Alternative Estimates of Lifetime Prevalence Of Abortion from Indirect Survey Questioning Methods

By Sarah K. Cowan, Lawrence L. Wu, Susanna Makela and Paula England

Sarah K. Cowan is assistant professor, and Lawrence L. Wu and Paula England are professors, Department of Sociology; New York University. Susanna Makela is Ph.D. candidate, Department of Statistics, Columbia University.

Additional supporting information may be found in the online version of this article at the publisher's website.

Abortion is a frequent medical procedure, which is undergone by diverse women in the United States and has profound demographic and political implications.[1–5] Despite the frequency with which it is performed, estimates of lifetime prevalence of abortion—that is, how many American women alive today have had an abortion—are lacking. We argue that this basic fact is critical knowledge for several reasons. First, journalists, participants in social movements and politicians routinely want to use the lifetime prevalence as a basis for arguing that abortion is common and therefore normal, or common and therefore a tragedy. In its absence, they misinterpret estimates based on the synthetic cohorts constructed from the most recent period age-specific first-abortion rates as measuring lifetime prevalence.[6] This error spreads misinformation and tarnishes the work of abortion researchers.[7]

Second, if estimates of true lifetime prevalence were available, they would provide important clues about the abortion histories not only of women who are currently of reproductive age, but also of those who were of reproductive age before abortion was legalized federally. While older women are no longer at risk of needing an abortion, their histories provide the closest picture we can get to abortion prevalence when the procedure was not legal or when its availability differed markedly across states—a scenario that the United States may be fast approaching again because of increased legal restrictions on access to abortion.[8] In addition, data regarding women currently of reproductive age have their own weaknesses, some of which can be overcome by a relatively new method, which we propose and pilot here, the list experiment.

Finally, knowledge of lifetime prevalence is necessarily a matter of basic science. Abortion is important—not just because of deep political conflicts over it, but also for the lives of women, their partners and their families. More than one million abortions are performed each year in the United States,[1] and yet we do not know the rudimentary fact of how many women have ever had one. It is a fundamental truth of demography that considering prevalence from various angles—such as period, synthetic cohort, cohort—reveals new patterns, trends and disparities. For other vital events, such as births and deaths, we routinely calculate all of these measures. Yet we do not for abortion, despite its demographic, social and political relevance.

The primary reason that estimates of abortion prevalence are lacking is that women seriously underreport abortions in virtually all surveys.[9,10] In this comment, we propose a list experimental design for estimating lifetime prevalence. This method was designed explicitly to explore sensitive topics and has improved the accuracy of responses to sensitive items in a variety of settings.[11] We conducted a pilot study using a slight variant, a double list design, that has some desirable properties. Results from our pilot, while preliminary in nature, suggest both the potential promise and some disadvantages of this approach.

CURRENT MEASURES

Comparing survey responses of women with external measures of abortion rates reveals that women underreport abortions in nationally representative surveys.[9,12] The National Survey of Family Growth (NSFG), for instance, gathers a complete pregnancy history in both face-to-face and self-administered portions of the interview. Underreporting is ubiquitous in both modes: The number of abortions reported by women as occurring in particular years is less than half the number that abortion providers report.[13] The underreporting is so severe that the developers of the NSFG themselves advise against drawing substantive conclusions about abortion prevalence from these data.

In principle, one could estimate abortion prevalence among women alive today by combining a series of period age-specific first-abortion rates from surveys of patients with census estimates of population counts for different age-groups. Thus, for example, such data would provide first-abortion rates for a cohort of women who were 15 in 2000, were 20 in 2005 and so forth, yielding a cumulative rate that could be applied to the number of 30-year-olds in 2015. Doing these calculations for each birth cohort alive today, with appropriate cohort corrections for mortality, would yield an estimate of lifetime abortion prevalence if the population of interest was not subject to in- or out-migration.

No previous study of which we are aware has pursued this strategy, in large part because the necessary series of age-specific first-abortion rates do not exist. In particular, no data exist that would give age-specific first-abortion rates for cohorts of women who had abortions before the procedure became legal at the federal level in 1973, despite evidence of substantial numbers of legal and illegal abortions in those cohorts.*[14–16] The data currently used for age-specific abortion rates come from the Guttmacher

*Prior work suggests stability in the distribution of number of prior abortions by cohort, but this work considers only legal abortions after 1973 (source: Cowan SK, Cohort abortion measures for the United States, *Population and Development Review,* 2013, 39(2):289–307). Underreporting of illegal abortions may be quite different and violate the assumptions made in this work.

Abortion Patient Survey, which was first conducted in 1987 and is updated periodically. Since 1973, the Centers for Disease Control and Prevention has collected data on abortion in the United States, but coverage is not complete, and it does not provide the data needed to estimate age-specific first-abortion rates. If lifetime prevalence were estimated using the time series, it would likely be more accurate than estimates from surveys such as the NSFG. Still, estimates from surveys of abortion patients may be subject to biases from site or survey nonresponse, or from missing data on respondents' age and whether the abortion was a first one.*

A further potential source of bias concerns medication abortion. Surveys of abortion providers often do not include facilities that provide only medication abortions, and while these procedures currently represent a small proportion of total abortions, estimates of lifetime abortion prevalence obtained from provider data will be biased if women increasingly have their first or only abortion in medication-only facilities.[1]

These issues motivate our consideration of an alternative method for estimating lifetime prevalence. We propose turning to established techniques to indirectly ask about sensitive items. Indirect techniques provide greater anonymity than survey questionnaires—even self-administered ones—by masking individual responses.

INDIRECT QUESTIONING TECHNIQUES

Two commonly used techniques indirectly elicit prevalence of sensitive information: the randomized response method and the list experiment. The randomized response method includes random noise, which masks the individual response. Respondents are given a randomizing device, such as a coin flip, but the interviewer does not know the outcome of the randomization. For example, if a coin flip is used, respondents are asked to answer the sensitive item truthfully in the case of heads, and to give a predetermined response in the case of tails. Because the randomization has a known distribution, researchers can infer deviations from that distribution result from the sensitive item. This technique has been used to obtain indirect estimates of abortion in a variety of places, including Mexico[17,18] and Botswana.[19]

The randomized response method suffers from a few weaknesses. First, half the sample does not answer the sensitive question, so the statistical power obtained from a given sample size is diminished. Second, the randomized response method primarily involves an interviewer, and American women reveal more abortions in self-administered questionnaires than in face-to-face interviews.[9,20]

We argue that the more promising technique for estimating lifetime abortion prevalence is the list experiment (also known as the item-count or the unmatched-count technique). This method obscures individual responses by asking not about a particular item, but rather about a number of items at once. Respondents are divided into treatment and control groups, each of which is given a list of items; the two lists contain all of the same nonsensitive items, but the treatment group's list also includes the sensitive item of interest. Respondents are asked how many of the statements are true for them, not which are true for them. The difference in the mean response for the treatment list and the mean response for the control list is an estimate of the overall prevalence of the sensitive characteristic. List experiments rely upon assumptions regarding the anticipated direction of bias in direct questioning. For instance, we would anticipate that respondents would underreport stigmatized or illegal behaviors, so if the list experiment increases reporting, then we assume that it yields an improved estimate of the lifetime prevalence of such behaviors.

The list experiment is easily conducted using a self-administered questionnaire, thus capitalizing on American women's increased willingness to reveal their abortions when an interviewer is not present.†[9,20] Additionally, it can easily be incorporated into existing large-sample surveys.

Three other indirect questioning techniques are worth mentioning. The first, the endorsement method, asks nonsensitive questions and randomly pairs them with a sensitive object. It is not appropriate for estimating the prevalence of behaviors, however, but is better suited for assessing attitudes. The second technique, the anonymous third-party reporting method (also known as the confidant's method or the best-friend report), asks respondents about other people's behavior. It has been used to study abortion rates in non-Western settings[21,22] and self-induced abortion in Texas.[23] In its simplest form, it asks women about their best friend's behavior. A limitation is that women may keep their abortion secret even from their best friend.[24,25] Alternatively, respondents can be asked to report on other members of their social network, but this requires including a large battery of questions to establish respondents' social network,[26] and these questions are known to be subject to interviewer effects, including fatigue.[27] The third technique, not strictly an indirect questioning technique, is the sealed envelope method.[26,28] Respondents answer the sensitive question on a paper form, which they then seal in an envelope and deposit in a bag or box filled with other envelopes. Respondents are thus keenly aware that their response to the sensitive question will not be linked to them or to their responses to other survey questions. A limitation is that this method requires additional surveyor time, as the envelopes and forms need to be transported, stored and processed.

THE DOUBLE LIST EXPERIMENT

In the list design, respondents are randomly assigned to the control list or to the treatment list. Subtracting the mean of the responses to the control list from the mean

*For example, the 2008 Guttmacher survey had a response rate of 74%. In addition, facility staff provided some information for nonrespondents, but not information on whether they were having a first abortion.[6]

†The list experiment improves upon direct questioning even when both are conducted via computer (source: Heerwig JA and McCabe BJ, Education and social desirability bias: the case of a black presidential candidate, *Social Science Quarterly,* 2009, 90(3):674–686).

of the responses to the treatment list reveals the proportion of respondents who have the sensitive characteristic. A weakness of the design is that statistical power is reduced because the control group is not asked about the sensitive item. A variation on the design, the double list design, increases statistical power by adding another set of control statements and a corresponding set of treatment statements; each respondent answers the control items from one list and the treatment items from the other. Just as in the single list design, the response to each list indicates how many items are true for the respondent. Given that the double list design entails two treatment lists and two control lists, it produces two estimates, which can be compared or averaged. Notably, with this technique, respondents see both sets of control items, but see the sensitive statement only once. Because all respondents are assigned to one treatment and one control group, statistical power is improved. The method's disadvantage is that it adds to the length of the survey and the attendant cognitive burden.

For our pilot, we constructed lists using items that would be appropriate in a general survey on health. If all the items on a treatment list are true for a respondent, then she has revealed that the sensitive item (in this case, a history of abortion) is true for her. Similarly, if none of the items is true for the respondent, then she has revealed her answer to the sensitive item. Thus, the control statements should be chosen so that the likelihood of ceiling or floor effects is minimized—that is, the control items should not likely be all true or all false. The most effective way to ensure this is to have items that are negatively correlated with each other. In selecting items for our experiment, we found few established health questions that are negatively correlated,* and so we chose at least one item that had a high prevalence and another that had a low prevalence in pretests and external surveys.[29] Several of the items are similar to those in the National Health and Nutrition Examination Survey (NHANES), and hence can provide an external set of estimates by which to assess a variety of potential sources of bias.

Our first set of control statements, control list A, consisted of the following: "I have gotten a vaccine of any sort, including as a child"; "I have visited a dentist"; "I have been told by a doctor or other health professional that I had a stroke"; and "I have used a hearing aid." Respondents were asked "How many of the above are true for you?" and could reply 0–4. The vaccine and the dentist questions represented high-prevalence behaviors and prevented floor effects; the stroke and hearing aid questions represented low-prevalence experiences and prevented ceiling effects. Treatment list A consisted of the control items plus the sensitive item, "I have had an abortion."

The second set of control statements, control list B, consisted of the following: "I have had a cold or the flu"; "I have had a stomach virus or food poisoning"; "I have been told by a doctor or other health professional that I had cancer"; and "I have been told by a doctor or other health professional that I had rheumatoid arthritis." Again, we had two high-prevalence items (cold and stomach virus), which prevented floor effects, and two low-prevalence items (cancer and rheumatoid arthritis), which prevented ceiling effects. Treatment list B consisted of the control items plus the abortion item.

We administered our pilot to a diverse convenience sample of more than 1,200 U.S. resident adult women who were recruited and participated online. Data collection was approved by the institutional review boards at Columbia University and New York University. Given that we used a convenience sample, the results can be used only to assess the method and not substantively as a lifetime prevalence estimate.

Respondents were randomly assigned to one of three tracks: In one, women were asked direct questions about all of the list items and about abortion; in another, they saw control list A and treatment list B; and in the third, they saw treatment list A and control list B. (Asking directly about the list items in the first track maintained survey length, question ordering effects and cognitive fatigue across all three tracks. In addition, it yielded direct estimates that could be used as an external check against the lists.) Details regarding the sample and the balance among tracks can be found in the appendix (Supporting Information).

There were two additional sources of randomization. First, items in all tracks were randomly ordered. Second, respondents assigned to the list tracks were randomly assigned to see treatment or control lists first.

RESULTS FROM THE DOUBLE LIST PILOT

On average, women reported that in list A, 2.20 treatment items and 1.98 controls were true for them (Table 1). Similarly, for list B, on average, 2.02 treatment items and 1.80 controls were true for them. The difference between the treatment and control means for each list gives the abortion prevalence estimate. Results from both lists indicate that 22% of women in the sample had had an abortion. This estimated prevalence is four percentage points higher than the 18% obtained through the direct question.

This double list experiment shows promising results in two ways. First, the method is functioning as anticipated, showing that women are more likely to reveal an abortion in a list design than when asked directly (even though the question was not asked by a live interviewer).† Second, lifetime prevalence estimates derived from the two lists are identical, as we would anticipate.

However, the results of our pilot should be interpreted with caution, for two reasons. First, our pilot used a convenience sample; hence our 22% estimate should not be

*In looking for negatively correlated health items, we considered questions from the National Health and Nutrition Examination Survey regarding hearing aid usage, blood pressure, diabetes, asthma, obesity, arthritis, congestive heart failure, coronary heart disease, heart attack, stroke, cancer, smoking, blood cholesterol tests, dentist visits and flossing. However, no pair of these is known to be strongly, or even moderately, negatively correlated.

†We do not find order effects based on whether the treatment list was seen before or after the control list. This was based on a two-sample t test.

TABLE 1. Results of a pilot experiment comparing estimates of lifetime abortion prevalence obtained through a double list indirect questioning technique and a direct question among a convenience sample of U.S. resident adult women

Questioning technique	Mean no. of items endorsed†		Estimated abortion prevalence‡
	Treatment group	Control group	
List A			
Measure	2.20	1.98	0.22***
N	395	429	na
List B			
Measure	2.02	1.80	0.22***
N	430	395	na
Direct question			
Measure	na	na	0.18***
N	na	na	408

***Estimate is different from zero at p<.001. †The range is 0–5 for the treatment group and 0–4 for the control group. ‡For the list designs, prevalence is the difference between treatment and control means. *Note:* na=not applicable.

regarded as a substantively correct estimate of lifetime prevalence. Second, the two-sample test required for the difference between the direct estimate and the list estimates shows that the four-point difference is not statistically significant; a much larger sample would be required to detect a significant difference.

SAMPLE SIZE CALCULATIONS

What is the minimum sample size needed to conclude that with prespecified probabilities of type I and type II error levels, the estimate from the list experiment is significantly different from the estimate from direct questioning? While Glynn provides formulas for sample size calculations for the double list experiment, he does not provide formulas for determining whether the estimates from the double list experiment are different from our direct estimates.[30] Building on his calculations, we derived formulas for the sample size needed to determine this (Supporting Information). In other words, we want to calculate the sample size such that with 80% power (i.e., a 20% type II error level), the 95% confidence interval (i.e., a 5% type I error level) for the difference in the two estimates excludes zero.[30]

We can think of a test of the difference in the estimates from the two experiments as a test of the difference in two proportions. This test requires that in each sample, we know the mean, standard deviation and sample size. For the direct questioning estimate, the mean and standard deviation are the familiar quantities for proportions. For the double list estimate, we use the variance formula provided in Glynn,[30] which involves the variances of the item counts for the four lists in the experiment, as well as the covariance of the item counts between the two lists seen by each respondent. The covariance terms appear in the formula because the double list estimate is the average of the estimates from lists A and B, and these estimates are not independent: Each respondent is in the treatment group for one list and the control group for the other. We then need to set values for each of these six terms (four variances, two covariances). In the absence of a pilot study, a researcher can choose a worst-case scenario in which all of the variances are the same and equal to some given quantity.

Additionally, the researcher must choose a value for the expected difference between the two estimates. We assume that the list experiment estimate is larger than the direct estimate and use the observed difference of four percentage points from our pilot study. We also assume that the sample is evenly divided among the two list and the direct questioning groups.

Given the quantities described above, we can calculate the sample size needed for a one-sided 95% confidence interval for the difference between the two estimates to exclude zero with 80% power. Using the values observed in our study, the experiment we presented here requires a sample of more than 6,500 women to detect a difference between an estimate of the lifetime abortion prevalence obtained by asking directly and one obtained through the list experiment.

CONCLUSION

On both theoretical and practical grounds, the list technique is an attractive approach to obtaining estimates of lifetime abortion prevalence. Prior studies have used this method for a number of sensitive items, including risky sexual behaviors,[31] breaches of professional ethics[32] and voting for a racial minority candidate.[33] Because it adds only a few survey questions, it has a relatively minimal impact on respondent burden and survey length; it likewise can be easily incorporated into the structure of existing surveys with minimal effort and thus could result in better and more frequent estimates of lifetime prevalence than are currently available. We used a slight variant, a double list design, and our pilot showed the promise of the method. But because of the convenience sample used in our pilot, our findings should be interpreted as, at best, suggestive of the potential of a list design for estimating lifetime abortion prevalence.

Could the design of our pilot be improved? In retrospect, the absence of negatively correlated items may have decreased statistical power substantially. Additionally, one list contained two items (hearing aid and stroke) that have low prevalence in general, but that may be both more common and positively correlated for older women. If so, all items in this list may have been true for some older women, who would then hit a ceiling that would reveal that they had had an abortion. It is also possible that the list items we used did not sufficiently mask the sensitive question—whether a woman had ever had an abortion.* We suggest tests to determine whether adding sensitive items affects reporting. More generally, researchers should develop control lists that best fit the content of their surveys. (The control lists from our pilot, for example, would likely not

*As a check, we analyzed data from NHANES, which revealed that hearing aid and stroke are not well correlated, even among older women, so we are not particularly concerned about these two items, but more generally caution against using two low-prevalence items that are correlated with each other.

work well in a fertility survey, given that questions on dental health would seem jarring and out of context.) One weakness of the design, however, is inescapable: It does not reveal any particular woman's abortion history; it therefore does not allow her abortion history to be linked to her other characteristics.

These issues should also be weighed in light of the potential biases in estimates of lifetime abortion prevalence obtained in conventional ways. Available data exclude any possibility of estimating lifetime abortion prevalence for older women; by contrast, our list design would include those women in the estimates. With increasing restrictions on abortion access,[8] these women's experiences may illuminate not just America's past, but its future.

In addition, estimates relying on data from abortion providers may become less reliable if the number of medication abortions performed by nontraditional abortion providers increases. It is therefore likely that all methods for estimating lifetime abortion prevalence will be subject to bias, albeit from different sources.

Yet, not knowing the facts about lifetime abortion prevalence is, we have argued, a glaring gap in our understanding of abortion. For these reasons, we think it important that the research community engage in active efforts, such as ours, to investigate alternative methods for obtaining estimates of lifetime abortion prevalence.

REFERENCES

1. Jones RK and Jerman J, Abortion incidence and service availability in the United States, 2011, *Perspectives on Sexual and Reproductive Health,* 2014, 46(1):3–14.

2. Jones RK, Finer LB and Singh S, *Characteristics of U.S. Abortion Patients, 2008,* New York: Guttmacher Institute, 2010.

3. Cook EA et al., *Between Two Absolutes: Public Opinion and the Politics of Abortion*, Boulder, CO: Westview Press, 1992.

4. Killian M and Wilcox C, Do abortion attitudes lead to party switching? *Political Research Quarterly,* 2008, 61(4):561–573.

5. Abramowitz AI, It's abortion, stupid: policy voting in the 1992 presidential election, *Journal of Politics,* 1995, 57(1):176–186.

6. Jones RK and Kavanaugh ML, Changes in abortion rates between 2000 and 2008 and lifetime incidence of abortion, *Obstetrics & Gynecology,* 2011, 117(6):1358–1366.

7. Lee MYH, The stale claim that "one in three" women will have an abortion by age 45, *Washington Post,* 2015, https://www.washingtonpost.com/news/fact-checker/wp/2015/09/30/the-stale-claim-that-one-in-three-women-will-have-an-abortion-by-age-45/.

8. Boonstra HD and Nash E, A surge of state abortion restrictions puts providers—and the women they serve—in the crosshairs, *Guttmacher Policy Review,* 2014, 17(1):9–15.

9. Jones RK and Kost K, Underreporting of induced and spontaneous abortion in the United States: an analysis of the 2002 National Survey of Family Growth, *Studies in Family Planning,* 2007, 38(3):187–197.

10. Peytchev A, Peytcheva E and Groves RM, Measurement error, unit nonresponse, and self-reports of abortion experiences, *Public Opinion Quarterly,* 2010, 74(2):319–327.

11. Rosenfeld B, Imai K and Shapiro JN, An empirical validation study of popular survey methodologies for sensitive questions, *American Journal of Political Science,* 2015, doi: 10.1111/ajps.12205.

12. Fu H et al., Measuring the extent of abortion underreporting in the 1995 National Survey of Family Growth, *Family Planning Perspectives,* 1998, 30(3):128–133, 138.

13. National Center for Health Statistics (NCHS), *Public Use Data File Documentation, 2011–2013 National Survey of Family Growth, User's Guide,* Hyattsville, MD: NCHS, 2014.

14. Tietze C, Two years' experience with a liberal abortion law: its impact on fertility trends in New York City, *Family Planning Perspectives,* 1973, 5(1):36–41.

15. Abernathy JR, Greenberg BG and Horvitz DG, Estimates of induced abortion in urban North Carolina, *Demography,* 1970, 7(1):19–29.

16. Krannich RS, Abortion in the United States: past, present, and future trends, *Family Relations,* 1980, 29(3):365–374.

17. Lara D et al., Measuring induced abortion in Mexico: a comparison of four methodologies, *Sociological Methods & Research,* 2004, 32(4):529–558.

18. Lara D et al., The measure of induced abortion levels in Mexico using random response technique, *Sociological Methods & Research,* 2006, 35(2):279–301.

19. Oliveras E and Letamo G, Examples of methods to address underreporting of induced abortion: preceding birth technique and randomized response technique, in: Singh S, Remez L and Tartaglione A, eds., *Methodologies for Estimating Abortion Incidence and Abortion-Related Morbidity: A Review,* New York: Guttmacher Institute; and Paris: International Union for the Scientific Study of Population, 2010, pp. 63–70.

20. Lessler JT, Weeks MF and O'Reilly JM, Results from the National Survey of Family Growth Cycle V pretest, in: American Statistical Association, *1994 Proceedings of the Section on Survey Research Methods,* Alexandria, VA, 1994. pp. 64–70, http://www.amstat.org/sections/srms/Proceedings/papers/1994_009.pdf.

21. Yeatman S and Trinitapoli J, Best-friend reports: a tool for measuring the prevalence of sensitive behaviors, *American Journal of Public Health,* 2011, 101(9):1666–1667.

22. Rossier C et al., Estimating clandestine abortion with the confidants method—results from Ouagadougou, Burkina Faso, *Social Science & Medicine,* 2006, 62(1):254–266.

23. Grossman D et al., Knowledge, opinion and experience related to abortion self-induction in Texas, Texas Policy Evaluation Project, *Research Brief,* 2015, https://www.utexas.edu/cola/txpep/_files/pdf/TxPEP-Research-Brief-KnowledgeOpinionExperience.pdf.

24. Cowan SK, Secrets and misperceptions: the creation of self-fulfilling illusions, *Sociological Science,* 2014, doi: 10.15195/v1.a26.

25. Foster DG et al., Attitudes and decision making among women seeking abortions at one U.S. clinic, *Perspectives on Sexual and Reproductive Health,* 2012, 44(2):117–124.

26. Rossier C, Measuring abortion with the anonymous third party reporting method, in: Singh S, Remez L and Tartaglione A, eds., *Methodologies for Estimating Abortion Incidence and Abortion-Related Morbidity: A Review,* New York: Guttmacher Institute; and Paris: International Union for the Scientific Study of Population, 2010, pp. 107–124.

27. Paik A and Sanchagrin K, Social isolation in America: an artifact, *American Sociological Review,* 2013, 78(3):339–360.

28. Juarez F, Cabigon J and Singh S, The sealed envelope method of estimating induced abortion: how much of an improvement? in: Singh S, Remez L and Tartaglione A, eds., *Methodologies for Estimating Abortion Incidence and Abortion-Related Morbidity: A Review,* New York: Guttmacher Institute; and Paris: International Union for the Scientific Study of Population, 2010, pp. 99–106.

29. Special tabulations of data from the 2011–2012 National Health and Nutrition Examination Survey and the 2012 National Health Interview Survey.

30. Glynn AN, What can we learn with statistical truth serum? Design and analysis of the list experiment, *Public Opinion Quarterly,* 2013, 77(S1):159–172.

31. LaBrie JW and Earleywine M, Sexual risk behaviors and alcohol: higher base rates revealed using the unmatched-count technique, *Journal of Sex Research,* 2000, 37(4):321–326.

32. Dalton DR, Wimbush JC and Daily CM, Using the unmatched count technique (UCT) to estimate base rates for sensitive behavior, *Personnel Psychology,* 1994, 47(4):817–829.

33. Heerwig JA and McCabe BJ, Education and social desirability bias: the case of a black presidential candidate, *Social Science Quarterly,* 2009, 90(3):674–686.

Acknowledgments

The authors acknowledge the research assistance of Jessie Ford, Daniela Urbina and Eliza Brown, and support during part of the data collection from the Robert Wood Johnson Health and Society Scholars program.

Author contact: *sarahkcowan@nyu.edu*

doi: *10.1363/48e11216*