HHS Public Access

Author manuscript

Womens Health Issues. Author manuscript; available in PMC 2023 May 30.

Published in final edited form as:

Womens Health Issues. 2022 ; 32(1): 33–40. doi:10.1016/j.whi.2021.08.003.

Estimating Prevalence of Abortion Using List Experiments: Findings from a Survey of Women in Delaware and Maryland

Alexandra Kissling,

Heide Jackson

Maryland Population Research Center, University of Maryland

Abstract

Background—Widespread under-reporting of abortion persists in survey data. The list experiment, a measurement tool designed to elicit truthful responses to sensitive questions, may alleviate under-reporting.

Methods—Using The Statewide Survey of Women of Reproductive Age in Delaware and Maryland (N= 2747), we estimate the prevalence of abortion in Maryland and Delaware using a double list experiment.

Results—We find 21% (95% CI: 16.8%, 25.3%) of respondents aged 18–44 ever had an abortion and disparities in abortion prevalence by age, race, education, income, marital status, and insurance status. Respondents who were Black (37.0%; 95% CI: 27.1%, 46.8%), had less than a college degree (24.8%; 95% CI: 18.3%, 31.3%), were in a cohabiting relationship (39.0%; 95% CI: 29.1%, 48.9%), were living in households with incomes under $50,000 (28.6%; 95% CI: 19.7%, 37.5%), and were currently covered by Medicaid (42.8%; 95% CI: 27.6%, 58.0%) were more likely than their counterparts to have ever had an abortion.

Conclusions—List experiments yield estimates of abortion substantially higher than those obtained from direct questions. Findings demonstrate external validity through consistency with estimates from administrative data sources and gold standard abortion provider survey data.

Keywords

abortion; list experiment; experimental survey data

Introduction

Abortion is a common but stigmatized procedure (Norris et al., 2011).[1] Researchers, policy makers, and advocates seek estimates of abortion prevalence over the lifecourse because such estimates demonstrate how common abortion is, provide histories of abortion regardless of age or the legal status of abortion, and because lifetime prevalence of abortion contributes to basic demography and science (Jones & Kost, 2007). However, we lack accurate measures of lifetime abortion due to the limited availability of administrative records and because survey respondents may not answer truthfully due to stigma (Jones

& Kost, 2007). Recent research on Add Health, NLSY97, and NSFG found that the underestimation of abortion is not a result of failing to capture the appropriate sample, women who have had an abortion. Rather, underreporting by respondents drives incomplete abortion information (Lindberg et al., 2020).

Researchers need innovative survey designs to obtain more accurate abortion information. Some surveys have shifted from interviewer administered questions to computer-based or audio-based questions to increase respondent privacy thus potentially reducing the effects of abortion stigma (for a complete review, see Lindberg et al., 2020). While alternate modes can improve reporting (Lindberg et al., 2020), they may produce lower reports of other sensitive behaviors (e.g., sexual activity) (Mensch et al., 2008) and increase reporting errors because of misunderstood questions and filling out the survey too quickly (Ghanem et al., 2005; Jaya et al., 2008).

Indirect third-party reporting is another approach used to collect information on this sensitive topic.[2] The “best friend approach” and “confidante method” ask respondents to think of their best friend (or multiple friends) and answer questions about outcomes of that person’s pregnancies (Sedgh & Keogh, 2019; Yeatman & Trinitapoli, 2011). The Network Scale Up Method (NSUM) uses information about respondents’ social networks to estimate the size of “hidden populations” (e.g., people who have had an abortion) (Sully et al., 2020). Though useful, third-party approaches may be expensive to implement, complicated to analyze, and laden with unrealistic assumptions (Rossier, 2003; Singh et al., 2010). For example, the best friend approach assumes a reciprocal relationship between friends, that individuals will agree on their best friend when in reality there may not be such a 1:1 association (Yeatman & Trinitapoli, 2011). The NSUM requires respondents to accurately describe their social network, which may be challenging (Sully et al., 2020).

Because of limitations in third-party and ACASI approaches, abortion researchers have begun to use another indirect method: list experiments. In list experiments, respondents are randomly assigned to a treatment or control list. The control list contains a series of non-sensitive items often having to do with the respondent’s health. The treatment list contains the same items as the control list, with the addition of a sensitive item, having had an abortion (Blair & Imai, 2012; Moseson et al., 2015). Respondents are asked to report how many *but not which* items apply to them. The prevalence of the sensitive item is determined by subtracting the mean of responses in the control list from the mean of responses in the treatment list.

By not asking directly which items in the list are true, list experiments may improve reporting of sensitive items and protect respondent privacy (Glynn, 2013). If this supposition holds, list experiments may also be less time intensive and less expensive to implement than alternative approaches. However, list experiments require some conditions be met. The presence of the sensitive item must not affect respondent answers to the control items and respondents must provide truthful answers for the sensitive item (Blair & Imai, 2012). If

[2]Direct methods ask respondents about their own abortion history and are typically used to get individual level data. Indirect methods ask about others’ abortion histories and are typically used to assess prevalence.

respondents are concerned that their responses may be disclosed, this may affect the veracity of responses (Kuklinski et al., 1997). Thus, there should be a low proportion of women for whom none or all of the control items apply.

Though well established in the political science literature (Blair & Imai, 2012; Gonzalez-Ocantos et al., 2012), list experiments have only recently been used to measure abortion. Some international studies successfully estimated the lifetime prevalence of abortion using a list experiment (Huber-Krum, Hackett, et al., 2020; Huber-Krum, Karadon, et al., 2020; Moseson et al., 2015, 2019, 2021) while others could not (Bell & Bishai, 2019; Elewonibi et al., 2021). A pilot study in the US used a double list experiment where all respondents answered to a treatment list to increase sample size and improve statistical power (Cowan et al., 2016).[3] However, the study (Cowan et al.(2016) used a convenience sample and did not check that the assumptions of list experiments were met. We are unaware of any study that has used both representative survey data and the list experiment method to estimate lifetime prevalence of abortion in the United States or any other high income country setting.

Using a recent survey of patients at abortion providers, when combined with American Community Survey data, Jones & Jerman (2017) found that about a quarter of women[4] in the US have ever had an abortion. However, due to difficulties in accurately measuring abortion using survey data, we know less about how lifetime prevalence varies across social and demographic groups. Administrative sources show that the annual abortion rate has generally declined since 2010 (Jones & Jerman, 2017). Younger women have a lower lifetime chance of having an abortion and declines in the yearly abortion rate are largest among those 19 years old and younger (Jones & Jerman, 2017). Further, the annual abortion rate is lowest among non-Hispanic White women and highest among Black women (Jatlaoui, 2018). Abortion rates also vary by education and income as college graduate women have the lowest abortion rates and women with incomes below the federal poverty level having the highest abortion rates (Jones & Jerman, 2017). While informative, administrative sources are limited by incomplete information. Some states, like Maryland, do not report abortion administrative data either via its Department of Public Health or the CDC (Jatlaoui, 2018). Even when such administrative records are available, they may still be incomplete and provide limited information about the characteristics of people seeking abortion. While Jones and Jerman (2017) combine US Census data with a provider based survey, this may yield different results from those drawn from a single survey data source.

We ask, what is the lifetime prevalence of abortion in this survey and does the lifetime abortion prevalence differ by age, race, education, income, marital status, pregnancy history, and current health insurance status? When possible, we also compare our estimates with Centers for Disease Control data (Jatlaoui, 2018), Guttmacher and American Community Survey data (Jones & Jerman, 2017), the Delaware Vital Statistics Annual Report (Delaware Health Statistics Center, 2020), and results from a pilot study using list experiment data

[3]One questionnaire contains treatment list A and control list B and the other contains control list A and treatment list B (see Table 1 for an example).

[4]Though those who are not cis-women can undergo abortions, we use the term women here to reflect the language of the data presented in the article.

(Cowan et al., 2016). Such sources allow us to evaluate the external validity of the list experiment for abortion data.

Methods

Data Source

The present study uses The Statewide Survey of Women of Reproductive Age in Delaware and Maryland (SWS) (Boudreaux & Rendall, 2020; see Steinberg et al., 2021 for a previous use of this survey). The SWS was administered to a probability sample of nearly 3000 women aged 18 to 44 in Delaware and Maryland from November 2016 through March 2017. The survey captures current and lifetime contraceptive use, as well as a range of other related information, such as beliefs and attitudes about pregnancy prevention and pregnancy intention; sexual activity; pregnancy history; abortion attitudes and experiences; and socio-demographic characteristics. The overall AAPOR response rate was 23.0% and 98% of started interviews were completed (National Opinion Research Center, 2019, p. 19). Post-stratification sample weights calibrated to Census Public Use Microdata enable estimates from the SWS to be representative of the Delaware and Maryland populations of women of reproductive ages (National Opinion Research Center, 2019, p. 19).

To measure the lifetime prevalence of abortion, the SWS included a double list experiment, modeled on prior studies (Bell & Bishai, 2019; Cowan et al., 2016; Elewonibi et al., 2021; Huber-Krum, Hackett, et al., 2020; Huber-Krum, Karadon, et al., 2020; Moseson et al., 2015, 2019, 2021; Moseson, Gerdts, et al., 2017; Moseson, Treleaven, et al., 2017). Respondents were randomly assigned to one of two versions of the SWS. Approximately half of respondents were assigned to receive Treatment List A and Control List B and half were assigned to receive Control List A and Treatment List B. Table 1 shows the list experiment items (control items were based on those used in Moseson et al. (2019)). Lists included two high prevalence control items and one low prevalence item to minimize the chances that respondents would have zero or all items apply to them, thus revealing whether they had or had not had an abortion (Moseson, Gerdts, et al., 2017; Moseson, Treleaven, et al., 2017). The survey was predominantly fielded via web with follow-up modes conducted by mail and telephone. We restricted our analytic sample to individuals who responded to both Lists A and B resulting in an analytic sample of N=2,747 (91.8% of the recruited sample).

Analysis

As a check that randomization was performed correctly, we examine the characteristics of respondents who received Treatment List A versus Treatment List B. Significant differences across groups, as assessed using a series of t-tests, would suggest a possible failure to randomize. To examine whether there is evidence that, for each list and mode of administration, list experiment assumptions are met, we use the Blair and Imai (Blair & Imai, 2012) design effect test. This test looks at the cumulative affirmative responses across treatment and control groups and uses a likelihood ratio test to determine if the observed pattern indicates a failure to reject no design effect. Following Blair and Imai (2012), a Bonferroni correction is applied to p-values resulting from this test which results in some loss of statistical power but is needed because directly testing the null hypothesis is problematic (Wolak, 1991). If the difference in cumulative proportions between treatment and control groups is always positive, design effect assumptions are met. However, any or all negative differences may suggest a violation of list experiment assumptions. Another assumption of the list experiment method is that respondents are honest about the number of items that apply to them. While we cannot test this assumption directly, looking at abortion prevalence by pregnancy history provides a useful falsification check. If respondents are reliable reporters of their reproductive history, we would expect that the difference in means between treatment and control groups for both lists will not be statistically different from 0 among individuals who report having never been pregnant.

To increase statistical power and follow past research (Bell & Bishai, 2019; Cowan et al., 2016) we pool treatment and control groups across lists and take the weighted difference in means between treatment and control groups to estimate the lifetime prevalence of abortion overall, by age (18–29, 30–44), by race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, and other race/ethnicity), by education (college graduate or not), by income (annual household income below $50,000 or at or above $50,000), by marital status (currently married, cohabiting, single), and by health insurance status (currently covered by Medicaid or any other insurance status). Differences between groups are ascertained using a series of t-tests. We focus on these social and demographic characteristics as past work has shown lifetime prevalence of abortion may vary by sociodemographic characteristics (Jatlaoui, 2018; Jones & Jerman, 2017).

To increase statistical power, in the main analysis, we use measures of age, race/ethnicity, education, marital status, and income that have been imputed by the survey via hotdeck imputation. While item non-response was generally low, less than 10% for most items including the abortion list experiment, 17.8% of respondents had missing income information (National Opinion Research Center, 2019, p. 28). In sensitivity analyses, we calculate prevalence for these subgroups when only unimputed measures are used. Another concern arises from respondents who are dropped in the main analysis due to having missing information on the list experiment questions. If respondents are differentially not responding because the sensitive item applies to them, we may be underestimating abortion prevalence. As a bounding exercise, we re-estimate the overall prevalence of abortion if all respondents with missing information in fact had an abortion (Manski, 2009). As another check on the robustness of our estimates, we perform an internal consistency check in which we estimate the effect of treatment separately for each list in a linear regression model and use a Wald test to determine if the treatment coefficients (capturing abortion prevalence) differ across lists (Lépine et al., 2020). If we fail to reject the null hypothesis, then we can confirm the internal consistency of abortion prevalence across lists. The Blair and Imai (2012) design effect test is performed on unweighted respondent data using the list package in R 4.02. All other analyses are done using appropriate survey weights using Stata Version 15.

Results

Table 2 shows weighted characteristics of the analytic sample overall and by treatment group. The sociodemographic characteristics of the weighted sample reflect respondents living in these two states; 55.0% of respondents in the sample are non-Hispanic White, 26.4% are non-Hispanic Black, 6.7% are Hispanic, and the remaining 11.9% are of some other race/ethnicity. 42.3% of respondents in the sample are under the age of 30 and 57.7% are between the ages of 30 and 44. There are no statistically significant differences in the characteristics of respondents across lists suggesting that randomization of the questionnaire was successfully implemented.

Next, we tested whether the design effect assumption is met for Lists A and B. Table 3A shows the cumulative proportion of affirmative responses for treatment and control groups in Lists A and B. In both lists, under 10% of respondents report that none or all of the list items apply to them suggesting a limited threat from floor or ceiling effects, individuals modifying their reporting to avoid disclosure of sensitive information (Kuklinski et al., 1997). For List A, the cumulative proportion is always greater for the treatment group than for the control group (Table 3A) suggesting that the assumption of the list experiment method is met (P=1.000). For List B, the proportion of respondents with at least one affirmative response is lower in the treatment group than in the control group (13.2% versus 15.7%), suggesting a possible violation of list experiment assumptions (P=.048). We also checked for design effects stratifying by survey mode. When we stratified by mode of administration, we did not find significant design effects for either list (P>.05 for all lists and modes) (Table 3B). As a final check that respondents appear reliable reporters of their reproductive history, we take the difference in means between treatment and control groups for Lists A and B by whether respondents had ever been pregnant. In both lists, the estimated prevalence of abortion is not statistically different from 0 among respondents who report having never been pregnant (Appendix Table 2A. List A 1.6%; 95% CI: −12.0%, 15.9%; List B 0.7%; 95% CI: −12.2%, 13.7%) and is positive among respondents who report ever having been pregnant (Appendix Table 2. List A 35.9%; 95% CI: 25.9%, 45.8%; List B 26.6%; 95% CI: 16.5%, 36.7%).

Given results from these tests, we then proceeded to pool estimates across lists and take the difference in means between treatment and control groups to obtain a lifetime prevalence of abortion (Table 4). For the overall sample, 21.0% (95% CI: 16.8%, 25.3%) of respondents ever experienced an abortion. The lifetime prevalence of abortion varies significantly by age, race, education, income, marital status, and current health insurance status. Respondents under 30 are less likely to have had an abortion (p<.05; 14.1%; 95% CI: 7.2%, 21.8%) than respondents 30 and older (26.6%; 95% CI: 20.8%−30.9%). Compared to all other race/ethnic groups, Non-Hispanic Black respondents are most likely to have had an abortion (p<.001; 37.0%; 95% CI: 27.1%, 46.8%) and Non-Hispanic White respondents are least likely to have had an abortion (p<.01; 14.5%; 95% CI: 9.5%, 19.5%). Hispanic respondents have a lifetime prevalence of abortion between White and Black respondents (24.2%; 95% CI: 5.3%, 43.0%) but due to small sample size we are unable to detect whether this estimate differs significantly from other groups. We found that respondents with a college degree have a lower lifetime prevalence of abortion (p<.05; 15.4%; 95% CI: 11.4%, 19.4%) than respondents with less education (24.8%; 95% CI: 18.3%, 31.3%). Abortion prevalence also varies by household income. Respondents living in households with incomes above $50,000 are less likely to have had an abortion (p<.05; 17.3%; 95% CI: 12.8%, 21.7%) than those living in households with incomes below $50,000 (28.6%; 95% CI: 19.7%, 37.5%). Respondents currently cohabiting are more likely to have had an abortion (p<.001; 39.0%; 95% CI: 29.1%, 48.9%) when compared with those currently married (17.1%; 95% CI: 11.6%, 22.7%) or single (16.1%; 95% CI: 8.3%, 23.8%). Consistent with our findings that respondents living in low income households are more likely to have had an abortion, respondents currently covered by Medicaid insurance are substantially more likely to have had an abortion (p<.01; 42.8%; 95% CI: 27.6%, 58.0%) than respondents with any other insurance status (15.1%; 95% CI: 10.3%, 20.0%).

We perform a number of supplemental analyses to check the robustness of findings. First, as measures of age, race, education, marital status, and income were imputed, we show prevalence results excluding imputed values (Appendix Table A.1). In our bounding exercise where we re-calculate the prevalence of abortion assuming all respondents with missing information on the list experiment in fact had an abortion, our abortion prevalence rises from 21.0% to 27.6%. Finally, as a check on the internal consistency of estimates, we present the prevalence for the overall sample and for subgroups using each list separately (Appendix Table A.2). The Wald test on the estimate of treatment across lists (here capturing abortion prevalence) suggests a failure to reject the null hypothesis that the treatment is the same across lists for the overall sample and for all but one of our subgroup analyses. We found that prevalence across lists differs for college graduates. Given that we made 18 comparisons, these findings suggest that internal consistency of lists is generally good.

Discussion

Research is beginning to establish the use of list experiments in US abortion research. While prior work uses convenience samples (Cowan et al., 2016) or population-representative samples from international contexts (Bell & Bishai, 2019; Moseson et al., 2015; Moseson, Gerdts, et al., 2017), this study is the first to estimate the lifetime prevalence of abortion using a list experiment administered to a representative sample of women living in the United States (specifically Maryland and Delaware).

We estimated the overall prevalence of abortion among all women 18–44 at 21%, like the 22% estimate found by Cowan et al. (2016) from their convenience sample. The list experiment approach yielded substantially higher estimates than direct question methods used in nationally-representative surveys. According to Lindberg et al. (2020), the NLSY97 and Add Health surveys, which both ask respondents about their abortion history, estimate that just 4.6% and 6.1% (respectively) had an abortion. These estimates miss over two thirds of abortions expected from estimates derived from abortion provider surveys. For example, the NSFG estimated 8,272,507 abortions (weighted) during a five year recall period. while the Abortion Provider Census estimated 11,413,954 during that same period (Lindberg et al., 2020), a 72% difference between the NSFG and APC counts. Further, our evidence indicates reliability in respondents' reporting of reproductive histories. Only among women who report ever having been pregnant is the difference between list experiment treatment and control groups statistically different from 0. That respondents who report having never been pregnant are also not classified as having an abortion has two possible interpretations. One possibility is that respondents are responding honestly and consistently that they have never had an abortion and have never been pregnant. Another is that respondents whose only pregnancy ended in abortion are reporting neither the abortion in the list experiment nor the pregnancy in later survey questions. Understanding reporting patterns of abortion and pregnancy history is an important area for future research.

However, our findings on abortion prevalence differentials by age, race, marital status, education, and income align with those from other surveys and administrative data sources. Past estimates combining abortion-clinic survey data with other data sources find that younger individuals are less likely to have had an abortion during their lifetimes (Jones & Jerman, 2017). Our results also align with annual abortion surveillance from the Centers for Disease Control (Jatlaoui, 2018) and the Delaware Vital Statistics Report (Delaware Health Statistics Center, 2020), finding that White individuals are the least likely to have had an abortion and Black individuals are most likely to have had an abortion. Consistent with Jones and Jerman (2017), we also find that college graduates and those living in households with higher incomes are the least likely to have had an abortion. Together, our findings are consistent with survey and administrative sources suggesting enduring disparities in abortion (Dehlendorf et al., 2013).

The similarities between our results and administrative estimates (Delaware Health Statistics Center, 2020; Jatlaoui, 2018; Jones & Jerman, 2017), as well as the work of other scholars (Cowan et al., 2016; Moseson et al., 2019; Moseson, Treleaven, et al., 2017), suggest that list experiments are a promising method for estimating abortion prevalence in the United States. However, there are a few limitations of note. First, we found possible design effects for List B when looking at the full sample—follow up work should investigate whether this effect is systematic or the product of random chance. Additionally, confidence intervals surrounding our estimates are rather large, particularly for the subgroup analyses. While we gain statistical power by pooling across lists, precision could be improved by administering these questions to a larger sample of women. Future work should also examine whether the list experiment method could be used to assess trends in abortion over time and by state, given differences in abortion trends over the last five years (wherein Delaware has substantially declined, and Maryland has slightly increased) (*State Facts About Abortion: Delaware*, 2020; *State Facts About Abortion: Maryland*, 2020). Finally, unlike some prior studies (Bell & Bishai, 2019; Cowan et al., 2016), the SWS did not directly ask about abortion thus we cannot directly compare the direct question and list experiment approaches.

Implications for Practice and Policy

Policy makers and practitioners, such as health departments, need accurate abortion estimates to ensure sufficient access to reproductive health services and identify subgroups who may need expanded contraceptive or abortion access. Improved survey data are of particular importance to those working at the state level, where most abortion regulation occurs and administrative records may be limited. Further, understanding how common abortion is important to reducing abortion stigma.

Conclusions

We conclude that list experiments appear a promising way to estimate the prevalence of abortion in a US context yielding estimates substantially higher than direct question approaches. Further, mode of administration does not appear to impact the results of a list experiment. However, more work is needed on the study of design effects and handling non-response to list experiment questions. The use of a list experiment in US survey data may improve our measurement of abortion prevalence while protecting respondent privacy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

1

Add Health	National Longitudinal Survey of Adolescent to Adult Health
CDC	The Centers for Disease Control and Prevention
HH	household
NLSY97	National Longitudinal Survey of Youth, 1997 cohort
NSFG	National Survey of Family Growth
SWS	The Statewide Survey of Women of Reproductive Age in Delaware and Maryland

Bibliography

Bell SO, & Bishai D. (2019). Can a List Experiment Improve Validity of Abortion Measurement? Studies in Family Planning, 50(1), 43–61. 10.1111/sifp.12082 [PubMed: 30675727]

Blair G, & Imai K. (2012). Statistical Analysis of List Experiments. Political Analysis, 20(1), 47–77. JSTOR.

Boudreaux M, & Rendall MS (2020). Statewide Survey of Women of Reproductive Age in Delaware and Maryland (“DELAWARE / MARYLAND SURVEY OF WOMEN”) [computer file] [Data set]. College Park, MD: University of Maryland- Collge Park. https://popcenter.umd.edu/delcaneval/survey

Choi YS, Rendall MS, Boudreaux M, & Roby DH (2020, June 4). Summary of the Delaware Contraceptive Access Now (DelCAN) Initiative. DelCAN Initiative Evaluation. https://popcenter.umd.edu/delcaneval/summary-init

Cowan SK, Wu LL, Makela S, & England P. (2016). Alternative Estimates of Lifetime Prevalence Of Abortion from Indirect Survey Questioning Methods. Perspectives on Sexual and Reproductive Health, 48(4), 229–234. 10.1363/48e11216 [PubMed: 27513590]

Dehlendorf C, Harris LH, & Weitz TA (2013). Disparities in abortion rates: a public health approach. American Journal of Public Health, 103(10), 1772–9. 10.2105/AJPH.2013.301339Delaware [PubMed: 23948010]

Health Statistics Center. (2020). Delaware Vital Statistics Annual Report, 2018. Delaware Department of Health and Social Services, Division of Public Health. https://www.dhss.delaware.gov/dhss/dph/hp/files/rept_preg18.pdf

Elewonibi B, Amour C, Gleason S, Msuya S, Canning D, & Shah I. (2021). Estimating the lifetime incidence of induced abortion and understanding abortion practices in a Northeastern Tanzania community through a household survey. Contraception, 103(2), 127–131. [PubMed: 33098850]

Gerdts C, & Ahern J. (2016). Multiple-bias analysis as a technique to address systematic error in measures of abortion-related mortality. Population Health Metrics, 14(1), 9. 10.1186/s12963-016-0075-3 [PubMed: 27006645]

Ghanem KG, Hutton HE, Zenilman JM, Zimba R, & Erbelding EJ (2005). Audio computer assisted self interview and face to face interview modes in assessing response bias among STD clinic patients. Sexually Transmitted Infections, 81(5), 421–425. 10.1136/sti.2004.013193 [PubMed: 16199744]

Glynn AN (2013). What Can We Learn with Statistical Truth Serum?Design and Analysis of the List Experiment. Public Opinion Quarterly, 77(S1), 159–172. 10.1093/poq/nfs070

Gonzalez-Ocantos E, De Jonge CK, Meléndez C, Osorio J, & Nickerson DW (2012). Vote buying and social desirability bias: Experimental evidence from Nicaragua. American Journal of Political Science, 56(1), 202–217. 10.1111/j.1540-5907.2011.00540.x

Huber-Krum S, Karadon D, Kurutas S, Rohr J, Baykal SS, Okcuoglu BA, ... & Shah I. (2020). Estimating abortion prevalence and understanding perspectives of community leaders and providers: Results from a mixed-method study in Istanbul, Turkey. Women's Health, 16, 1745506520953353. 10.1177/1745506520953353

Huber-Krum S, Hackett K, Kaur N, Nausheen S, Soofi S, Canning D, & Shah I. (2020). An Application of the List Experiment to Estimate Abortion Prevalence in Karachi, Pakistan. International Perspectives on Sexual and Reproductive Health, 46(Supplement 1), 13–24. 10.1363/46e0520 [PubMed: 33326396]

Jatlaoui TC (2018). Abortion Surveillance—United States, 2015. MMWR. Surveillance Summaries, 67. 10.15585/mmwr.ss6713a1

Jaya Hindin, M. J., & Ahmed S. (2008). Differences in Young People's Reports of Sexual Behaviors According to Interview Methodology: A Randomized Trial in India. American Journal of Public Health, 98(1), 169–174. 10.2105/AJPH.2006.099937 [PubMed: 18160677]

Jones RK, & Jerman J. (2017). Population Group Abortion Rates and Lifetime Incidence of Abortion: United States, 2008–2014. American Journal of Public Health, 107(12), 1904–1909. 10.2105/AJPH.2017.304042 [PubMed: 29048970]

Jones RK, & Kost K. (2007). Underreporting of Induced and Spontaneous Abortion in the United States: An Analysis of the 2002 National Survey of Family Growth. Studies in Family Planning, 38(3), 187–197. 10.1111/j.1728-4465.2007.00130.x [PubMed: 17933292]

Kost K. (2015). Unintended Pregnancy Rates at the State Level: Estimates for 2010 and Trends Since 2002 (p. 19). Guttmacher Institute.

Kuklinski JH, Cobb MD, & Gilens M. (1997). Racial Attitudes and the "New South." New South, 59(2), 27.

Lépine A, Treibich C, & d'Exelle B. (2020). Nothing but the truth: Consistency and efficiency of the list experiment method for the measurement of sensitive health behaviours. Social Science & Medicine, 113326. 10.1016/j.socscimed.2020.113326

Lindberg L, Kost K, Maddow-Zimet I, Desai S, & Zolna M. (2020). Abortion Reporting in the United States: An Assessment of Three National Fertility Surveys. Demography, 57(3), 899–925. 10.1007/s13524-020-00886-4 [PubMed: 32458318]

Manski CF (2007).Identification for Prediction and Decision. Cambridege,MA: Harvard University Press.

Mensch BS, Hewett PC, Gregory R, & Helleringer S. (2008). Sexual Behavior and STI/HIV Status Among Adolescents in Rural Malawi: An Evaluation of the Effect of Interview Mode on Reporting. Studies in Family Planning, 39(4), 321–334. [PubMed: 19248718]

Moseson H, Filippa S, Baum SE, Gerdts C, & Grossman D. (2019). Reducing underreporting of stigmatized pregnancy outcomes: Results from a mixed-methods study of self-managed abortion in Texas using the list-experiment method. BMC Women's Health, 19(1), 113. 10.1186/s12905-019-0812-4 [PubMed: 31481033]

Moseson H, Gerdts C, Dehlendorf C, Hiatt RA, & Vittinghoff E. (2017). Multivariable regression analysis of list experiment data on abortion: Results from a large, randomly-selected population based study in Liberia. Population Health Metrics, 15(1), 40. 10.1186/s12963-017-0157-x [PubMed: 29268794]

Moseson H, Jayaweera R, Huber-Krum S, Garver S, Norris A, & Gerdts C. (2021). Reducing underreporting of abortion in surveys: Results from two test applications of the list experiment method in Malawi and Senegal. PloS one, 16(3), e0247201. 10.1371/journal.pone.0247201

Moseson H, Massaquoi M, Dehlendorf C, Bawo L, Dahn B, Zolia Y, ... & Gerdts, C. (2015). Reducing under-reporting of stigmatized health events using the List Experiment: results from a randomized, population-based study of abortion in Liberia. International Journal of Epidemiology, 44(6), 1951–1958. [PubMed: 26342584]

Moseson H, Treleaven E, Gerdts C, & Diamond-Smith N. (2017). The List Experiment for Measuring Abortion: What We Know and What We Need. Studies in Family Planning, 48(4), 397–405. 10.1111/sifp.12042 [PubMed: 29148056]

NORC. (2019) Delaware Contraceptive Access Now (DelCAN): Baseline Survey (Methodology report). Report prepared for The University of Maryland- College Park, College Park, MD. https://popcenter.umd.edu/delcaneval/files/SoW_Baseline_Survey_Methodology_202005

Norris A, Bessett D, Steinberg JR, Kavanaugh ML, De Zordo S, & Becker D. (2011). Abortion Stigma: A Reconceptualization of Constituents, Causes, and Consequences. Women's Health Issues, 21(3), S49–S54. 10.1016/j.whi.2011.02.010 [PubMed: 21530840]

Roberts SCM, Fuentes L, Berglas NF, & Dennis AJ (2017). A 21st-Century Public Health Approach to Abortion. American Journal of Public Health, 107(12), 1878–1882. 10.2105/AJPH.2017.304068 [PubMed: 29048963]

Rossier C. (2003). Estimating Induced Abortion Rates: A Review. Studies in Family Planning, 34(2), 87–102. 10.1111/j.1728-4465.2003.00087.x [PubMed: 12889341]

Sedgh G, & Keogh SC (2019). Novel approaches to estimating abortion incidence. Reproductive health, 16(1), 1–10. 10.1186/s12978-019-0702-0 [PubMed: 30621726]

Singh S, Remez L, & Tartaglione A. (2010). Methodologies for estimating abortion incidence and abortion-related morbidity: A review. Guttmacher Institute. http://clacaidigital.info:8080/xmlui/handle/123456789/448

Smith JC, & Cates W. (1978). The public health need for abortion statistics. Public Health Reports, 93(2), 194–197. [PubMed: 635096]

State Facts About Abortion: Delaware (p. 1). (2020). Guttmacher Institute.

State Facts About Abortion: Maryland (p. 1). (2020). Guttmacher Institute.

Steinberg JR, Marthey D, Xie L. & Boudreaux M. (2021). "Contraceptive method type and satisfaction, confidence in use, and switching intentions." Contraception. Published online.

Sully E, Giorgio M, & Anjur-Dietrich S. (2020). Estimating abortion incidence using the network scale-up method. Demographic Research, 43, 1651–1684. 10.4054/DemRes.2020.43.56

Tourangeau R, & Plewes TJ (Eds.). (2013). Nonresponse in Social Science Surveys: A Research Agenda. The National Academies Press. 10.17226/18293

Treleaven E, Pham TN, Nguyen AD, & Diamond-Smith N. (2021). The list experiment: An approach to measuring stigmatized behaviours related to sex-selective abortion. Population Studies, 1–18. 10.1080/00324728.2021.1900589

Wolak FA (1991). The local nature of hypothesis tests involving inequality constraints in nonlinear models. Econometrica: Journal of the Econometric Society, 981–995.

Yeatman S, & Trinitapoli J. (2011). Best-Friend Reports: A Tool for Measuring the Prevalence of Sensitive Behaviors. American Journal of Public Health, 101(9), 1666–1667. 10.2105/AJPH.2011.300194 [PubMed: 21778489]

Table 1.

Double list experiment items

On the following list of health experiences, how many of these have you personally experienced? You don't need to say which ones, just how many.	
Version A- Control	• Ever used or taken medication for which a prescription is needed • Ever had a pap smear • Diagnosed with breast cancer in the past 10 years
Version A- Treatment	• Ever had an abortion (ended a pregnancy on purpose) • Ever used or taken medication for which a prescription is needed • Ever had a pap smear • Diagnosed with breast cancer in the past 10 years
Version B- Control	• Ever used a birth control method (such as: pills, an IUD or implant, condoms or the shot) • Had a tubal or ectopic pregnancy in the past year • Ever had your blood pressure measured
Version B- Treatment	• Ever had an abortion (ended a pregnancy on purpose) • Ever used a birth control method (such as: pills, an IUD or implant, condoms or the shot) • Had a tubal or ectopic pregnancy in the past year • Ever had your blood pressure measured

Table 2.

Characteristics of Respondents Overall and by Treatment Group

	Received Treatment A N=1359		Received Treatment B N=1388		Overall Sample N=2747	
	No	%[d]	No	%[d]	No	%[d]
Delaware	681	48.3	717	53.3	1398	50.9
Maryland	678	51.7	671	46.7	1349	49.1
African American Non-Latino	285	27.2	283	25.6	568	26.4
Asian Non-Latino	100	7.5	78	6.7	178	7.1
Latino	83	6.8	89	6.6	172	6.7
Multiple/Other	48	5.0	53	4.6	101	4.8
White Non-Latino	843	53.5	885	56.5	1728	55.0
18–24	192	21.7	229	25.6	421	23.7
25–29	259	18.5	264	18.7	523	18.6
30–34	284	20.5	304	20.3	588	20.4
35–39	312	19.6	288	17.2	600	18.4
40–44	312	19.6	303	18.3	615	18.9
LHS	31	4.2	35	4.0	66	4.1
HS	142	14.7	141	14.0	283	14.3
SCO	378	40.8	413	42.4	791	41.6
BA+	808	40.3	799	39.6	1607	40.0
HH Income Below 50000	407	32.1	480	34.5	887	33.3
HH Income At or Above 50000	952	67.9	908	65.5	1860	66.7
Not Currently Covered by Medicaid[a]	832	79.1	854	80.7	1686	79.9
Currently Covered by Medicaid[a]	148	20.9	152	19.3	300	20.1
Married[b]	678	42.1	686	43.1	1364	42.6
Cohabiting[b]	259	22.5	221	18.2	480	20.3
Single[b]	400	35.5	464	38.6	864	37.1
Never Been Pregnant[b]	439	35.2	510	40.2	949	37.7
Ever Been Pregnant[b]	843	64.8	791	59.8	1634	62.3
Self Administered Paper	518	40.9	524	38.4	1042	39.6
Web or Computer Assisted Telephone Interview[c]	841	59.1	864	61.6	1705	60.4

[a.] Only asked of women not pregnant or trying to get pregnant

[b.] Counts will not match total in sample due to item missingness

[c.] Counts for web and computer assisted telephone interview collapsed due to small sample size (unweighted n<10) of computer assisted telephone interview

[d.] Calculated percentages use normalized survey weights

Table 3A.

List experiment estimates of lifetime experience of abortion among Delaware and Maryland women aged 18–44 using the piecewise estimator

List A					
Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**
Proportion of Treatment Group Responding	0.035	0.118	0.650	0.176	0.022
Proportion of Treatment Group who Respond at Least	1.000	0.965	0.848	0.198	0.022
Proportion of Control Group Responding	0.052	0.137	0.764	0.048	0.000
Proportion of Control Group who Respond at Least	1.000	0.948	0.811	0.048	0.000
Treatment – Control Group Cumulative Response	0.000	0.017	0.037	0.150	0.022
Bonferroni Corrected P Value 1					
List B					
Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**
Proportion of Treatment Group Responding	0.035	0.132	0.628	0.173	0.032
Proportion of Treatment Group who Respond at Least	1	0.965	0.834	0.205	0.032
Proportion of Control Group Responding	0.022	0.157	0.765	0.055	0
Proportion of Control Group who Respond at Least	1	0.978	0.82	0.055	0
Treatment – Control Group Cumulative Response	0	−0.013	0.014	0.150	0.032
Bonferroni Corrected P Value 0.048					

Table 3B.

List experiment estimates of lifetime experience of abortion among Delaware and Maryland women aged 18–44 using the piecewise estimator, by survey mode

List A Web

Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**
Proportion of Treatment Group Responding	0.041	0.102	0.67	0.171	0.017
Proportion of Treatment Group who Respond at Least	1.000	0.959	0.858	0.188	0.017
Proportion of Control Group Responding	0.059	0.142	0.761	0.038	0.000
Proportion of Control Group who Respond at Least	1.000	0.941	0.799	0.038	0.000
Treatment – Control Group Cumulative Response	0.000	0.018	0.059	0.150	0.017
Bonferroni Corrected P Value 1					

List A Self Administered Paper

Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**
Proportion of Treatment Group Responding	0.025	0.139	0.622	0.183	0.031
Proportion of Treatment Group who Respond at Least	1.000	0.975	0.836	0.214	0.031
Proportion of Control Group Responding	0.04	0.128	0.769	0.063	0.000
Proportion of Control Group who Respond at Least	1.000	0.96	0.832	0.063	0.000
Treatment – Control Group Cumulative Response	0.000	0.015	0.004	0.151	0.031
Bonferroni Corrected P Value 1					

List B Web

Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**
Proportion of Treatment Group Responding	0.036	0.144	0.623	0.162	0.035
Proportion of Treatment Group who Respond at Least	1.000	0.964	0.820	0.197	0.035
Proportion of Control Group Responding	0.023	0.154	0.776	0.047	0.000
Proportion of Control Group who Respond at Least	1.000	0.977	0.823	0.047	0.000
Treatment – Control Group Cumulative Response	0.000	−0.013	−0.003	0.150	0.035
Bonferroni Corrected P Value 0.207					

List B Web					
Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**

List B Self Administered Paperh					
Group	**Response to List Experiment**				
	0	**1**	**2**	**3**	**4**
Proportion of Treatment Group Responding	0.032	0.111	0.637	0.191	0.029
Proportion of Treatment Group who Respond at Least	1.000	0.968	0.857	0.219	0.029
Proportion of Control Group Responding	0.017	0.164	0.753	0.066	0.000
Proportion of Control Group who Respond at Least	1.000	0.983	0.819	0.066	0.000
Treatment – Control Group Cumulative Response	0.000	−0.015	0.038	0.153	0.029
Bonferroni Corrected P Value 0.119					

Table 4.

Estimate of lifetime experience of abortion among women from Delaware and Maryland aged 18–44, by sociodemographic characteristics

Group	Estimated Prevalence	95% CI		
Overall	21.0	16.8	,	25.3
Non-Hispanic White	14.5	9.5	,	19.5
Non-Hispanic Black	37.0	27.1	,	46.8
Hispanic	24.2	5.3	,	43.0
Other	14.3	2.8	,	25.7
HH Income at Least 50,000	17.3	12.8	,	21.7
HH Income Below 50,000	28.6	19.7	,	37.5
Under Age 30	14.5	7.2	,	21.8
Over Age 30	25.8	20.8	,	30.9
Less than College	24.8	18.3	,	31.3
College Graduate	15.4	11.4	,	19.4
Currently Covered by Medicaid	42.8	27.6	,	58.0
Not Currently Covered by Medicaid	15.1	10.3	,	20.0
Currently Married	17.1	11.6	,	22.7
Currently Cohabiting	39.0	29.1	,	48.9
Single	16.1	8.3	,	23.8
Never Been Pregnant	2.5	−4.4	,	9.3
Ever Been Pregnant	31.2	25.7	,	36.7