Title: A New Method for Recording Responses to Questions about Traumatic Sexual Life Events: Evidence of the Use of the Non-Verbal Response Card Method in a Survey of Adolescents in Rural Burkina Faso

Authors: Guy Harling, Mamadou Bountogo, Ali Sié, Till Bärnighausen, David Lindstrom

ABSTRACT

In an effort to limit social desirability bias in face-to-face survey interviews, we conducted an individually-randomized experiment by assigning 50% of respondents to the conventional verbal response method and 50% to the non-verbal response card (NVRC) method. NVRCs provide a low-tech method for respondents to non-verbally respond to sensitive questions without revealing the nature of their response to the interviewer. As part of a larger health questionnaire, we asked 1544 12-20 years olds in rural Burkina Faso about their history of physical violence, sexual debut, sexual unwanted attention/violence, and post-traumatic stress disorder (PTSD) symptoms. NVRC respondents provided logically consistent responses. NVRC respondents reported similar patterns of physical assault and sexual debut as verbal respondents, but significantly higher levels of sexual assault, forced sex and PTSD symptoms. Our findings suggest that NVRC may be practical and beneficial in a low-literacy population for reducing underreporting of stigmatized and traumatic experiences.

INTRODUCTION

Sexual violence, both by intimate partners and others, is both highly prevalent and has substantial effects on the lives of survivors. The World Health Organization estimates that 35% of women have experienced either physical violence from an intimate partner or sexual violence (World Health Organization, 2013). While fewer men are survivors of sexual violence, 23.4% of men (compared to 43.9% of women) in the United States are estimated to have ever experienced sexual violence in the form of being made to penetrate, being sexually coerced, or facing unwanted sexual contact or noncontact sexual experiences (Breiding et al., 2014). Such violence has myriad physical and psychological impacts (Dillon et al., 2013; Jordan et al., 2010). Violence against children can be particularly debilitating insofar as it affects them during a key developmental period, affecting health, learning and future life opportunities (Pinheiro, 2006). A meta-analysis of population-based studies estimated that cumulative experience of sexual abuse alone by age 18 was 18.0% for girls and 7.6% for boys (Stoltenborgh et al., 2011).

The prevalence of, and risk factors for, sexual violence is difficult to validly estimate however, since it is almost always based on retrospective self-reports and thus susceptible to recall bias – both inadvertent and intentional. Intentional underreporting may be exacerbated by the highly sensitive and stigmatizing nature of sexual violence, and the degree to which it is inflicted by family members and other loved individuals (Jejeebhoy et al., 2013). As a result, the true prevalence of sexual violence, and risk factors predicting it, remain somewhat uncertain.

This 'social desirability' bias, i.e. the tendency to under-report stigmatized (and over-report socially normative) behaviors, is well-known in the sexual behavior and survey literature (Langhaug et al., 2010; Phillips et al., 2010). Several methods have been proposed for reducing the risk of intentional mis-reporting in surveys. One such approach is to ask respondents to record their answers either on paper or on a computer screen, i.e. complete a 'self-interview', or to post their answer in a ballot box (Gregson et al., 2004). These methods require literacy and experience and comfort with computers, which is not universal in many low-income settings. Another approach is to hide the respondents' answer to the sensitive question among other answers. This approach includes 'random response' techniques, where respondents answer truthfully some known proportion of the time but the interviewer does not know when (Lensvelt-Mulders et al., 2005), and 'list

randomization' methods, where the respondent is asked how many items from a larger set of statements they agree with (Haber et al., 2018). These hidden response methods, while potentially effective at estimating population prevalence, do not allow post-hoc identification of individual's responses that can be correlated with other attributes.

The non-verbal response card (NVRC) is a method for reducing social desirability bias that can be used in low-literacy settings and allows individual responses to be linked to other attributes. These cards were developed for soliciting responses to questions about sexual knowledge, attitudes and practices among adolescents in (Lindstrom et al., 2010; Lindstrom et al., 2012). The NVRC method allows respondents to indicate their response to interviewers, without the interviewers knowing the nature of the response. The cards are low-tech, lightweight, privacy-preserving and require minimal literacy – and thus appropriate to being carried for fieldwork and used in low-literacy populations. Past trials comparing these cards to conventional verbal responses found significantly higher reported prevalence of sexual coercion, rape, and non-marital sex, and lower reported levels of condom knowledge among respondents who used the card method (Lindstrom, et al., 2010; Lindstrom, et al., 2012).

The NVRC method has not, however, been tested outside Ethiopia. We conducted an individually randomized experiment of using NVRCs to ask adolescents in rural and semi-rural Burkina Faso about several potentially highly stigmatized behaviors in the context of a wider health and behavior questionnaire.

METHODS

Setting and sample

Our sample data comprises baseline responses from a cohort of adolescents aged 12-20 in rural Burkina Faso. The cohort forms part of the Africa Research, Implementation Science and Education (ARISE) network, a collaboration between the Harvard T.H. Chan School of Public Health and nine sub-Saharan African institutions in seven countries (Darling et al., 2019). The Burkinabé cohort was drawn from the Health and Demographic Surveillance System (HDSS) site overseen by the Centre de Recherche en Santé de Nouna (CRSN). The HDSS site (total population ~107,000 in 2015) is centered on the town of Nouna in the Boucle du Mouhoun province in north-western Burkina Faso

and comprises the town of Nouna (population approximately 30,000) and 58 surrounding rural villages (Sié et al., 2010). The site contains a mixture of ethnic and religious groups.

ARISE Burkina Faso used a two-part stratified sampling procedure to ensure representativeness of ethnicity/religion and urbanicity. We first purposively selected two Nouna HDSS villages in which each of one of the five main ethnicities formed a majority. We then drew a random sample of 1795 youth from a 2015 census of residents of these 10 villages who were age-eligible on 1 October 2017, ensuring this sample respected the ethnic make-up of all age-eligible HDSS adolescents. Second, we drew a simple random sample of 749 age-eligible adolescents from one of the seven sectors of Nouna town.

Baseline interviews were conducted in November and December 2017 in the villages or compounds where the adolescents lived, with interviews conducted in either French or a local language if the respondent did not speak French. The study collected information on socio-demographics, behaviors, health practices and health outcomes using tablet computers. Approval for this study were obtained from the Institutional Ethics Committee of the CRSN, village elders, participants (written consent/assent) and parents/guardians (written consent if participant aged <18). Approval for ARISE overall was obtained from the Harvard T.H. Chan School of Public Health Institutional Review Board.

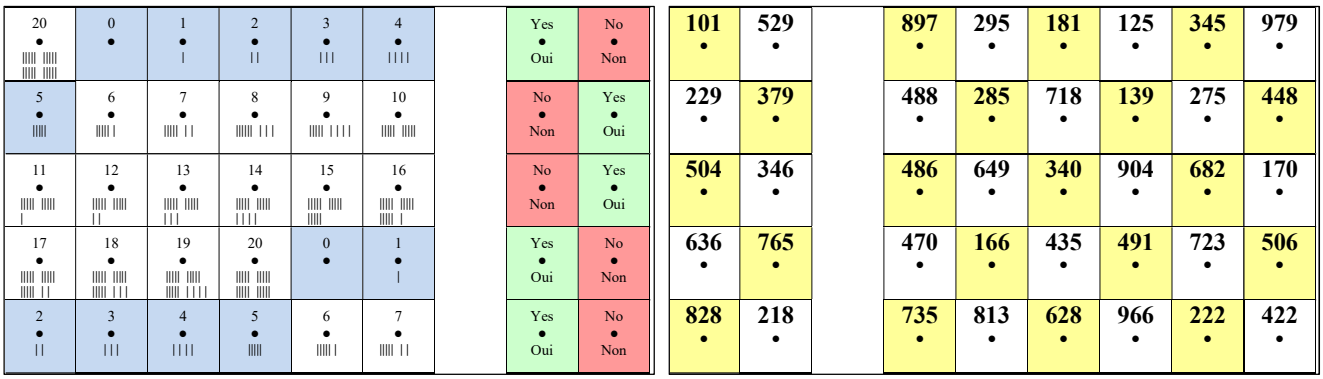
Non-verbal response card

In this baseline survey, all respondents provided verbal responses to non-sensitive questions. For questions regard sexual experiences, violence, and mental health, one-half of the 2544 sampled individuals were randomized to provide verbal responses and one-half to use the NVRC method. Randomization of the two methods was completed prior to the interview.

The NVRC is a two-sided, laminated card; each side of the card is divided into 42 cells with a small hole punched through the center of each cell (Figure 1). Cells on the respondent side of the card contain written and color-coded responses. Non-numeric responses are written in French and are color coded – green for “Yes”, red for “No”. Numeric responses range from 0 to 20 (e.g. for number of sexual partners and age at sexual initiation) and include hash marks for those unable to read Arabic numerals. The cells colored blue were also used for scale responses, e.g. varying from 0 for

'never' to 5 for 'always'. Cells on the interviewer side of the card each contain a unique three-digit number.

Figure 1. Non-verbal response card



A. Side of card facing respondent B. Side of card facing interviewer

The respondent holds the card so only they can see their side of the card, and indicates their response to each question by inserting a stick through the hole in the relevant cell. The interviewer records only the three-digit number they see in the cell through which the stick is protruding. To minimize the risk, or appearance of risk, that the interviewer might recognize a response based on the position of the response cell, multiple cells containing the same value are provided on each card and the respondent is always offered their choice of four cards with differing response positions. Respondents are free to change the card they are using at any point in the interview. The card is divided into two panels, one with 'Yes/No' responses and one with the numeric responses. If the respondent is distracted or not concentrating on the questions, the division of the card into the two panels potentially alerts the interviewer in the event that the respondent is providing a 'Yes/No' response to a numeric response question or a numeric response to a 'Yes/No' question. After data collection is complete, the three-digit numeric codes are recoded to their corresponding response. At the start of the sensitive question section, the interviewer demonstrates to the respondent how the cards work. The interviewer also holds his/her own small demonstration card through the sensitive questions section to remind the respondent how the card works.

Measurement of key variables

The NVRCs were used with a battery of questions regarding sexual experience, violence, trauma and mental health. These included 15 'Yes/No' questions based on the Life Events Checklist in the PTSD Checklist for DSM-5 (Blevins et al., 2015; Weathers et al., 2013), in the format 'in all your life, have you ever experienced…', including: (a) physical assault; (b) sexual assault; and (c) other unwanted or uncomfortable sexual experiences. We also asked the four-question Primary Care PTSD screen, on which responding affirmatively to three or four items is considered indicative of probable PTSD (Cameron & Gusman, 2003). We asked all individuals if they had ever had sexual intercourse, and we asked four questions of increasing sensitivity about sexual assault, 'have any of the following ever happened to you': (1) 'Someone made verbal jokes about wanting to have sex with you'; (2) 'Someone touched you on your genital or breast when you did not want to be touched'; (3) 'Someone forced you to have sex against your will but you escaped'; 'Someone forced you to have sex against your will'.

As part of general field activities, the study team reviewed data every week for the first three weeks of data collection, and provided interviewer-specific feedback where interviewers were generating responses that were marked different from those of their peers, or in the case of the NVRCs generating impossible values (e.g. an answer of '7' to the question 'have you ever had sex?').

To assess respondents' general sensitivity to social desirability bias, we asked all respondents to complete the Balanced Inventory of Desirable Responding (BIDR) (Paulhus, 1984), as translated into French by Sabourin and colleagues (1989) and shortened from 40 to 16 items by Hart and colleagues (2015).

Statistical analysis

After describing the data for 12 questions on the key NVRC section variables outlined above, we evaluated the likely validity of the NVRC survey responses. Interviewer error in the form of data entry errors (limited in this case by automated data entry checks on the tablet computers), and respondent non-compliance through declining to provide a valid response, can occur with both the NVRC and verbal response methods. The improper use of the card in the form of providing the wrong type of response, e.g. a numeric answer to the question 'have you ever had sex', is unique to

the card. We measured non-response and error rates by arm, testing for significant differences using χ^2 tests. We also evaluated how these rates changed over the nine-week survey period. As a an additional check on the relative error rates of the two response methods we looked at the number of married respondents who reported never having had sex, which we expect to be zero.

We tested the internal reliability of the card method by estimating Cronbach's alpha for each of three groups of questions (lifetime traumatic events, the PC-PTSD screen, and sexual abuse) stratified by response method. We also estimated Cronbach's alpha for 16 questions that measure general personality traits that capture susceptibility to social desirability bias and to which all respondents provided verbal responses; our expectation was that there should be no difference in Cronbach's alpha between the two response groups for this latter set of questions. For all alpha values we calculated 95% confidence intervals based on 1000 bootstrap replications.

After evaluating validity, we compared the level of affirmative responses across arms for each of the 12 questions in the NVRC section, testing for significant differences using χ^2 tests. We then conducted multivariable hierarchical regression analysis (respondents nested within interviewers) for the outcome of forced sex. We first evaluated whether any associations between NVRC use and affirmative responses were due to confounding by age, gender, marital status, potential PTSD status (for non-PTSD questions) or social desirability bias. We did not expect to see confounding here since NVRC arm was randomly assigned. We then considered whether the effect of the card was modified by the same set of variables, evaluating effect-modification by comparing the Akaike Information Criteria (AIC) for models without and with interaction terms for the relevant covariates.

RESULTS

Of the 2544 sampled adolescents, 1644 were found and consented to participate. The great majority of those who did not participate had either moved out of the HDSS area (30.1%) or were not available for interview due to travel or work responsibilities (58.5%). In only 39 cases (4.4% of non-participants) did either the sampled individual or their parent decline consent. Among the 1644 respondents, 785 (47.7%) were randomly assigned to the NVRC arm and 859 (52.3%) to the

verbal response method. No respondents refused to use the cards. Descriptive statistics for respondents are provided in Table 1.

Table 1. Descriptive statistics of ARISE Burkina Faso respondents

	All respondents	Verbal response	NVRC
N	1644	859	785
Female			
12-15 years	25.0%	27.1%	22.8%
16-17 years	9.8%	10.7%	8.8%
18+ years	7.5%	7.1%	7.9%
Male			
12-15 years	32.4%	29.3%	35.7%
16-17 years	13.8%	13.9%	13.8%
18+ years	11.5%	11.9%	11.1%
Education			
Currently in school	50.0%	52.0%	47.8%
Marital status			
Engaged to be married	2.7%	2.8%	2.6%
Ever married	6.8%	6.9%	6.8%
Religion			
Muslim	69.2%	67.5%	70.1%
Catholic	20.7%	21.9%	19.4%
Protestant	6.9%	7.7%	6.1%
Animist	3.2%	2.9%	3.6%
Ethnicity			
Bwaba	19.9%	21.5%	18.1%
Dafin	38.1%	36.1%	40.3%
Mossi	17.6%	17.4%	17.8%
Peulh	10.1%	9.4%	10.8%
Samo	12.5%	13.4%	11.6%
Other	1.8%	2.2%	1.4%

Table 2 presents the error rates for selected questions for both the NVRC and verbal response methods. Error rates for the card method were very low, ranging from 0.3 to 1.7 percent across questions. Although NVRC non-response rates were generally higher than those for verbal response, only two differences were statistically significant, and in both cases a single interviewer garnered 50% of the NVRC errors. The error rate for the NVRC method also declined over the course of the survey from a rate of 2.3% during the first week of interviewing, to a rate of 0.2% after the fourth week of interviewing (results not shown in the table). We additionally found only four married individuals who reported never having had sex: four among card respondents and two among verbal respondents; the difference was not statistically significant.

Table 2. Comparison of non-response and error rates in NVRC survey section by randomization arm

Arm	Verbal	NVRC	χ^2	p-value
N				
Lifetime traumatic events				
Physical Assault	0.1%	0.4%	1.2	0.275
Sexual Assault	0.5%	0.3%	0.5	0.479
Other unwanted sexual experience	0.3%	0.3%	0.1	0.728
PC-PTSD questions				
Nightmares	0.1%	1.5%	10.4	0.001
Avoid being reminded	0.3%	1.3%	4.5	0.034
Constantly on guard	0.1%	1.7%	11.5	0.001
Numb, detached	1.1%	1.3%	0.2	0.668
Ever had sexual intercourse	0.7%	0.4%	0.8	0.385
Someone joked about wanting to have sex	0.1%	0.6%	3.1	0.080
Someone touched genitals/breasts	0.9%	0.9%	<0.1	0.933
Someone tried to force sex but failed	0.3%	0.8%	1.3	0.255
Someone forced sex	0.5%	1.0%	1.7	0.188

Table 3 shows that internal reliability for the social desirability scale was the same across response groups, but that on the three sets of sensitive questions, Cronbach's alpha is higher among the card respondents than among the verbal respondents. The differences in Cronbach's alpha are especially notable for the traumatic events and mental health questions, where the 95% confidence intervals for the two groups do not cross. The differences in internal reliability suggest that the verbal respondents may be selectively underreporting some stigmatized experiences with violence and trauma that tend to cluster with other, less stigmatized, experiences.

Table 3. Cronbach's alpha for series of questions in the NVRC survey section

Arm	Verbal		NVRC	
N	859		785	
Social desirability scale (k=16) †	0.57	[0.52 - 0.61]	0.56	[0.52 - 0.61]
Traumatic events (k=15)	0.60	[0.54 - 0.65]	0.82	[0.77 - 0.85]
PC-PTSD screen (k=4)	0.55	[0.50 - 0.61]	0.71	[0.66 - 0.75]
Sexual abuse questions (k=4)	0.65	[0.60 - 0.71]	0.69	[0.64 - 0.75]

Values are Cronbach's alpha and [95% confidence intervals] from 1000 bootstraps.
† Questions in this scale were not asked as part of the NVRC section.

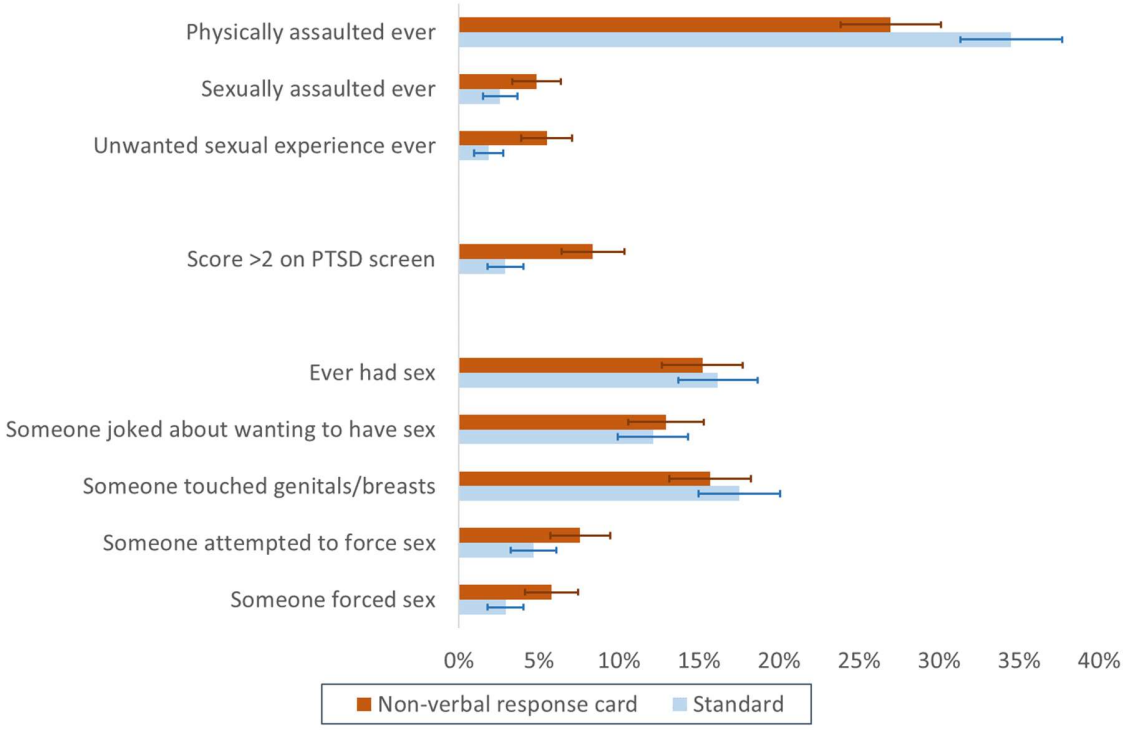
Table 4 presents the percentages of affirmative responses to the questions on traumatic events, PC-PTSD and sexual experience (Figure 2 presents the same with confidence intervals). Reports of

sexual assault and other unwanted sexual experience are significantly higher among respondents who used the card method compared to respondents who gave verbal responses. On the other hand, verbal respondents reported significantly higher levels of physical assault than respondents who used the cards. Respondents who used the card method also reported significantly higher levels of three of the four symptoms of PC-PTSD compared to verbal respondents, and were significantly more likely to screen positive for PTSD. For sexual experience questions, both card and verbal respondents reported similar levels of ever had sexual intercourse, joking about wanting to have sex, and sexual touching. However, respondents using the card method reported significantly higher levels of attempted forced sex and forced sex, both of which are highly stigmatized.

Table 4. Comparison of affirmative responses in NVRC survey section by randomization arm

	Affirmative responses		χ^2	**p-value**
Arm	**Verbal**	**NVRC**		
N	859	785		
Lifetime traumatic events				
Physical Assault	34.5%	27.0%	10.8	0.001
Sexual Assault	2.6%	4.9%	6.0	0.014
Other unwanted sexual experience	1.9%	5.5%	15.5	<0.001
Total score from 15 traumatic events	1 [0 - 2]	1 [0 - 2]	0.7	0.402
PC-PTSD questions				
Nightmares	21.6%	23.5%	0.9	0.338
Avoid being reminded	13.3%	20.0%	13.2	<0.001
Constantly on guard	7.7%	13.2%	13.4	<0.001
Numb, detached	4.4%	8.8%	13.1	<0.001
PTSD screen positive (>2/4)	2.9%	8.4%	23.3	<0.001
Ever had sexual intercourse	16.2%	15.2%	0.3	0.594
Someone joked about wanting to have sex	12.1%	12.9%	0.3	0.613
Someone touched genitals/breasts	17.5%	15.7%	1.0	0.322
Someone tried to force sex but failed	4.7%	7.6%	6.0	0.014
Someone forced sex	2.9%	5.8%	8.2	0.004

Figure 2. Proportion of valid responses in the NVRC section given in the affirmative by randomization arm



Finally, in models predicting ever having been forced to have sex (Table 5), we found the odds of reporting forced sex were roughly twice as high in the NVRC arm as the verbal response arm. This was not affected by adjustment for covariates, which confirms the randomization of the response method. As expected, forced sex rose with age, with higher rates reported by adolescent women than men at all ages. Rates were also non-significantly higher for engaged and married individuals (Model 2). Scoring highly on the social desirability scale was associated with less reporting of forced sex, which might be expected if such experiences were stigmatized (Model 3). Finally, those reporting forced sex were substantially more likely to screen positive for probable PTSD (Model 4). We found no evidence that any of these associations were significantly modified by NVRC arm, as indicated by the higher values of AIC for interaction models than their non-interacted partner models at the bottom of Table 5.

Table 5. Regression analyses of reported lifetime history of unwanted forced sex

	Model 1		Model 2		Model 3		Model 4	
NVRC vs Standard arm	1.95	[1.09, 3.49]	2.15	[1.24, 3.70]	2.09	[1.17, 3.75]	1.75	[1.02, 3.00]
Female, 12-15			1.00		1.00		1.00	
Female, 16 or 17			1.14	[0.46, 2.83]	1.16	[0.47, 2.87]	1.09	[0.45, 2.61]
Female, 18-20			2.70	[1.35, 5.42]	2.69	[1.37, 5.26]	2.64	[1.45, 4.79]
Male, 12-15			0.25	[0.07, 0.83]	0.23	[0.07, 0.74]	0.28	[0.08, 1.04]
Male, 16 or 17			0.39	[0.16, 0.94]	0.38	[0.16, 0.91]	0.40	[0.15, 1.04]
Male, 18-20			0.83	[0.38, 1.83]	0.87	[0.40, 1.89]	0.83	[0.34, 1.98]
Never married, never engaged			1.00		1.00		1.00	
Engaged to be married			1.75	[0.47, 6.53]	1.60	[0.38, 6.74]	2.08	[0.57, 7.63]
Married			1.49	[0.66, 3.40]	1.54	[0.67, 3.56]	1.43	[0.61, 3.34]
PC-PTSD screen positive (≥3/4)							5.84	[2.35, 14.5]
Social desirability score:								
Lowest quintile					1.00			
2nd quintile					0.67	[0.30, 1.51]		
Middle quintile					0.61	[0.24, 1.50]		
2nd highest quintile					0.35	[0.08, 1.43]		
Highest quintile					0.26	[0.10, 0.67]		
N	1632		1630		1630		1617	
Intraclass Correlation Coefficient	0.08		0.16		0.14		0.17	
Akaike Information Criterion (AIC)	608.6		581.2		578.6		551.4	
AIC for interaction with: †								
Age, gender, marital status			584.0					
Social desirability quintiles					581.2			
PC-PTSD screen positive							553.3	

All results are from hierarchical (respondents nested in interviewers) logistic regression models. NVRC: non-verbal response cards; PC-PTSD: Primary care post-traumatic stress disorder screening tool. † AIC values below here are from models in which the named variables were interacted with NVRC arm.

DISCUSSION

In our survey of adolescents in a poor, rural part of Burkina Faso, we found that the use of a response method that protected the confidentiality of responses, by avoiding verbalized answers and hiding the response given from the interviewer, led to a higher level of reporting of the most sensitive questions we asked. A history of experiencing physical violence was common, with over 30% of respondents reporting this, as were having had someone joke about having sex with the respondent and unwanted sexual touching (12.5% and 16.6% respectively). NVRC arm prevalence was not significantly higher for any of these experiences. However, for less common experiences, notably sexual assault, having someone try to force sex but fail and having someone force sex (overall prevalence 3.7%, 6.1% and 4.3% respectively), NVRC arm respondents were between 1.5 and 2 times more likely to answer affirmatively. Similarly, NVRC arm respondents were 2.9 times more likely to affirm three or four items in the PC-PTSD screen and thus screen positive for probable PTSD.

Perhaps surprisingly, when we examined one variable in detail – that relating to forced sex – we did not find evidence that the effect of NVRCs was modified by other factors that also affected affirmative response rates on the multiplicative scale. Thus the effect appeared to apply equally to males and females, to married and unmarried individuals, to those screening positive for probable PTSD and for those displaying greater and lesser levels of social desirability. In practice, however, we may be more interested in additive interaction. For example, affirmative responses to the forced sex question for males were 1.5% in the Standard arm and 4.5% in the NVRC arm, compared to 4.7% and 8.7% for females respectively; while the prevalence ratio was larger for males (3.0 vs 1.9), the absolute difference was larger for females (4.2% vs 3.2%). When considering the impact of NVRCs it is important to consider whether the absolute or relative change is the meaningful one for the research question of interest.

Importantly, the use of these non-verbal response cards did not appear to lead to substantial errors in the responses given, as reflected in the low level of illogical answers. The already-low error rate fell as the survey process progressed. Since the study

researchers provided feedback to the field supervisors each week on which respondents were acquiring illogical responses (e.g. 'yes' in response to the question 'how many times have you harmed yourself in the past 12 months?'), this pattern of ever-lower levels of illogical responses suggests that such errors were due to insufficient interviewer training, rather than respondents incomprehension. Indeed, informal feedback from the interviewers suggested that most adolescent respondents quickly understood the method once it was explained and they had completed the four training questions at the beginning of the NVRC section.

As well as producing few illogical responses to individual questions, the NVRCs did not lead to less coherent sets of responses. Respondents in the NVRC arm in fact tended to provide groups of responses that were a little more internally consistent, as indicated by values of Cronbach's alpha, than respondents in the normal, verbal response arm. This may reflect greater willingness to provide answers that are internally consistent when they are kept private from the interviewer.

Strengths and limitations

Clearly NVRCs have some important limitations. Centrally, the interviewers have less control over the interview process when using these cards than they usually would: they are unable to identify illogical responses (although this was a relatively rare occurrence in this study) and cannot use skip patterns since they do not know the response to preceding questions. We allowed for skip patterns by telling respondents to point to 'zero' for Yes/No questions that depended on preceding responses if the question was not relevant (e.g. 'how old were you when you first had sex' for those who had not reached sexual debut). It remains unclear how broadly our finding that NVRCs can generate improved response patterns can be generalized – in terms of geography beyond sub-Saharan Africa, in terms of literacy status beyond largely illiterate populations or in terms of age beyond adolescents. These are testable in future work, however.

Conclusions

We have demonstrated the practicality and potential benefit of using of non-verbal response cards as a way of reducing socially desirable mis-reporting on sensitive topics in a

population with relatively low education and literacy. Our work extends previous use of this method in Ethiopia. There are several potential extensions to this work. These include: testing how any differences in reporting using these cards change on repeated use by the same respondents; evaluating whether it is the non-verbal or confidentiality aspects of the cards that lead to greater reporting of more socially unacceptable responses; and extending the use of these cards to older adults who may also face literacy and social desirability biases and may or may not be able to understand the novel methodology as quickly as did the adolescents in our study.

References

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress, 28*(6), 489-498.

Breiding, M. J., Smith, S. G., Basile, K. C., Walters, M. L., Chen, J., & Merrick, M. T. (2014). Prevalence and characteristics of sexual violence, stalking, and intimate partner violence victimization—National Intimate Partner and Sexual Violence Survey, United States, 2011. *Morbidity and Mortality Weekly Report, 63*(8), 1-18.

Cameron, R. P., & Gusman, D. (2003). The primary care PTSD screen (PC-PTSD): development and operating characteristics. *Primary Care Psychiatry, 9*(1), 9-14.

Darling, A.-M., Assefa, N., Bärnighausen, T. W., Berhane, Y., Canavan, C. R., Guwatudde, D., . . . Fawzi, W. W. (2019). Design and field methods of the ARISE Network Adolescent Health Study. *Submitted for publication*.

Dillon, G., Hussain, R., Loxton, D., & Rahman, S. (2013). Mental and physical health and intimate partner violence against women: A review of the literature. *International Journal of Family Medicine, 2013*, 313909.

Gregson, S., Mushati, P., White, P., Mlilo, M., Mundandi, C., & Nyamukapa, C. (2004). Informal confidential voting interview methods and temporal changes in reported sexual risk behaviour for HIV transmission in sub-Saharan Africa. *Sexually Transmitted Infections, 80*(suppl 2), ii36-ii42.

Haber, N., Harling, G., Cohen, J., Mutevedzi, T., Tanser, F., Gareta, D., . . . Fink, G. (2018). List randomization for eliciting HIV status and sexual behaviors in rural KwaZulu-Natal, South Africa: a randomized experiment using known true values for validation. *BMC Medical Research Methodology, 18*(1), 46.

Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The balanced inventory of desirable responding short form (BIDR-16). *Sage Open, 5*(4), 2158244015621113.

Jejeebhoy, S. J., Shah, I., & Thapa, S. (2013). *Sex Without Consent: Young People in Developing Countries*: Zed Books Ltd.

Jordan, C. E., Campbell, R., & Follingstad, D. (2010). Violence and women's mental health: The impact of physical, sexual, and psychological aggression. *Annual Review of Clinical Psychology, 6*, 607-628.

Langhaug, L. F., Sherr, L., & Cowan, F. (2010). How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. [Review]. *Tropical Medicine & International Health, 15*(3), 362-381. doi: 10.1111/j.1365-3156.2009.02464.x

Lensvelt-Mulders, G. J., Hox, J. J., Van der Heijden, P. G., & Maas, C. J. (2005). Meta-analysis of randomized response research: Thirty-five years of validation. *Sociological Methods & Research, 33*(3), 319-348.

Lindstrom, D. P., Belachew, T., Hadley, C., Hattori, M. K., Hogan, D., & Tessema, F. (2010). Nonmarital sex and condom knowledge among Ethiopian young people: improved estimates using a nonverbal response card. *Studies in Family Planning, 41*(4), 251-262.

Lindstrom, D. P., Hattori, M. K., Belachew, T., & Tessema, F. (2012). Lifting the curtain on the conditions of sexual initiation among youth in Ethiopia. *Journal of Adolescent Health, 50*(6), 614-620.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598.

Phillips, A. E., Gomez, G. B., Boily, M. C., & Garnett, G. P. (2010). A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *International Journal of Epidemiology, 39*(6), 1541-1555. doi: 10.1093/ije/dyq114

Pinheiro, P. S. (2006). World Report on Violence Against Children. Geneva, Switzerland: United Nations.

Sabourin, S., Laferrière, N., Sicuro, F., Coallier, J.-C., Cournoyer, L.-G., & Gendreau, P. (1989). Social desirability, psychological distress, and consumer satisfaction with mental health treatment. *Journal of Counseling Psychology, 36*(3), 352.

Sié, A., Louis, V., Gbangou, A., Müller, O., Niamba, L., Stieglbauer, G., . . . Becher, H. (2010). The Health and Demographic Surveillance System (HDSS) in Nouna, Burkina Faso, 1993–2007. *Global Health Action, 3*(1), 5284.

Stoltenborgh, M., Van Ijzendoorn, M. H., Euser, E. M., & Bakermans-Kranenburg, M. J. (2011). A global perspective on child sexual abuse: meta-analysis of prevalence around the world. *Child Maltreatment, 16*(2), 79-101.

Weathers, F., Blake, D., Schnurr, P., Kaloupek, D., Marx, B., & Keane, T. (2013). The life events checklist for DSM-5 (LEC-5), from https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp

World Health Organization. (2013). Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence. Geneva, Switzerland: WHO.