for DSM-5 (LEC-5); 2013 from https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp
- Weaver B, Wuensch KL. SPSS and SAS Programs for Comparing Pearson Correlations and OLS Regression Coefficients. *Behavioral Research*. 2013;45:880-895.
- Weathers FW, Keane TM. The Criterion A Problem Revisited: Controversies and Challenges in Defining and Measuring Psychological Trauma. *Journal of Traumatic Stress*. 2007;20(2):107-121.
- Winders SJ, Murphy O, Looney K, O'Reilly G. Self-compassion, Trauma, and Posttraumatic Stress Disorder: A Systematic Review. *Clinical Psychology and Psychotherapy*. 2020;27:300-329.
- Wolford G, Rosenberg SD, Rosenberg HJ, Swartz MS, Butterfield MI, Swanson JW, Jankowski MK. A Clinical Trial Comparing Interviewer with Computer-Assisted Assessment among Clients with Severe Mental Illness. *Psychiatric Services*. 2008;59:769-775.

Figure 1. Nonverbal Response Card (NVRC), ARISE 2017.

A. Side of card facing respondent

20 • 	0 •	1 • 	2 • 	3 • 	4 • 		Yes • Oui	No • Non
5 • 	6 • 	7 • 	8 • 	9 • 	10 • 		No • Non	Yes • Oui
11 • 	12 • 	13 • 	14 • 	15 • 	16 • 		No • Non	Yes • Oui
17 • 	18 • 	19 • 	20 • 	0 •	1 • 		Yes • Oui	No • Non
2 • 	3 • 	4 • 	5 • 	6 • 	7 • 		Yes • Oui	No • Non

B. Side of card facing interviewer

101 •	529 •		897 •	295 •	181 •	125 •	345 •	979 •
229 •	379 •		488 •	285 •	718 •	139 •	275 •	448 •
504 •	346 •		486 •	649 •	340 •	904 •	682 •	170 •
636 •	765 •		470 •	166 •	435 •	491 •	723 •	506 •
828 •	218 •		735 •	813 •	628 •	966 •	222 •	422 •

Table 1. Selected Respondent Characteristics by Response Method, ARISE 2017.

Variable	% (No.)		<i>p</i> ^a <
	Verbal	NVRC	
Gender			0.026
Female	44.9 (386)	39.5 (310)	
Male	55.1 (473)	60.5 (475)	
Age			0.925
12-14	45.3 (389)	44.8 (352)	
15-17	33.9 (291)	34.8 (273)	
18-20	20.8 (179)	20.4 (160)	
Student status			0.084
Not in school	48.0 (412)	52.2 (410)	
In school	52.0 (447)	47.8 (375)	
Marital status			0.802
Never married	90.3 (776)	90.7 (712)	
Ever married	9.7 (83)	9.3 (73)	
Religion			0.265
Muslim	67.5 (580)	70.0 (557)	
Catholic	21.9 (188)	19.4 (152)	
Protestant	7.7 (66)	6.1 (48)	
Animist, other	2.9 (25)	3.6 (28)	

Note. ^aFrom Pearson χ^2 test.

Table 2. Affirmative Reports of Life Time Trauma and PTSD in the Last 30 Days by Response Method, ARISE 2017.

	Yes, % (No.) or Mean (SD)		$p^a <$
	Verbal	NVRC	
Trauma types			
Natural disaster	34.8 (857)	32.6 (777)	0.345
Fire or explosion	11.6 (859)	16.9 (781)	0.002
Vehicle accident	14.3 (858)	16.9 (783)	0.159
Chemical exposure	1.6 (858)	5.9 (782)	0.001
Physical assault	34.5 (858)	27.0 (782)	0.001
Assault with weapon	4.8 (858)	7.3 (781)	0.032
Sexual assault	2.6 (855)	4.9 (783)	0.014
Other unwanted sexual experience	1.9 (856)	5.5 (783)	0.001
Combat/war zone	0.5 (857)	4.5 (781)	0.001
Captivity	1.1 (856)	4.6 (784)	0.001
Life-threatening injury/illness	3.7 (854)	10.2 (784)	0.001
Witnessed sudden intentional death	8.4 (859)	10.0 (780)	0.257
Witnessed sudden unintentional death	13.4 (859)	12.8 (782)	0.174
Witnessed harm caused to someone else	4.7 (856)	7.2 (780)	0.031
Other stressful event	11.3 (856)	12.5 (783)	0.460
Mean number of trauma types	1.50 (1.64)	1.77 (2.49)	0.010
Mean number of trauma types ^b : Females	1.36 (1.51)	1.86 (2.51)	0.001
Mean number of trauma types ^b : Males	1.61 (1.73)	1.71 (2.48)	0.489
PTSD symptoms			
Nightmares	21.6 (858)	23.5 (773)	0.338
Avoid being reminded	13.3 (856)	20.0 (775)	0.001
Constantly on guard	7.7 (858)	13.2 (772)	0.001
Numb, detached	4.4 (850)	8.8 (775)	0.001
3-4 symptoms	3.0 (846)	8.4 (765)	0.001
3-4 symptoms ^c : Females	5.6 (376)	10.6 (301)	0.015
3-4 symptoms ^c : Males	0.9 (470)	6.9 (464)	0.001

Note. ^aFrom Pearson χ^2 test for difference of proportions and F test for difference of means. ^b $p = 0.026$ from F test for difference of verbal female and male means, $p = 0.411$ from F test for difference of NVRC female and male means. ^c $p = 0.001$ from Pearson χ^2 test for difference of verbal female and male proportions, $p = 0.068$ from Pearson χ^2 test for difference of NVRC female and male proportions.

Table 3. Reliability for Trauma and PTSD Questions by Response Method, ARISE 2017.

	Verbal (No.)	NVRC (No.)	$p^a <$
Cronbach's alpha			
Disaster/accident (4 questions)	0.488 (835)	0.629 (756)	0.001
Assault (4 questions)	0.285 (835)	0.655 (756)	0.001
War/violence (6 questions)	0.411 (835)	0.753 (756)	0.001
Trauma (15 questions)	0.599 (835)	0.844 (756)	0.001
Trauma: Females ^b	0.563 (371)	0.845 (295)	0.001
Trauma: Males ^b	0.621 (464)	0.844 (461)	0.001
PTSD (4 questions)	0.518 (846)	0.708 (765)	0.001
PTSD: Females ^c	0.560 (376)	0.666 (301)	0.053
PTSD: Males ^c	0.410 (470)	0.735 (464)	0.001
Correlation coefficient (r)			
Trauma, PTSD	0.298 (824)	0.543 (738)	0.001
Trauma, PTSD: Females ^d	0.324 (363)	0.513 (287)	0.004
Trauma, PTSD: Males ^d	0.335 (461)	0.567 (451)	0.001

Note. ^aFrom χ^2 test for difference of alpha coefficients (Feldt et al. 1987), and from Fischer's r -to- z transformation to test for difference of correlation coefficients (Weaver and Wuensch 2013). ^b $p = 0.176$ from χ^2 test for difference of verbal female and male alpha coefficients, $p = 0.956$ from χ^2 test for difference of NVRC female and male alpha coefficients. ^c $p = 0.021$ from χ^2 test for difference of verbal female and male alpha coefficients, $p = 0.085$ from χ^2 test for difference of NVRC female and male alpha coefficients. ^d $p = 0.433$ from z test for difference of verbal female and male correlation coefficients, $p = 0.159$ from z test for difference of NVRC female and male correlation coefficients.

Table 4. Bivariate Correlations for Ever Self-harm and Felt Life Not Worth Living with PTSD Conditions by Response Method, ARISE 2017.

	Verbal (No.)	NVRC (No.)	$p^a <$
Ever self-harm with:			
Nightmares	0.178 (858)	0.203 (751)	0.603
Avoid being reminded	0.137 (856)	0.156 (753)	0.697
Constantly on guard	0.175 (858)	0.179 (750)	0.936
Numb, detached	0.037 (850)	0.168 (753)	0.008
Felt life not worth living with:			
Nightmares	0.045 (856)	0.097 (749)	0.298
Avoid being reminded	0.075 (854)	0.114 (751)	0.430
Constantly on guard	0.078 (856)	0.199 (748)	0.014
Numb, detached	0.089 (848)	0.156 (752)	0.174

Note. ^aFrom Fischer's *r*-to-*z* transformation to test for difference of correlation coefficients (Weaver and Wuensch 2013).

Table 5. Question-specific Nonresponse/Error Rate by Response Method, ARISE 2017.

Question order	Yes/no questions	Nonresponse/error %		<i>p</i> ^a <
		Verbal (n = 859)	NVRC (n = 785)	
2	Listened to radio in last 7 days		5.4	
6	Ever self-harm	0.0	3.4	0.001
17	Natural disaster	0.2	1.0	0.041
18	Fire or explosion	0.0	0.5	0.036
19	Vehicle accident	0.1	0.3	0.511
20	Chemical exposure	0.1	0.4	0.275
21	Physical assault	0.1	0.4	0.275
22	Assault with weapon	0.1	0.5	0.148
23	Sexual assault	0.5	0.3	0.479
24	Other unwanted sexual experience	0.3	0.3	0.728
25	Combat/war zone	0.2	0.5	0.353
26	Captivity	0.3	0.1	0.362
27	Life-threatening injury/illness	0.6	0.1	0.127
28	Sudden intentional death	0.0	0.6	0.019
29	Sudden unintentional death	0.0	0.4	0.070
30	Harm caused to someone else	0.3	0.6	0.402
31	Other stressful event	0.3	0.3	0.728
32	Nightmares	0.1	1.5	0.001
33	Avoid being reminded	0.3	1.3	0.034
34	Constantly on guard	0.1	1.7	0.001
35	Numb, detached	1.0	1.3	0.668
36	Ever had sexual intercourse	0.7	0.4	0.385
44	Someone made jokes about sex with you	0.1	0.6	0.080
45	Unwanted touch of genitals or breasts	0.9	0.9	0.933
46	Escaped forced sex	0.3	0.8	0.255
47	Forced sex	0.5	1.0	0.188

Note. ^aFrom Pearson χ^2 test for difference of proportions.

Table 6. Interviewer-specific Nonresponse/Error Rate by Response Method, ARISE 2017.

Interviewer	Nonresponse/error, % ^a (No.) ^b		<i>p</i> ^c <
	Verbal	NVRC	
1	0.1 (56)	0.0 (55)	0.324
2	0.1 (60)	3.5 (74)	0.001
3	0.1 (58)	1.6 (43)	0.206
4	0.3 (57)	0.4 (51)	0.643
5	1.6 (42)	1.1 (47)	0.558
6	1.3 (52)	0.2 (42)	0.170
7	0.0 (41)	0.1 (62)	0.419
8	0.0 (44)	0.3 (48)	0.051
9	0.5 (56)	0.0 (59)	0.014
10	0.1 (63)	0.3 (62)	0.168
11	0.1 (72)	0.0 (41)	0.453
12	0.0 (83)	0.2 (52)	0.073
13	0.4 (52)	1.7 (50)	0.133
14	0.4 (72)	0.8 (47)	0.224
15	0.1 (50)	0.6 (51)	0.068
Total	0.3 (859)	0.8 (785)	0.002
Correlation (Verbal, NVRC) = 0.02 (15)			0.953

Note. ^aNon-response/error rate for 25 questions listed in Table 5 that were asked of both verbal and NVRC respondents. ^bNumber of interviews. ^cFrom *F* test for difference of means, *t* test for null hypothesis that correlation coefficient equals zero.