IJADR International Journal of Alcohol and Drug Research

The Official Journal of the Kettil Bruun Society for Social and Epidemiological Research on Alcohol

doi: 10.7895/ijadr.246 IJADR, 2018, 7(1), 29 – 39 ISSN: 1925-7066

Fieldworker effects on substance use reporting in a rural South African setting

Brian Houle[1,2,3], Nicole Angotti[2,3,4], F. Xavier Gómez-Olivé[3,5,7], and Samuel J. Clark[2,3,6,7]

[1] School of Demography, The Australian National University, Canberra, ACT, Australia
[2] CU Population Center, Population Program, Institute of Behavioral Science, University of Colorado at Boulder, Boulder, Colorado, United States
[3] MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[4] Department of Sociology and Center on Health, Risk and Society, American University, Washington, D.C., United States
[5] Center for Population and Development Studies, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
[6] Department of Sociology, The Ohio State University, Columbus, Ohio, United States
[7] INDEPTH Network, Accra, Ghana

Abstract

Aims: Fieldworkers capturing reports of sensitive behaviors, such as substance use, may influence survey responses and represent an important factor in response validity. We explored the effects and interaction of fieldworker and respondent characteristics (age and gender) in substance (tobacco and alcohol) use reporting. We aim to further the literature on conditional social attribution effects on substance use reporting in the context of South Africa, where accurate estimates of modifiable risk factors are critical for medical and public health practitioners and policy-makers in efforts to reduce chronic disease burden and mortality.

Design: We modeled substance use reporting using binary logistic regression. We also tested if fieldworker effects remained, allowing for correlation in reporting for respondents with the same fieldworker using multi-level logistic regression.

Setting: Agincourt Health and Socio-Demographic Surveillance System site, rural South Africa.

Participants: We used data from a 2010–2011 study on HIV and cardiometabolic risk, ages 15+ (N = 4,684).

Measures: Lifetime and current alcohol and tobacco use.

Findings: Respondents reported higher lifetime smoking use to older fieldworkers. Male respondents reported higher lifetime alcohol use to older fieldworkers. No fieldworker effects were significant on reports of current smoking. An older, male fieldworker increased the probability of reports of current alcohol use. Adjusting for intra-fieldworker correlation explained many of the observed fieldworker effects.

Conclusions: Our results highlight the importance of adjusting for interviewer characteristics to improve the accuracy of chronic disease risk factor estimates and validity of inferred associations. We recommend that surveys collecting information that may be subject to response bias routinely include anonymized fieldworker identifiers and demographic information. Analysts can then use these additional fieldworker data as a tool in evaluating probable bias in respondent reporting.

Introduction

Accurate estimates of behavioral chronic-disease risk factors are critical for medical and public health practitioners and policy-makers in efforts to reduce excess morbidity and mortality. Two important contributors to the global burden of chronic disease are tobacco and alcohol use (Lim et al., 2013). In much of sub-Saharan Africa, alcohol and tobacco use are key contributors to health loss (Institute for Health Metrics and Evaluation, Human Development Network, & The World Bank, 2013). In South Africa in particular, cumulative occurrence of alcohol and tobacco use in 2002–2004 were estimated at

Correspondence: Brian Houle; #9 Fellows Road, Acton ACT 2601, Australia; Telephone: +61 6125 7355; fax +61 6125 2992; E-mail: brian.houle@anu.edu.au
Financial support: We gratefully acknowledge funding from the National Institute on Aging (R24AG0 32112), National Institute of Child Health and Human Development (K01 HD057246; R01 HD054511), Wellcome Trust, UK (085477/Z/08/Z), the William and Flora Hewlett Foundation, and the Anglo American Chairman's Fund. This project also received funding and administrative support from the University of Colorado Population Center, funded by the Eunice Shriver National Institute of Child Health and Human Development (NICHD R21 HD051146). The funders had no role in the design of the study nor in the collection, analysis, and interpretation of data and writing the manuscript.
Keywords: South Africa, rural, interviewer effects, substance use, respondent reporting

38.7% and 30.0%, respectively, with men far more likely to have become alcohol and tobacco users than women (Van Heerden et al., 2009). Smoking prevalence was also much more common among men (35% compared to 10% of women) (Department of Health, Medical Research Council, & ORC Macro, 2007). In 2003, about 39% of men and 16% of women reported drinking alcohol in the past year (Department of Health et al., 2007). These estimates were based on survey reports, however, and compared to more objective consumption estimates, appear to underestimate substance use in South Africa (Department of Health et al., 2007), perhaps due in part to social pressures to underreport and to reporting bias. Other influencing factors may include recall bias, underestimation of standard drinks, and undercoverage in surveys of the heaviest drinkers due to selective non-response (Gmel & Rehm, 2004).

Estimates of individual risk behaviors are often based on surveys using respondent reports. Under-reporting of socially undesirable behaviors, including substance use, may be driven by social desirability, in which responses are adjusted to be closer to perceived norms governing acceptable behavior (Johnson & Parsons, 1994; Tourangeau & Yan, 2007). While capturing respondent reports of potentially sensitive behaviors or excessive substance use, interviewers may also influence the survey-response process; thus, interviewers represent an important factor to consider when producing estimates and conducting inference (Davis, Couper, Janz, Caldwell, & Resnicow, 2010; Elliott & West, 2015). These two factors may also interact: for instance, respondents may adjust their responses based on perceived norms or values they attribute to the interviewer. Respondents adapting their responses based on inferences of observable characteristics of interviewers such as age, gender, and race/ethnicity may yield systematic differences in respondents' reported behaviors that vary by interviewer characteristics.

Direct or social attribution effects are observable characteristics of interviewers that respondents may evaluate in their reporting (Fendrich, Johnson, Shaligram, & Wislar, 1999; Johnson & Moore, 1993). Johnson and Parsons (1994), for example, found that respondents of both genders were more likely to report substance use to male interviewers. Additionally, conditional social attribution effects represent judgments of interviewer characteristics that vary according to subject characteristics (Fendrich et al., 1999). Fendrich et al. (1999) found that for drug-use reporting, the impact of interviewer race/ethnicity varied according to the respondent's race/ethnicity: Black participants had lowered odds of reporting drug use to Black interviewers, while the responses of White participants and those of other races/ethnicities did not vary by interviewer race/ethnicity. Another study on respondent-reported alcohol consumption found an interaction effect between interviewer and respondent age: younger respondents reported lower alcohol consumption to older interviewers, while older respondents reported higher alcohol consumption to older interviewers (Heeb & Gmel, 2001). Finally, interviewer influence and social desirability biases may also be culturally determined, highlighting the need for further research on interviewer effects in different settings (Bernardi, 2006; Kim & Kim, 2016; Lalwani, Shrum, & Chiu, 2009; McCombie & Anarfi, 2002).

In this paper we explore the effects and interaction of interviewer and respondent characteristics in substance use reporting. We use data from the Ha Nakekela ("We Care") cross-sectional study, conducted in 2010–2011 in the Agincourt subdistrict in rural South Africa, that included respondent-reported tobacco and alcohol use. We aim to explore social attribution effects by including interviewer characteristics, and to test whether these effects persist after allowing for correlation among respondents with the same interviewer. Based on a prior study exploring interviewer effects on sexual-behavior reporting in the same setting (see Houle et al., 2016), we hypothesize that respondents will report less substance use to older interviewers, and that male respondents will report higher substance use to male interviewers. The present inquiry is particularly important given the marked gender disparity in reported substance use in South Africa, and the potential interactions with social norms and desirability around associated risk behaviors. It also furthers the literature on social attribution effects on substance use reporting in the context of South Africa, and provides a comparative basis for other studies using different survey procedures.

Methods

The study received ethical approvals from the University of the Witwatersrand Human Research Ethics Committee (M10458) and the Mpumalanga Provincial Research and Ethics Committee. Written informed consent (or assent for minors) was obtained for all participants.

Sample

We conducted a cross-sectional, community-level HIV and chronic diseases prevalence and risk factors survey in 2010–2011 in the Agincourt subdistrict in rural South Africa. The area has been under demographic and health surveillance since 1992 using an annual household census, including collection of vital events and household, social, and economic factors (Kahn et al., 2012). From an eligible population of 34,413 residents based on the 2009 census, we randomly sampled 7,662 individuals ages 15+, stratified by age and sex.

Data Collection

Ten trained fieldworkers, randomly assigned to different villages and households, visited sampled participants in their homes and invited them to participate in the study. The field team consisted of five men and five women aged between 28–44, with six fieldworkers under age 35 and four fieldworkers 35 years of age and over; mid-study, one male fieldworker left and was replaced by a female fieldworker of a different age group. A total of 4,684 individuals consented to participation (Gómez-Olivé et al., 2013). The fieldworkers administered a questionnaire on cardiometabolic diseases risk and substance use, adapted from the World Health Organization STEPwise approach to chronic disease risk factor surveillance (World Health Organization, 2017). Each home visit lasted approximately 45 minutes. Fieldworkers were similar in many characteristics: all completed secondary school, were predominately Christian, and per Agincourt site guidelines, were Xitsonga/Shangaan speakers living in the study site. We did not have additional information available on other fieldworker characteristics.

Statistical Analysis

We modeled four outcomes from the questionnaire to explore fieldworker age and gender effects and their interaction with respondent characteristics on lifetime and current substance use:

Ever smoked: Have you ever smoked any tobacco product such as cigarettes, cigars, or pipes?

Currently smokes: Do you currently smoke (you will smoke if you have the possibility) any tobacco products, such as cigarettes, cigars, or pipes?

Ever drank: Have you ever consumed an alcoholic drink such as beer, wine, spirits, fermented cider, thothotho [a high-proof, distilled spirit], or traditional beer?

Currently drinks: Have you consumed an alcoholic drink within the past 30 days?

We modeled each of the four outcomes using complete-case binary logistic regression and built the model in stages by testing for improvements in model fit using nested likelihood ratio tests. First, we modeled each outcome using respondent characteristics only, including sex, age (and age^2 when indicated to model changes along the life course—i.e., allowing age to have a nonlinear association on substance use reporting), quintiles of household socioeconomic status (SES), education, employment and union status, and village of residence.

Second, we included fieldworker characteristics (age categorized as < 35 and 35+ years of age and sex) and tested interactions between respondent and fieldworker characteristics to assess their impact on tobacco- and alcohol-use reporting. As we have limited variability on fieldworker effects, we modeled them as fixed effects. We selected our fieldworker age cut-offs to give variation for comparison, while not categorizing fieldworkers into unrealistic age categories; our cut-offs for "younger" and "older" fieldworkers takes into account that life expectancy in Agincourt is 55 years for males and 62 years for females (Kahn et al., 2012), and that the average age at first birth is 20 (Williams et al., 2013); and these age cut-offs were used previously in a similar study exploring fieldworker effects on sexual behavior reporting (Houle et al., 2016). While in this analysis we tested the effects of social categories of fieldworker age, we also tested models including differences in respondent and fieldworker age, finding most effects to be non-significant. We also tested if any outlying fieldworker(s) unduly influenced our results by including an indicator for each fieldworker in each model. We summarize each model using average marginal effects, including variation by significant fieldworker characteristics (if indicated).

Third, we estimated a multi-level model including a fieldworker random intercept to allow for correlation in respondent reporting to the same fieldworker (as in Dailey & Claus, 2001). As we included respondent sex and age as independent variables in our models, we report unweighted estimates. All analyses were completed using Stata 14.

Results

Respondent sample characteristics and fieldworker characteristics are presented in Table 1. The sample was approximately 60% female, with a mean age of 42 years. Unemployment was high (79%) as well as having a previous migration history (57%). Reporting lifetime and current substance use was much higher for males compared to females.

Table 1

Respondent sample characteristics by sex and fieldworker characteristics: age–sex stratified random sample of ages 15+, Agincourt, South Africa, 2010–2011.

	Male (*n* = 1840)		**Female (*n* = 2771)**		**Total (*n* = 4611)**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Age	41.7	20.3	42.2	18.7	42.0	19.3
	n	**Proportion**	***n***	**Proportion**	***n***	**Proportion**
2009 SES quintiles						
First (lowest)	292	16	436	16	728	16
Second	353	19	537	19	890	19
Third	379	21	597	22	976	21
Fourth	362	20	558	20	920	20
Fifth (highest)	454	25	643	23	1097	24
Past migration history						
No	915	50	1052	38	1967	43
Yes	925	50	1719	62	2644	57

	Male (n = 1840)		Female (n = 2771)		Total (n = 4611)	
	Mean	SD	Mean	SD	Mean	SD
Age	41.7	20.3	42.2	18.7	42.0	19.3
	n	**Proportion**	***n***	**Proportion**	***n***	**Proportion**
Education						
None	369	20	747	27	1116	24
Less than 6 years	235	13	314	11	549	12
6+ years	1236	67	1710	62	2946	64
Employed						
No	1340	73	2297	83	3637	79
Yes	500	27	474	17	974	21
Union status						
Never	824	45	916	33	1740	38
Current	765	42	1072	39	1837	40
Previous	251	14	783	28	1034	22
Ever smoked						
No	1190	65	2728	98	3918	85
Yes	650	35	43	2	693	15
Ever drank						
No	519	28	2072	75	2591	56
Yes	1321	72	699	25	2020	44
Currently smokes						
No	279	43	32	74	311	45
Yes	371	57	11	26	382	55
Currently drinks[a]						
No	228	23	244	61	472	34
Yes	769	77	158	39	927	66

[a] In the questionnaire, "currently drinks" was conditioned also on "Have you consumed an alcoholic drink within the past 12 months?" to account for the difference between those indicating they ever drank and those currently drinking in the past month.

Ever Smoked

Table 2 (column a) shows the results of the binary logistic regression for reporting having ever smoked, including respondent characteristics only. First including age^2 ($p < .001$), next interacting sex and age ($p = .010$), and finally sex and age^2 ($p < .001$) significantly improved model fit and resulted in the base model with respondent characteristics only (Table 2a). Males had higher odds of ever smoking, while higher SES and being in a current union lowered the odds of ever smoking.

We next included fieldworker sex and age, shown in Table 2 (column b). Interacting fieldworker and respondent age significantly improved model fit ($p = .017$; Table 2b). Figure 1 shows the predicted probability of ever smoking by respondent and fieldworker age. Respondents had higher odds of reporting ever smoking to older fieldworkers, and this effect increased among older respondents.

Including a random intercept for the fieldworker significantly improved model fit ($p < .001$; Table 2c). The intraclass correlation coefficient (ICC) was .06, representing the total variance shared among individuals with the same fieldworker. The total effect of fieldworker age on reporting ever smoking remained significant, accounting for correlation in respondent reporting with the same fieldworker ($p = .029$).

Currently smokes

Table 2 (column d) shows the results of the binary logistic regression for reporting currently smoking, including respondent characteristics only. First including age^2 ($p = .018$) and then interacting sex and age ($p < .001$) significantly improved model fit and resulted in the base model with respondent characteristics only (Table 2d). Males had higher odds of currently smoking, while higher SES and being in a current union lowered the odds of currently smoking.

We next included fieldworker sex and age, shown in Table 2 (column e). Including fieldworker sex and age showed no effect on the odds of currently smoking. Figure 2 shows the predicted probability of currently smoking by respondent sex and age. For males, the probability of currently smoking declined with age, while for females the probability increased with age. Including a random intercept for the fieldworker significantly improved model fit ($p < .001$; ICC = .07; Table 2f).

Table 2

Binary logistic regression of reporting ever smoking (columns a–c) and current smoking (columns d–f), by Base Model (respondent characteristics), added fieldworker effects (sex and age), and added fieldworker effects including a random intercept for the fieldworker. All models adjusted for village.

	Ever smoke					
	(A) Base (*N* = 4611)		**(B) With fieldworker effects (*N* = 4611)**		**(C) With random intercept (*N* = 4611)**	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
			Respondent characteristics			
Male	89.159	[56.803, 139.947]	86.238	[55.222, 134.675]	92.917	[59.136, 145.994]
Age	1.033	[1.006, 1.060]	1.000	[1.000, 1.000]	1.036	[1.008, 1.064]
Age^2	1.000	[0.999, 1.001]	0.999	[0.999, 1.000]	0.999	[0.999, 1.000]
Male X age	1.017	[0.991, 1.043]	1.015	[0.989, 1.042]	1.015	[0.989, 1.042]
Male X age^2	0.998	[0.997, 0.999]	0.998	[0.998, 0.999]	0.998	[0.998, 0.999]
2009 SES quintiles						
First (lowest)	–	–	–	–	–	–
Second	0.880	[0.630, 1.229]	0.896	[0.640, 1.253]	0.857	[0.608, 1.207]
Third	0.688	[0.490, 0.965]	0.703	[0.500, 0.988]	0.653	[0.461, 0.926]
Fourth	0.730	[0.515, 1.034]	0.764	[0.538, 1.085]	0.730	[0.510, 1.045]
Fifth (highest)	0.556	[0.395, 0.784]	0.555	[0.393, 0.784]	0.519	[0.365, 0.738]
Past migration history	0.977	[0.789, 1.210]	0.983	[0.793, 1.219]	1.010	[0.811, 1.257]
Education						
None	–	–	–	–	–	–
Less than 6 years	1.292	[0.918, 1.819]	1.305	[0.924, 1.842]	1.374	[0.967, 1.953]
6+ years	0.803	[0.583, 1.106]	0.805	[0.583, 1.112]	0.810	[0.583, 1.125]
Employed	0.937	[0.739, 1.187]	0.926	[0.728, 1.178]	0.904	[0.705, 1.160]
Union status						
Never	–	–	–	–	–	–
Current	0.600	[0.448, 0.803]	0.599	[0.447, 0.803]	0.572	[0.425, 0.772]
Previous	0.876	[0.620, 1.239]	0.873	[0.617, 1.235]	0.864	[0.606, 1.230]
Constant	0.024	[0.012, 0.049]	0.024	[0.012, 0.050]	0.023	[0.009, 0.054]
			Fieldworker and respondent effects			
Male fieldworker			1.020	[0.823, 1.265]	0.975	[0.526, 1.807]
Aged 35+ fieldworker			1.028	[0.778, 1.359]	0.990	[0.510, 1.921]
Aged 35+ fieldworker X respondent age			0.996	[0.982, 1.009]	0.998	[0.984, 1.012]
Aged 35+ fieldworker X respondent age^2			1.001	[1.000, 1.001]	1.001	[1.000, 1.001]
					σ^2	95% CI
					0.212	[0.078, 0.576]

	Currently smoke					
	(D) Base (*n* = 693)		**(E) With fieldworker effects (*n* = 693)**		**(F) With random intercept (*n* = 693)**	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
			Respondent characteristics			
Male	12.439	[3.556, 43.519]	12.204	[3.471, 42.902]	11.590	[3.256, 41.257]
Age	1.069	[1.016, 1.125]	1.000	[1.000, 1.000]	1.000	[1.000, 1.000]
Age^2	0.999	[0.999, 1.000]	0.999	[0.999, 1.000]	0.999	[0.999, 1.000]
Male X age	0.920	[0.877, 0.965]	0.919	[0.876, 0.965]	0.919	[0.875, 0.964]
Male X age^2						
2009 SES quintiles						
First (lowest)	–	–	–	–	–	–
Second	0.372	[0.208, 0.667]	0.368	[0.205, 0.660]	0.383	[0.211, 0.693]
Third	0.319	[0.176, 0.580]	0.313	[0.172, 0.569]	0.326	[0.177, 0.601]
Fourth	0.347	[0.189, 0.636]	0.339	[0.184, 0.625]	0.358	[0.192, 0.668]
Fifth (highest)	0.360	[0.198, 0.656]	0.367	[0.202, 0.669]	0.378	[0.205, 0.695]
Past migration history	0.906	[0.631, 1.303]	0.911	[0.632, 1.313]	0.890	[0.613, 1.293]
Education						
None	–	–	–	–	–	–
Less than 6 years	1.295	[0.751, 2.235]	1.330	[0.769, 2.300]	1.280	[0.732, 2.237]
6+ years	0.980	[0.572, 1.678]	1.008	[0.588, 1.730]	1.055	[0.608, 1.831]
Employed	1.109	[0.747, 1.647]	1.070	[0.713, 1.605]	1.144	[0.748, 1.749]
Union status						
Never	–	–	–	–	–	–
Current	0.533	[0.333, 0.853]	0.523	[0.326, 0.839]	0.519	[0.319, 0.842]
Previous	0.712	[0.408, 1.243]	0.698	[0.398, 1.222]	0.692	[0.390, 1.228]

	Currently smoke					
	(D) Base (n = 693)		(E) With fieldworker effects (n = 693)		(F) With random intercept (n = 693)	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Constant	0.682	[0.150, 3.099]	0.802	[0.172, 3.739]	0.989	[0.190, 5.153]
Fieldworker and respondent effects						
Male fieldworker			0.888	[0.617, 1.278]	0.950	[0.456, 1.977]
Aged 35+ fieldworker			0.721	[0.507, 1.025]	0.646	[0.307, 1.360]
Aged 35+ fieldworker X respondent age						
Aged 35+ fieldworker X respondent age^2						
					σ^2	95% CI
					0.236	[0.065, 0.853]

Figure 1

Predicted probability of ever smoking using average marginal effects, by respondent and fieldworker age, Agincourt, South Africa, 2010–2011.

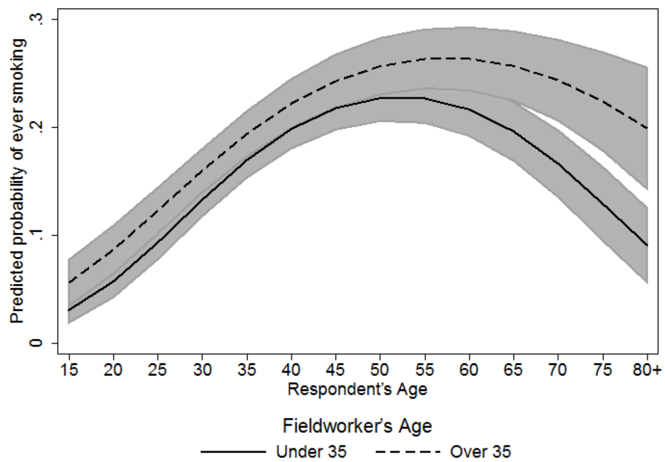
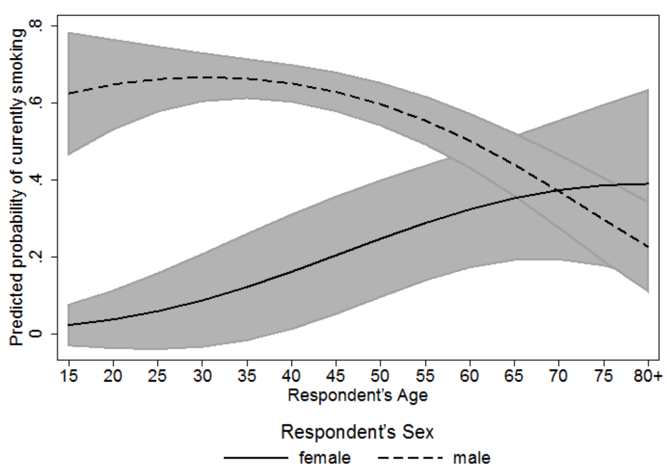


Figure 2

Predicted probability of currently smoking using average marginal effects, by respondent sex and age, Agincourt, South Africa, 2010–2011.



Ever Drank

Table 3 (column a) shows the results of the binary logistic regression for reporting having ever drank, including respondent characteristics only. Including age^2 significantly improved model fit (p = .036) and resulted in the base model with respondent characteristics only (Table 3a). Males had higher odds of ever drinking, while higher SES and education, and being in a current union, lowered the odds of ever drinking. The probability of ever drinking increased with age.

We next included fieldworker sex and age, shown in Table 3 (column b). Interacting fieldworker age and respondent sex significantly improved model fit (p = .013; Table 3b). Figure 3 shows the predicted probability of ever drinking by respondent sex and age, as well as by fieldworker age. Male respondents had a higher probability of reporting ever drinking to older fieldworkers.

Including a random intercept for the fieldworker significantly improved model fit (p < .001; ICC = .03; Table 3c). The overall effect of fieldworker age on reporting ever drinking was no longer significant after accounting for intra-fieldworker correlation (p = .089).

Currently Drinks

Table 3 (column d) shows the results of the binary logistic regression for reporting currently drinking, including respondent characteristics only. Including age^2 significantly improved model fit (p < .001) and resulted in the base model with respondent characteristics only (Table 3d). Males had higher odds of currently drinking compared to females. The probability of currently drinking increased with respondent age.

We next included fieldworker sex and age, shown in Table 3 (column e). Figure 4 shows the predicted probability of currently drinking by respondent age and fieldworker sex and age. Having a male or older fieldworker increased the probability of reporting currently drinking.

Including a random intercept for the fieldworker significantly improved model fit (p < .001; ICC = .15; Table 3f). The effects of fieldworker sex (p = .187) and age (p = .246) were no longer significant after accounting for intra-fieldworker correlation.

Table 3

Binary logistic regression of reporting ever drinking (columns a–c) and current drinking (columns d–f), by Base Model (respondent characteristics), added fieldworker effects (sex and age), and added fieldworker effects including a random intercept for the fieldworker. All models adjusted for village.

	Ever drink					
	(A) Base (*N* = 4611)		**(B) With fieldworker effects (*N* = 4611)**		**(C) With random intercept (*N* = 4611)**	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Respondent characteristics						
Male	8.295	[7.184, 9.578]	7.254	[6.086, 8.646]	7.848	[6.554, 9.399]
Age	1.001	[0.994, 1.008]	1.001	[0.994, 1.007]	1.002	[0.995, 1.009]
Age^2	1.000	[1.000, 1.000]	1.000	[1.000, 1.000]	1.000	[1.000, 1.000]
2009 SES quintiles						
First (lowest)	–	–	–	–	–	–
Second	0.940	[0.749, 1.180]	0.946	[0.754, 1.188]	0.928	[0.738, 1.167]
Third	0.757	[0.603, 0.950]	0.764	[0.608, 0.959]	0.740	[0.587, 0.931]
Fourth	0.832	[0.659, 1.050]	0.842	[0.667, 1.064]	0.845	[0.668, 1.070]
Fifth (highest)	0.749	[0.594, 0.944]	0.740	[0.586, 0.934]	0.744	[0.588, 0.942]
Past migration history	1.063	[0.917, 1.234]	1.060	[0.914, 1.230]	1.076	[0.926, 1.251]
Education						
None	–	–	–	–	–	–
Less than 6 years	0.596	[0.464, 0.764]	0.597	[0.466, 0.766]	0.601	[0.468, 0.773]
6+ years	0.603	[0.483, 0.754]	0.601	[0.481, 0.752]	0.597	[0.477, 0.748]
Employed	1.144	[0.960, 1.363]	1.124	[0.942, 1.341]	1.023	[0.853, 1.225]
Union status						
Never	–	–	–	–	–	–
Current	0.669	[0.549, 0.815]	0.673	[0.552, 0.820]	0.652	[0.534, 0.797]
Previous	0.951	[0.752, 1.203]	0.957	[0.756, 1.211]	0.948	[0.747, 1.203]
Constant	0.593	[0.390, 0.903]	0.598	[0.389, 0.919]	0.581	[0.340, 0.993]
Fieldworker and respondent effects						
Male fieldworker			0.935	[0.811, 1.079]	0.948	[0.630, 1.425]
Aged 35+ fieldworker			1.031	[0.854, 1.244]	1.017	[0.655, 1.577]
Aged 35+ fieldworker X respondent male			1.432	[1.079, 1.900]	1.359	[1.021, 1.809]
					σ^2	95% CI
					0.093	[0.035, 0.248]

	Currently drink					
	(D) Base (*n* = 1399)		**(E) With fieldworker effects (*n* = 1399)**		**(F) With random intercept (*n* = 1399)**	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
			Respondent characteristics			
Male	5.873	[4.420, 7.802]	6.069	[4.549, 8.097]	6.423	[4.698, 8.783]
Age	1.036	[1.022, 1.050]	1.036	[1.022, 1.050]	1.037	[1.022, 1.052]
Age^2	0.999	[0.999, 1.000]	0.999	[0.999, 1.000]	0.999	[0.999, 1.000]
2009 SES quintiles						
First (lowest)	–	–	–	–	–	–
Second	0.748	[0.499, 1.121]	0.768	[0.511, 1.154]	0.682	[0.440, 1.059]
Third	0.633	[0.420, 0.952]	0.664	[0.440, 1.003]	0.564	[0.362, 0.879]
Fourth	0.972	[0.633, 1.491]	0.995	[0.647, 1.532]	0.852	[0.536, 1.355]
Fifth (highest)	1.092	[0.710, 1.680]	1.089	[0.706, 1.680]	0.953	[0.599, 1.514]
Past migration history	0.855	[0.652, 1.120]	0.860	[0.655, 1.129]	0.864	[0.648, 1.151]
Education						
None	–	–	–	–	–	–
Less than 6 years	1.154	[0.682, 1.953]	1.142	[0.673, 1.940]	1.238	[0.709, 2.163]
6+ years	0.898	[0.552, 1.460]	0.873	[0.535, 1.427]	0.934	[0.555, 1.573]
Employed	0.994	[0.722, 1.369]	1.046	[0.756, 1.448]	1.154	[0.814, 1.636]
Union status						
Never	–	–	–	–	–	–
Current	0.752	[0.513, 1.103]	0.762	[0.518, 1.122]	0.785	[0.521, 1.183]
Previous	0.738	[0.475, 1.145]	0.772	[0.495, 1.205]	0.810	[0.503, 1.304]
Constant	0.91	[0.411, 2.012]	0.573	[0.252, 1.307]	0.532	[0.167, 1.695]
			Fieldworker and respondent effects			
Male fieldworker			1.652	[1.249, 2.186]	1.943	[0.725, 5.206]
Aged 35+ fieldworker			1.662	[1.258, 2.196]	1.829	[0.660, 5.070]
Aged 35+ fieldworker X respondent male						
					σ^2	95% CI
					0.561	[0.224, 1.406]

Figure 3

Predicted probability of ever drinking using average marginal effects, by respondent sex and age, and fieldworker age, Agincourt, South Africa, 2010–2011.

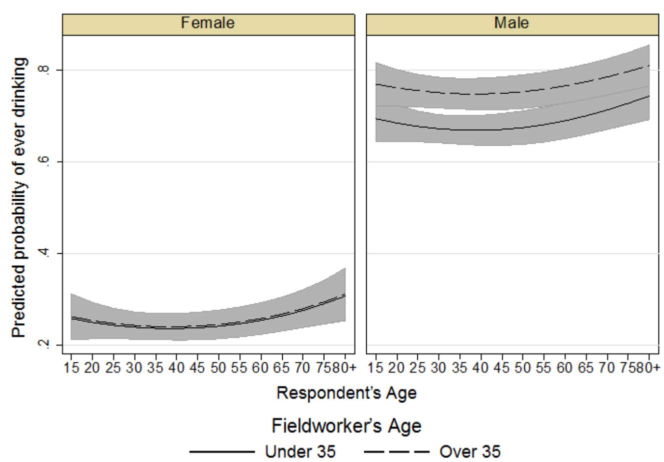
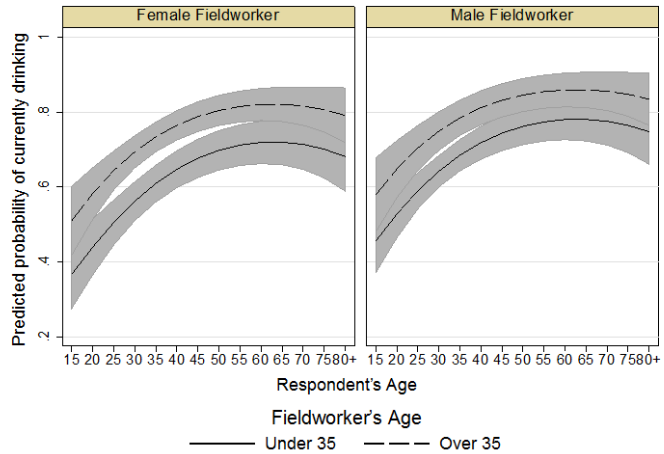


Figure 4

Predicted probability of currently drinking using average marginal effects, by respondent age and fieldworker sex and age, Agincourt, South Africa, 2010–2011.



Discussion

We found evidence for both direct and conditional social attribution effects of interviewers on respondent reporting of substance use in rural South Africa. For reporting lifetime substance use, we found conditional social attribution effects of interviewer age for smoking (conditional on respondent age) and drinking (conditional on respondent gender). For reporting current substance use we found direct social attribution effects of interviewer age and gender on drinking, but not for smoking. We also found that accounting for intra-interviewer correlation often made these interviewer effects non-significant, suggesting that the similarity of individual responses within interviewers explained many of the interviewer effects. In other words, many of the observed interviewer effects were explained by the correlation induced from individuals responding to the same interviewer.

Earlier work in Agincourt explored interviewer effects on sexual behavior reporting and found that respondents reported fewer sexual partners and "safer" sexual behaviors (such as condom use and discussing HIV with sexual partners) to older interviewers (Houle et al., 2016). Men also reported higher numbers of sexual partners to female interviewers (Houle et al., 2016). In the present study on substance use behaviors, however, we found striking differences: respondents had a higher probability of reporting ever and current substance use to older interviewers, and males being interviewed by older interviewers had a higher probability of reporting ever drinking. These contrasting results suggest several important considerations. Foremost, social desirability bias varies depending on the dimension of life being queried, and thus should be analyzed and interpreted separately. Notably, behaviors deemed "risky" for chronic diseases, like drinking and smoking, may not carry the same negative valence in the Agincourt setting as risk behaviors linked to other diseases, such as lack of or inconsistent condom use or having multiple sexual partners on risk for HIV infection, and thus are not subject to the same sorts of biases. Moreover, prominent social marketing efforts and public health campaigns in South Africa and elsewhere in the region where HIV prevalence is high have emphasized the link between modifying sexual behaviors to avert HIV infection. The link between substance use and mortality, however, even as the burden of chronic diseases becomes more profound, is not as widely established.

Second, drinking and smoking may also represent a socially sanctioned activity for men, and thus may be commonplace to discuss, particularly in the presence of other men. Men may also report this behavior differently to female interviewers because of the associations drinking has with irresponsibility and even violent behavior, such as drunk driving and intimate partner violence (Jewkes, 2002; Jewkes, Levin, & Penn-Kekana, 2002). The low levels of reported use by women in our study suggest that substance use may be considered a "male activity." Studies of substance use disorders among women and men in South Africa have also found harsher criticisms of women than men, which may be attributed to the gendered "moral discourses" around substance use, particularly the association with sexual deviance and subversion of traditional gender roles among female users (Myers, Fakier, & Louw, 2009).

We acknowledge several study limitations, the first three of which we have noted elsewhere in similar analyses (see Houle et al., 2016). First, due to the cross-sectional nature of the data, we can only make assumptions about respondent reporting. Future studies may employ more targeted techniques, such as eliciting perceived age versus using actual interviewer age (Davis et al., 2010), to explore interviewer effects on respondent reporting. Second, we may be detecting other unobserved interviewer effects (e.g., community reputation, degree of religiosity, marital status) that we were unable to measure. That most interviewer effects were no longer significant in our multi-level models indicates that correlation in reporting for respondents with the same interviewer reflects other shared factors. A strength of the data, however, is the lack of interviewer variability in other socio-demographic characteristics, suggesting that other factors during the survey process warrant further study, such as interviewer skill and personality in questionnaire delivery ("role-restricted interviewer effects") (also see Bignami-Van Assche, Reniers, & Weinreb, 2003; Weinreb, 2006). Third, it is unknown how respondents actually view the perceived age of the interviewer. While our focus was on social categories of age, we also modeled differences in interviewer and respondent age, with mostly non-significant results. We also attempted to accommodate for age differences by including interactions between respondent and interviewer age when it improved model fit. Fourth, given the low levels of lifetime and current use of substances reported by females, we were not able to examine gender-specific interactions in further detail. Similarly, the substance use behaviors available in the survey were limited, and other measures may be more sensitive to interviewer effects. Fifth, while interviewers were randomly assigned to respondents (avoiding confounding interviewer and respondent characteristics), this study used a small number of interviewers relative to a large number of respondents, which may increase the design effects of individual interviewers on study results as well as limit our ability to explore other sociodemographic factors (Davis et al., 2010). Finally, we lack systematic information on third-party presence, which with home-based interviews is likely to have affected interviewer variation in respondent reporting (Aquilino, Wright, & Supple, 2009).

Our results highlight the importance of adjusting for interviewer characteristics to improve the accuracy of chronic-disease risk-factor estimates and validity of inferred associations. This is particularly important in settings undergoing rapid social and epidemiological change, to provide a strong evidence base for effective prevention and intervention efforts, as well as effective targeting of health services and care management for those most in need (Houle, Clark, Gomez-Olive, Kahn, & Tollman, 2014; Tollman et al., 2008). Based on these results, we recommend that surveys collecting information that may be susceptible to social attribution and other biases routinely include anonymized interviewer identifiers and other demographic information (see also Elliott & West, 2015). Analysts can then use this information as a useful tool in assessing the possibility and extent of bias in respondent reporting, and, where possible, adjust for interviewer effects when consequential for their research question.

Acknowledgments

For helpful comments on the manuscript, we thank Vusumusi Goodwill Dlamini.

References

Aquilino, W. S., Wright, D. L., & Supple, A. J. (2009). Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use. *Substance Use & Misuse*, *35*(6–8), 845–867. doi:10.3109/10826080009148424

Bernardi, R. A. (2006). Associations between Hofstede's cultural constructs and social desirability response bias. *Journal of Business Ethics*, *65*(1), 43–53. doi:10.1007/s10551-005-5353-0

Bignami-Van Assche, S., Reniers, G., & Weinreb, A. A. (2003). An assessment of the KDICP and MDICP data quality: Interviewer effects, question reliability and sample attrition. *Demographic Research*, *Special 1*, 31–76. doi:10.4054/DemRes.2003.S1.2

Dailey, R. M., & Claus, R. E. (2001). The relationship between interviewer characteristics and physical and sexual abuse disclosures among substance users: A multilevel analysis. *Journal of Drug Issues*, *31*(4), 867–888. doi:10.1177/002204260103100404

Davis, R. E., Couper, M. P., Janz, N. K., Caldwell, C. H., & Resnicow, K. (2010). Interviewer effects in public health surveys. *Health Education Research*, *25*(1), 14–26. doi:10.1093/her/cyp046

Department of Health, Medical Research Council, & ORC Macro. (2007). South Africa Demographic and Health Survey 2003. Pretoria, South Africa: Department of Health.

Elliott, M. R., & West, B. T. (2015). "Clustering by interviewer": A source of variance that is unaccounted for in single-stage health surveys. *American Journal of Epidemiology*, *182*(2), 118–126. doi:10.1093/aje/kwv018

Fendrich, M., Johnson, T., Shaligram, C., & Wislar, J. S. (1999). The impact of interviewer characteristics on drug use reporting by male juvenile arrestees. *Journal of Drug Issues*, *29*(1), 37.

Gmel, G., & Rehm, J. T. (2004). Measuring alcohol consumption. *Contemporary Drug Problems*, *31*, 467–540.

Gómez-Olivé, F. X., Angotti, N., Houle, B., Klipstein-Grobusch, K., Kabudula, C., Menken, J., . . . Clark, S. J. (2013). Prevalence of HIV among those 15 and older in rural South Africa. *AIDS Care*, *25*, 1122–1128. doi:10.1080/09540121.2012.750710

Heeb, J. L., & Gmel, G. (2001). Interviewers' and respondents' effects on self-reported alcohol consumption in a Swiss health survey. *Journal of Studies on Alcohol and Drugs*, *62*(4), 434–442. doi:10.15288/jsa.2001.62.434

Houle, B., Angotti, N., Clark, S. J., Williams, J., Gomez-Olive, F. X., Menken, J., . . . Tollman, S. (2016). Let's talk about sex, maybe: Interviewers, respondents, and sexual behavior reporting in rural South Africa. *Field Methods*, *28*, 112–132. doi:10.1177/1525822X15595343

Houle, B., Clark, S. J., Gomez-Olive, F. X., Kahn, K., & Tollman, S. M. (2014). The unfolding counter-transition in rural South Africa: Mortality and cause of death, 1994–2009. *PLoS One*, *9*(6), e100420. doi:10.1371/journal.pone.0100420

Institute for Health Metrics and Evaluation, Human Development Network, & The World Bank. (2013). *The global burden of disease: Generating evidence, guiding policy - Sub-Saharan Africa regional edition*. Seattle, WA, United States: Institute for Health Metrics and Evaluation.

Jewkes, R. (2002). Intimate partner violence: causes and prevention. *The Lancet*, *359*(9315), 1423-1429.

Jewkes, R., Levin, J., & Penn-Kekana, L. (2002). Risk factors for domestic violence: Findings from a South African cross-sectional study. *Social Science & Medicine*, *55*(9), 1603–1617.

Johnson, T. P., & Moore, R. W. (1993). Gender interactions between interviewer and survey respondents: Issues of pornography and community standards. *Sex Roles*, *28*, 243–261.

Johnson, T. P., & Parsons, J. A. (1994). Interviewer effects on self-reported substance use among homeless persons. *Addictive Behaviors*, *19*(1), 83–93. doi:10.1016/0306-4603(94)90054-X

Kahn, K., Collinson, M. A., Gomez-Olive, F. X., Mokoena, O., Twine, R., Mee, P., . . . Tollman, S. M. (2012). Profile: Agincourt Health and Socio-demographic Surveillance System. *International Journal of Epidemiology*, *41*(4), 988–1001. doi:10.1093/ije/dys115

Kim, S. H., & Kim, S. (2016). National culture and social desirability bias in measuring public service motivation. *Administration & Society*, *48*(4), 444–476. doi:10.1177/0095399713498749

Lalwani, A. K., Shrum, L. J., & Chiu, C. Y. (2009). Motivated response styles: The role of cultural values, regulatory focus, and self-consciousness in socially desirable responding. *Journal of Personality and Social Psychology*, *96*(4), 870–882. doi:10.1037/a0014622

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., . . . Andrews, K. G. (2013). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260.

McCombie, S. C., & Anarfi, J. K. (2002). The influence of sex of interviewer on the results of an AIDS survey in Ghana. *Human Organization*, *61*(1), 51–57. doi:10.17730/humo.61.1.em6l865y3v9y7y2l

Myers, B., Fakier, N., & Louw, J. (2009). Stigma, treatment beliefs, and substance abuse treatment use in historically disadvantaged communities. *African Journal of Psychiatry*, *12*, 218–222. Retrieved from https://www.ajol.info/index.php/ajpsy/article/viewFile/48497/34850

Tollman, S. M., Kahn, K., Sartorius, B., Collinson, M. A., Clark, S. J., & Garenne, M. L. (2008). Implications of mortality transition for primary health care in rural South Africa: A population-based surveillance study. *The Lancet*, *372*, 893–901. doi:10.1016/S0140-6736(08)61399-9

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–883. doi:10.1037/0033-2909.133.5.859

Van Heerden, M. S., Grimsrud, A. T., Seedat, S., Myer, L., Williams, D. R., & Stein, D. J. (2009). Patterns of substance use in South Africa: Results from the South African Stress and Health Study. *South African Medical Journal*, *99*(5), 358–366.

Weinreb, A. A. (2006). The limitations of stranger-interviewers in rural Kenya. *American Sociological Review*, *71*(6), 1014–1039. doi:10.1177/000312240607100607

Williams, J., Ibisomi, L., Sartorius, B., Kahn, K., Collinson, M., Tollman, S., & Garenne, M. (2013). Convergence in fertility of South Africans and Mozambicans in rural South Africa, 1993–2009. *Glob Health Action*, *6*, 19236. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=23364078

World Health Organization. (2017). *STEPwise approach to surveillance (STEPS).* Geneva, Switzerland: World Health Organization. Retrieved from http://www.who.int/ncds/surveillance/steps/en/