Published in final edited form as:

Int J Soc Res Methodol. 2016 March 1; 19(2): 205–222. doi:10.1080/13645579.2014.979717.

Effects of intensive longitudinal data collection on pregnancy and contraceptive use

Jennifer S. Barber^a, Heather H. Gatny^{a,*}, Yasamin Kusunoki^a, and Paul Schulz^a

^a Institute for Social Research, University of Michigan, Ann Arbor, MI USA

Abstract

We use an experiment based on the Relationship Dynamics and Social Life (RDSL) study of sexual behavior, pregnancy, and contraceptive use to investigate whether participation in weekly self-reports over one year result in differential effects on related behaviors and attitudes, compared to participation in a one-time follow-up interview requiring retrospective self-reporting for the prior year. We randomly assigned 200 subjects to either a control group or a journal group. All subjects were interviewed at the beginning of the study (baseline interview) and 12 months later (closeout interview). Subjects in the journal group also completed a five-minute web- or phone-based survey every week during the 12-month study period. We found no statistically significant difference in pregnancy rates across the two groups at closeout. Contrary to our expectation, the control group experienced a slightly larger increase in having ever used a contraceptive method, although this was mainly due to increases in the least effective methods, such as condoms and withdrawal. Respondents in the weekly interview group became slightly more positive toward childbearing in terms of one specific attitude measure, but not for the vast majority of measures. We conclude that intensive longitudinal data collection does not appear to have a large or consistent impact on respondents' pregnancy, contraceptive use, or related attitudes, relative to the more standard longitudinal approach.

Keywords

social research; survey research; panel conditioning; diaries; intensive longitudinal data collection

1. Introduction

In the U.S., surveys measuring sexual behavior, contraceptive use, and pregnancy are used to generate important scientific information about levels and trends in these behaviors. In most such surveys, respondents self-report on these behaviors over lengthy retrospective periods. For example, the National Survey of Family Growth (NSFG) uses a life history calendar-based approach, asking respondents to recall the specific months over the past three years when they did not have sex, and the specific contraceptive method(s) used each month over the same period. The concern is that this kind of long-term retrospective reporting is subject to recall error. An alternative is to collect data via frequent, brief surveys (sometimes

* Corresponding author: Heather Gatny, Institute for Social Research, 426 Thompson St., Ann Arbor, MI 48106, USA. Phone: 734-615-3560. h gatny@umich.edu..

called “journals” or “diaries”). For example, the Relationship Dynamics and Social Life (RDSL) study – which we analyze here – collected *weekly* surveys from 18- and 19-year-old women for 2.5 years, shortening the period of recall to one *week*.

Diary-based methods of data collection have been used by social scientists to measure a wide variety of frequent behaviors such as alcohol consumption (Lemmens et al. 1988); food consumption (Stephen 2007); medication adherence (Johnson and Bytheway 2001); sexual behavior (Baer et al. 2002; Coxon 1999; Garry et al. 2002; Jaccard et al. 2002; Searles et al. 2000); and travel behavior (Axhausen et al. 2002). Schroder et al. (2003) review the advantages and disadvantages of diaries for measurement of sexual behaviors.

One obvious risk of methods with such frequent measurement is the potential for the data collection process itself to affect behavior, particularly the behaviors under study. The idea that participation in a study can itself change respondents’ behaviors is not new – the “Hawthorne effect” has been discussed for at least 50 years (French 1950, 1953), and “panel conditioning” is also a significant concern in survey research (Warren and Halpern-Manners 2012, Cantor 2008, Sturgis et al. 2009, and Waterton and Lievesley 1989). To the extent that participation in frequently repeated interviews influences respondents’ actual or reported behaviors/attitudes, it also introduces measurement error. Information about the effect of frequent interviewing on respondents has become increasingly important as the potential for nearly real-time data collection via mobile devices increases researchers’ interest in intensive longitudinal data collection.

Here we explore the potential influence of panel conditioning using an empirical example of collecting longitudinal panel data on contraceptive use and pregnancy. We examine how two different approaches – a one-time follow-up survey requiring retrospective recall over 12 months versus brief weekly surveys over the same 12 months – may influence self-reports of behaviors and attitudes.

2. Repeated measurement effects

Repeated survey measurement requiring self-reporting may affect behavior in several ways. (For excellent review articles, see Warren and Halpern-Manners 2012, Cantor 2008, Sturgis et al. 2009, and Waterton and Lievesley 1989.) Below, we focus our discussion on two concerns: the extent to which panel conditioning may result in *actual behavioral change*, and the extent to which it may change *reporting behavior*, independent of actual behavioral change. We use the example of the RDSL data collection on contraceptive use and pregnancy as a basis of discussion.

2.1 Changes in behaviors/attitudes

Repeated survey questions may lead to *actual* behavioral/attitudinal adaptations because respondents may experience changes in motivation toward, knowledge of, or cognitive attention to the measured subject over the course of the survey.

Changes in motivation may be fueled by cognitive dissonance – the uncomfortable feeling that individuals experience when they simultaneously hold conflicting attitudes, or when

their behavior conflicts with one or more of their attitudes (Festinger 1957). For example, if a young woman who does not desire pregnancy repeatedly reports sexual activity in the absence of contraception, this may force her to confront that dissonance, and may increase her motivation to bring her behavior in line with her stated desire. On the other hand, it may be easier to change attitudes than behavior, so it could instead force a change in attitude, which would also bring behavior and desire into alignment.

Even without cognitive dissonance at work, the simple act of asking about contraceptive use could spur its use. Some evidence suggests that asking young people about specific behaviors – both healthy and unhealthy – may actually increase those behaviors. For example, college students asked about exercise reported higher rates of exercise, and those asked about illegal drug use reported higher rates of illegal drug use, two months after the initial survey (Williams et al. 2006).

Research across a range of domains – from marital satisfaction to buying a car – suggests that survey questions may also impact behavior or attitudes by virtue of imparting new information to respondents or stimulating their thinking about the interview topics (Crespi 1948). Recursive questioning may only underline this impact. So, for example, repeatedly asking respondents about their contraceptive use by presenting a list of specific contraceptive options from which to choose may alert them to methods of they were not previously considering, which could alter either their attitude toward trying a new method or their actual contraceptive use. Previous social science research suggests that attitudes not fully formed or strongly held may be particularly susceptible to these effects (see Schuman and Presser 1981; for consumer research, see Kardes et al. 1993).

2.2 Changes in reporting

Even if repeated measurement does not affect actual behaviors or attitudes, it may affect *reports* of behaviors/attitudes via social desirability bias – or the tendency to under-report socially undesirable behavior and over-report normatively good behavior. Questions about socially sensitive issues are especially susceptible to this bias, because respondents are more likely to self-report in a manner they believe the interviewers will view favorably (Crowne and Marlowe 1964; Moorman and Podsakoff 1992; Paulhus 1984). This can be viewed as another form of cognitive dissonance: When respondents act or hold attitudes that conflict with societal norms, they may attempt to alleviate this dissonance by reporting actions/attitudes that are more normative. The risk of social desirability bias may increase when questions are repeated frequently.

Conversely, repeated interactions with the same interviewer, or even the same survey instrument, may increase respondents' comfort with the survey – for example, by reducing suspicions about interviewer bias or data confidentiality – and thus *reduce* the risk of social desirability bias (Fowler 1995; Nancarrow and Cartwright 2007; Tourangeau et al. 2000; Waterton and Lievesley 1989). In addition to varying by type of questions asked, the risk of social desirability bias may differ across social groups, contexts, and survey modes (e.g., web versus phone) (Warren and Halpern-Manners 2012). Social desirability bias may be weaker in web-based versus interviewer-administered surveys, and weaker in phone versus face-to-face surveys (Kreuter, Presser, and Tourangeau 2008). However, for young women

in the age range of the RDSL study sample – 18 to 22 years – having sex *without* contraception and becoming pregnant are almost universally perceived as highly socially undesirable behaviors. Thus repeated interactions would probably compound rather than reduce the risk of social desirability bias regardless of context or survey mode.

2.3 Closely spaced interviews

Evidence indicates that the risk of panel conditioning is greater when surveys are repeated at short as compared to long intervals (Bailar 1989; Warren and Halpern-Manners 2012; Zwane et al. 2011). For example, a study of the impact of using chlorine to improve water quality on childhood diarrhea randomly assigned households to biweekly surveys for approximately 9 months plus a follow-up survey 7 months later, or three surveys over the same period. The households in the frequently interviewed group had significantly higher levels of chlorine in their drinking water and reported significantly less diarrhea in their children (Zwane et al. 2011). The authors hypothesize that the biweekly interviews served as a reminder to chlorinate drinking water, even though the survey itself contained only three questions (in a 20- to 30-minute survey) on chlorine.

Panel conditioning bias may be more prevalent when interviews are longer. For example, a study randomly assigned newly married couples to an intensively interviewed group or a control group, and followed them for four years. In the intensively interviewed group, each spouse completed two 80-minute structured questionnaires with both open and fixed response options (years 1 and 3), two 15-minute survey interviews (years 2 and 4), and were interviewed together twice (years 1 and 3) for approximately 30 minutes. The control group completed nearly as many surveys, but they were briefer – the wife completed a 7-minute survey in year 1, and each spouse completed a 5-minute survey in year 2 and 15-minute survey in year 4. By the end of the study, the intensively interviewed group scored higher on several marital quality scales. Thus, we argue that the risk of panel conditioning bias in short-interval re-interviewing may be reduced when interviews are kept brief.

2.4 Hypotheses

Here we use an experiment based on the RDSL study of sexual behavior, pregnancy, and contraceptive use to investigate whether participation in weekly journal-type self-reports (the journal condition) affects related behaviors and attitudes, comparing to those found among participants in a one-time follow-up interview requiring retrospective self-reporting for the prior year (the control condition). We expect the following results, if panel conditioning is a factor in the study:

1. The journal group will report lower rates of pregnancy than the control group.
2. The journal group will report larger increases in contraceptive use than the control group.
3. The journal group will report more positive change in attitudes toward pregnancy and related behaviors than the control group.

3. Design of the experiment

The representative population-based sample of the RDSL study included 1,003 respondents who were age 18 or 19 and residing in a single Michigan county at the time of the baseline survey. The RDSL included weekly measures of prospective pregnancy desires, contraceptive use, and pregnancy, as well measures of relationship characteristics and behaviors such as commitment, conflict, and sexual activity, collected in a “journal” (a short, structured weekly survey interview). The overriding concern for the journal was to keep the interview short, minimizing the burden for respondents and thereby maximizing their participation.

An experiment to test for panel conditioning, distinct from the RDSL study itself, enrolled 200 women ages 18 and 19, who were randomly assigned to either (1) the journal group or (2) the control group. Because the main focus of the RDSL is early pregnancy – who gets pregnant and under what conditions – the experiment was designed to compare reports of contraceptive use and pregnancy experiences among participants who answered the follow-up questions weekly (journal group) and participants who had solely a one-time follow-up interview (control group).¹

Both groups received an initial 50-minute face-to-face interview using the RDSL baseline instrument. At the conclusion of the baseline interview, the journal group members were invited to enroll in the weekly follow-up surveys over the next 12 months, and all agreed to participate. The control group members were not invited to enroll in the weekly surveys, but they, along with the journal group members, were asked to participate in the close-out interview 12 months after the baseline interview. The close-out interview, conducted by phone using a subset of the questions in the baseline interview, had an overall 93% response rate (186 of the original 200 subjects were re-interviewed).

Both groups were paid \$35 for participation in the baseline interview. Respondents in the journal group were paid \$1 per weekly journal with \$5 bonuses for on-time completion of five weekly journals in a row. Journal incentives were distributed via reloadable cash cards. These respondents also received small tokens of appreciation for their continued participation in the journal, such as a pen, compact, lip balm, etc. They were allowed to complete their interview either online or by calling a survey research phone lab. Both groups received \$35 for completing the closeout interview.

4. Measures

Sociodemographic characteristics of the sample for the larger RDSL study, control, and journal groups are reported in Table 1. Although the experimental sample (N=200) was similar to the RDSL main sample (N=1,003) on these characteristics and experiences, we found several key differences between the randomly assigned control and journal groups in the experimental sample. First, a higher proportion of the journal group was enrolled in high

¹Although we are interested in the general effects of panel conditioning on contraceptive use and pregnancy, it is important to note that this design does not permit us to assess the effects of repeatedly asking about contraception and pregnancy relative to repeatedly asking about something else.

school (22% versus 12% of the control group) and was living with a sexual partner (15% versus 7%). A smaller proportion of the journal group had a teen mother (28% versus 46% of the control group).

At the baseline interview, all respondents were asked about prior sexual behavior, prior pregnancies, and current/prior contraceptive use, as well as a series of Likert-scale questions assessing attitudes related to pregnancy. Respondents in the journal group were asked about pregnancy and contraceptive use in each weekly interview, referencing the prior week. The closeout interview included questions about prior pregnancy and current/prior contraceptive use, repeated from the baseline interview. Table 2 presents the survey questions for measures of sex, pregnancy, and contraceptive use. Table 5 presents the questions for the attitude measures.

5. Analytic strategy

We compare the control and journal groups in two ways. First, we use independent samples t-tests to indicate whether differences in the two groups are statistically unlikely ($p < .10$) to be due to random chance. Second, because the two randomly assigned groups differed in a few important ways, we also use weighted estimates of the differences in the *changes* experienced by the two groups. The weighted estimates in Table 3 are computed using the estimated propensity of being in the journal (versus control) group, based on the three sociodemographic variables that were statistically different in Table 1: being enrolled in high school, living with a sexual partner, and having a biological mother who was < 20 at her first birth. The weighted estimates in Table 4 are computed using the estimated propensity of being in the journal (versus control) group, based on those sociodemographic differences, the statistically significant differences in sex and pregnancy presented in Table 3, and the corresponding baseline level of the variable (i.e., prior pregnancy, ever used contraception, current contraceptive use, specific method use). In Table 5, weighted estimates also account for the initial (baseline) value of the attitude measure. We used `teffects psmatch` in Stata 13.

6. Baseline differences in sex, pregnancy and contraceptive use

Table 3 presents measures of sex, pregnancy, and contraceptive use in the journal and control groups *prior* to the study period, as reported in the baseline interview. On these measures we found that journal group respondents had slightly more partners and were more likely to have had a prior pregnancy, on average. Although the proportion never pregnant was similar for the two groups, those with prior pregnancies in the journal group were split nearly evenly between one and two-plus prior pregnancies, while most of the women in the control group with a prior pregnancy had only one (not shown in tables).

It is important to emphasize that these differences between the journal and control groups are due solely to imperfect randomization, from random chance, and represent differences in a small number of respondents (e.g., 12 women in the journal group had two or more prior pregnancies, versus 5 in the control group). Overall, some of these differences predict a higher risk of pregnancy for the journal group, relative to the control group (higher proportion living with a sexual partner, more sexual partners, more prior pregnancies). Other

differences predict a higher risk of pregnancy for the control group (lower proportion still enrolled in high school, higher proportion with a teen mother).

Similar proportions in both groups reported *ever* using contraception (76% and 77%) and *currently* using any type of contraception (54% and 56%) on the baseline survey. In both groups, birth control pills were the most common contraceptive method, followed by other (unspecified)², Depo-Provera, IUD/implant, and the patch or ring. By random chance, a larger proportion of the journal group than the control group was using birth control pills (35% versus 26%). Additionally, although the difference is not statistically significant in this relatively small sample, the control group had a higher rate of using Depo-Provera. Note that these proportional differences translate into a difference of only 15 people for pill use (63 versus 48 respondents), and 7 people for Depo-Provera (13 versus 20). Nonetheless, it is important to account for these baseline differences when analyzing whether the weekly journal interviews increased (or decreased) overall contraceptive behavior as reported in the closeout interview.

7. Changes in pregnancy and contraceptive use

Table 4 presents *changes* in the control and journal groups, in terms of pregnancy and contraceptive use, over the course of the study. The first two columns refer to the change in proportion within each group. The next three columns are the unweighted differences between these two change estimates (indicating which group changed more), and the corresponding *p*-values and significance levels. The final three columns present the weighted (by estimated propensity of being in the journal group) estimate of the differences and the corresponding *p*-values and significance levels.

First, note that in terms of our primary outcome of concern – pregnancy – the journal group is not statistically significantly different from the control group: 15% of the control group (14 respondents) and 20% of the journal group (18 respondents) reported a pregnancy during the study period. This 5% difference is not in the hypothesized direction – that is, we expected that journal participation would reduce pregnancy rates because it would motivate some respondents to reduce the cognitive dissonance between reported desires to avoid pregnancy and irregular contraceptive use.

Contraceptive use increased for both groups during the 12 months of observation, which is consistent with contraceptive use increasing with age in the general population (Welti et al. 2011). The proportion of both groups that had *ever* used contraception increased (11% in the control group and 2% in the journal group), as did the proportion *currently* using contraception (19% in the control group and 9% in the journal group). For *ever* using contraception, the weighted and unweighted differences between the journal and control groups in these changes are statistically significant. For *current* contraceptive use, the weighted difference is significant but the unweighted difference is not. Differences between the control and journal groups in terms of non-users who began using a contraceptive

²Respondents who indicated that they did or used something to avoid pregnancy, but who indicated that they did not use birth control pills, the patch, the ring, Depo-Provera/other injectable, or an IUD were coded “other.” The most common of these methods were the “rhythm” (calendar) method, condoms, and withdrawal.

method, and switching to more effective methods were not statistically significant; however, those numbers indicate that a higher fraction of control group nonusers (versus journal group nonusers) began using some contraceptive method during the study period, but a higher fraction of journal group contraceptive users (versus control group contraceptive users) switched from a less to a more effective method during the study period.

Group-based differences for specific contraceptive methods are also explored in Table 3. Use of birth control pills, the most popular method in this age group, increased similarly in both groups during the study period. Although the unweighted difference between the journal group and the control group is quite small, the weighted difference is much larger because it accounts for the journal group's higher propensity to use the pill before the study began, and other factors that affect the risk of pregnancy but were not balanced in the two samples.

Changes in the proportion of respondents who were using “other” methods of contraception were large. These “other” method differences were opposite in sign for the two groups – increasing by 7% in the control group, but decreasing by 4% in the journal group – and statistically significant. Changes in other specific contraceptive methods were small and/or similar across the two groups, with no statistically significant unweighted or weighted differences.

The findings in Table 4 support neither hypothesis 1, which predicts that the journal group will report larger increases in contraceptive use than the control group, nor hypothesis 2, which predicts that the journal group will report lower rates of pregnancy. In neither case are the differences between the groups statistically significant, but they run counter to predictions for both hypotheses. There is one exception – respondents in the journal group experienced a decline in the use of the least effective methods (“other”), and the data suggest that they may have had larger increases in use of the most effective methods (particularly IUD/implant, and also pills).

8. Attitude change

Table 5 presents results for changes in attitudes over the course of the study for those in the control and journal groups. A change score was computed for each respondent for each attitude measure (score on the close-out interview minus score on the baseline interview). The numbers in Table 5 represent the means of these individual change scores, by group. Positive numbers indicate one of the following: an increase in agreement with one of the opinion statements (items 1, 2, 3, 5, 6, 7, 8, and 9), a more negative feeling about never marrying or never having children (items 4 and 10), or an increase in desired family size (item 11).

We found only two unweighted differences and one weighted difference between the control and journal groups that are statistically significant. For question 6, on average, respondents in the control group became more positive toward being a mother while respondents in the journal group become less positive. The weighted difference, however – which takes into account the extent to which the original sample was unbalanced in terms of the attitude measure, as well as key sociodemographic and behavioral measures – is smaller and is not significantly different from zero. For question 8, respondents in the control group came to

more strongly agree that waiting to have a baby may result in infertility, while those in the journal group came to more strongly disagree. Both the weighted and the unweighted differences are statistically greater than zero. This is not consistent with hypothesis 3, which predicted more positive attitude change toward childbearing in the journal group than in the control group. Further, one statistically significant difference in eleven tests does not represent a strong pattern of any kind.

9. Discussion

The journal group reported a slightly higher, but nonsignificant, rate of pregnancy than the control group, which runs counter to hypothesis 2. Although the results did not support hypothesis 1 either, given that the control group increased their use of contraception slightly more than the journal group did, we did find that the journal group reported greater increases in the most effective methods – IUD/Implanon and birth control pills, and the control group reported greater increases in the least effective methods (coital-dependent methods such as condoms and withdrawal). We did not find support for hypothesis 3 – respondents in the control group reported more positive attitudes toward motherhood and early childbearing, relative to the journal group, but only on one measure.

Overall, although there are some statistically significant differences between the control and journal groups, the results do not indicate that the frequent re-interviews have a clear impact on the behaviors and attitudes being monitored. The nonrandomness of the two groups in terms of pre-study pregnancy experiences and sociodemographics – especially the greater proportion of the journal group who reported having prior pregnancies and living with their sexual partners – was the surprise difference, and may have increased the journal group's risk of pregnancy during the study period, independent of the mode of survey participation.

9.1 Limitations

The experiment did not measure pregnancy desires, which may be particularly susceptible to influence by the frequent reporting of pregnancy-related behaviors. For example, if young women participating in the journal are reporting weekly on sexual intercourse without contraceptive use, they may report a stronger pregnancy desire, in order to rationalize their behavior. It is also possible that asking each week about relationships and pregnancy may stimulate desire for pregnancy.

The experiment did not force randomization on any variables at the time of enrollment, which would have produced a more powerful experiment – particularly randomization on key variables such as previous pregnancy and contraceptive use. Future experimental research on panel conditioning should do this, particularly when working with small samples.

The small sample size for this experiment also limits our ability to detect differences between the journal and control groups, particularly for a relatively rare event like teen pregnancy. The statistical power was only .24 at the 5% level for the observed level of difference in pregnancy rates (20% vs. 15%) between the two groups of 100 respondents each. The total sample for the experiment would have to be more than 1,400 total

respondents to increase the statistical power to 80% for the observed difference. The experiment was slightly better powered to detect differences in contraceptive use, which occur at a higher rate. For example, for a 71% versus 63% rate of current contraceptive use, the power is .33. However, although the experiment was under-powered, the *pattern* of differences between the two groups is not strong.

The narrow geographic focus (a single county in Michigan) of the experiment, which is the same as the main RDSL study, is also a limitation. Although neither sample is nationally representative, Michigan falls around the national median of cohabitation, marriage, age at first birth, completed family size, non-marital childbearing, and teenage childbearing (see Lesthaeghe & Neidert 2006). This is not, of course, to suggest that Michigan is representative of the nation, rather that it is not an outlier. In addition, the study includes only a small number of Latinas.

Collecting weekly measures among young women at high risk for pregnancy helps capture the dynamic nature of behaviors and attitudes associated with unintended pregnancy. The effectiveness of this method, however, is somewhat dependent upon the degree to which this survey mode influences responses over time – and in particular, the influence of panel conditioning. Ideally, future research on panel conditioning in similar surveys of pregnancy-related behaviors and attitudes would have larger and more diverse populations, increasing both the statistical power and the generalizability of results, and ultimately informing the design of future surveys.

9.2 Conclusions

As Warren and Halpern-Manners (2012) point out, this type of experiment does *not* address general, theoretically informed questions about panel conditioning. In addition, we cannot disentangle the consequences of a relatively high respondent burden, in the form of weekly interviews, from the consequences of repeated questions about sex, pregnancy desire, and contraceptive use. However, our results are particularly relevant to researchers who want to implement an intensive longitudinal data collection project focused on pregnancy or other family formation behaviors. Demographers are increasingly interested in studying dynamic processes – an interest served by the journal-type data collection method described here. We leave it to future research to investigate more nuanced consequences of repeated questions about pregnancy-related behaviors and attitudes. Our intent with this small-scale experiment was to motivate such research.

The results from our experiment – which indicate no discernable pattern of panel conditioning – allow us to conclude that the benefits of the frequent measurement strategy used in the RDSL are likely to outweigh the costs, at least in terms of potential effects on respondent behavior. The data produced using this method have many advantages for analyses of pregnancy and contraception. Weekly data provide the potential to closely track dynamic behaviors, such as condom use and sexual experiences, that are both important to unintended pregnancy rates and are not easily recollected after long periods of time. The data also provide the opportunity to examine sequential events, such as changes in relationships that precede changes in contraceptive method choice, which may offer additional insights into unintended pregnancy. And finally, weekly data permit dynamic

assessments of the associations among pregnancy-related attitudes and behaviors – assessments that can be used to inform family planning programs aimed at young women.

Acknowledgements

This research was supported by two grants from the National Institute of Child Health and Human Development (R01 HD050329, R01 HD050329-S1, PI Barber), a grant from the National Institute on Drug Abuse (R21 DA024186, PI Axinn), and a population center grant from the National Institute of Child Health and Human Development to the University of Michigan's Population Studies Center (R24 HD041028). The authors gratefully acknowledge the Survey Research Operations (SRO) unit at the Survey Research Center of the Institute for Social Research for their help with the data collection, particularly Vivienne Outlaw, Sharon Parker, and Meg Stephenson. The authors also gratefully acknowledge the intellectual contributions of other members of the original RDSL project team, including William Axinn, Mick Couper, and Steven Heeringa.

References

- Axhausen KW, Zimmermann A, Schonfelder S, Rindsfuser G, Haupt T. Observing the rhythms of daily life: A six-week travel diary. *Transportation*. 2002; 29:95–124.
- Baer A, Sariou S, Koutsky LA. Obtaining sensitive data through the web: An example of design and methods. *Epidemiology*. 2002; 13(6):640–645. [PubMed: 12410004]
- Bailar, BA. Information needs, surveys, and measurement errors.. In: Kasprzyk, D.; Duncan, GJ.; Kalton, G.; Singh, MP., editors. *Panel Surveys*. Wiley; New York: 1989. p. 1-24.
- Cantor, D. A review and summary of studies on panel conditioning.. In: Menard, S., editor. *Handbook of longitudinal research: Design, measurement, and analysis*. Academic Press; Burlington, MA: 2008. p. 123-138.
- Coxon APM. Parallel accounts? Discrepancies between self-report (diary) and recall (questionnaire) measures of the same sexual behaviour. *Aids Care*. 1999; 11(2):221–234. [PubMed: 10474624]
- Crespi LP. The interview effect in polling. *Public Opinion Quarterly*. 1948; 12:99–111.
- Crowne, D.; Marlowe, D. *The approval motive: Studies in evaluative dependence*. Wiley; New York: 1964.
- Festinger, L. *A theory of cognitive dissonance*. Row-Peterson; Evanston, IL: 1957.
- French, John R.P. Field experiments: Changing group productivity.. In: Miller, JG., editor. *Experiments in social process: A symposium on social psychology*. McGraw-Hill; 1950.
- French, John R.P. Experiments in field settings.. In: Festinger, L.; Katz, D., editors. *Research Methods in the Behavioral Sciences*. Dryden Press; 1953.
- Fowler, FJ. *Improving survey questions: Design and evaluation*. Sage; Thousand Oaks, CA: 1995.
- Garry M, Sharman SJ, Feldman J, Marlatt GA, Loftus EF. Examining memory for heterosexual college students' sexual experiences using an electronic diary. *Health Psychology*. 2002; 21(6): 629–634. [PubMed: 12433018]
- Jaccard J, McDonald R, Wan CK, Dittus PJ, Quinlan S. The accuracy of self-reports of condom use and sexual behavior. *Journal of Applied Social Psychology*. 2002; 32:1863–1905.
- Johnson J, Bytheway B. An evaluation of the use of diaries in a study of medication in later life. *International Journal of Social Research Methodology*. 2001; 4(3):183–204.
- Kardes FR, Allen CT, Pontes MJ. Effects of multiple measurement operations on consumer judgment - Measurement reliability or reactivity. *Advances in Consumer Research*. 1993; 20:280–83.
- Kreuter F, Presser S, Tourangeau R. Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*. 2008; 72:847–865.
- Ku L, Sonenstein FL, Pleck JH. The dynamics of young men's condom use during and across relationships. *Family Planning Perspectives*. 1994; 26:246–251. [PubMed: 7867771]
- Lemmens P, Tan ES, Knibbe RA. Bias due to nonresponse in a Dutch survey on alcohol consumption. *British Journal of Addiction*. 1988; 83:1069–1077. [PubMed: 3224190]
- Lesthaeghe RJ, Neidert L. The second demographic transition in the United States: Exception or textbook example? *Population and Development Review*. 2006; 32(4):669–698.

- Moorman RH, Podsakoff PM. A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology*. 1992;65:131–149.
- Nancarrow C, Cartwright T. Online access panels and tracking research: The conditioning issue. *International Journal of Market Research*. 2007; 49:573–94.
- Paulhus DL. Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*. 1984;46:598–609.
- Schroder KEE, Carey MP, Venable PA. Methodological challenges in research on sexual risk behavior: Accuracy of self-reports. *Annals of Behavioral Medicine*. 2003; 26(2):104–123. [PubMed: 14534028]
- Schuman, H.; Presser, S. Questions and answers in attitude surveys: Experiments on question form, wording, and context. Academic Press; New York: 1981.
- Searles JS, Helzer JE, Walter DE. Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behaviors*. 2000; 14(3):277–286. [PubMed: 10998953]
- Stephen AM. The case for diet diaries in longitudinal studies. *International Journal of Social Research Methodology*. 2007; 10(5):365–377.
- Sturgis, P.; Allum, N.; Brunton-Smith, I. Attitudes over time: The psychology of panel conditioning.. In: Lynn, P., editor. *Methodology of Longitudinal Surveys*. Wiley; New York: 2009. p. 113-126.
- Tourangeau, R.; Rips, LJ.; Rasinski, K. *The psychology of survey response*. Cambridge University Press; New York: 2000.
- Veroff J, Hatchett S, Douvan E. Consequences of participating in a longitudinal-study of marriage. *Public Opinion Quarterly*. 1992; 56:315–27.
- Waterton, J.; Lievesley, D. Evidence of conditioning effects in the British Social Attitudes Panel Survey.. In: Kasprzyk, D.; Duncan, GJ.; Kalton, G.; Singh, MP., editors. *Panel surveys*. Wiley; New York: 1989. p. 319-339.
- Warren JR, Halpern-Manners A. Panel conditioning in longitudinal social science surveys. *Sociological Methods & Research*. 2012; 41(4):491–534.
- Welti, K.; Wildsmith, E.; Manlove, J. *Child Trends Research Brief*. Child Trends; Washington, DC: 2011. Trends and recent estimates: contraceptive use among U.S. teens and young adults.. Publication #2011-23
- Williams P, Block LG, Fitzsimons GJ. Simply asking questions about health behaviors increases both healthy and unhealthy behaviors. *Social Influence*. 2006; 1:117–27.
- Wilson TD, Kraft D. Why do I love thee: Effects of repeated introspections about a dating relationship on attitudes toward the relationship. *Personality and Social Psychology Bulletin*. 1993; 19:409–18.
- Zwane AP, Zinman J, van Dusen E, Pariente W, Null C, Miguel E, Kremer M, Karlan DS, Hornbeck R, Gine X, Duflo E, Devoto F, Crepon B, Banerjee A. Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences*. 2011; 108:1821–26.

TABLE 1

Descriptive Statistics for Sociodemographic Characteristics among the Main RDSL Sample, Control Group, and Journal Group (standard errors in parentheses for non-dichotomous variables)

	RDSL (N=1003)	Experiment Sample (N=200)		All samples				
		Control (N=100)	Journal (N=100)	difference		Min.	Max.	
		Mean/Std. Dev.	Mean/Std. Dev.	Mean/Std. Dev.	p			sig
Age								
18 years old	.41	.41	.49	.26		0	1	
19 years old	.50	.48	.42	.40		0	1	
20 years old	.09	.11	.09	.64		0	1	
African American	.34	.34	.31	.65		0	1	
School enrollment/type								
Enrolled in high school	.13	.12	.22	.06	†	0	1	
Enrolled in 2 year college/vocational program	.29	.25	.24	.87		0	1	
Enrolled in 4 year college	.27	.25	.21	.50		0	1	
Dropped out of high school/not enrolled	.08	.09	.08	.80		0	1	
Completed high school/not enrolled	.22	.29	.25	.53		0	1	
Receiving public assistance	.27	.27	.30	.64		0	1	
Religious importance (1 = not important, 4 = more important than anything else)	2.69 (.03)	2.61 (.09)	2.69 (.09)	.53		1	4	
Living with sexual partner	.15	.07	.15	.07	†	0	1	
Biological mother <20 at first birth	.37	.46	.28	.01	**	0	1	
Family Structure								
Two parents	.52	.48	.59	.12		0	1	
One biological parent only	.40	.43	.35	.25		0	1	
Other	.08	.09	.06	.42		0	1	
Mother's education < high school graduate	.09	.13	.10	.51		0	1	
Parents' Income								
\$14,999	.15	.21	.24	.61		0	1	
\$15,000 - \$44,999	.28	.30	.27	.64		0	1	
\$45,000 - \$74,999	.19	.23	.25	.74		0	1	
\$75,000+	.18	.15	.18	.57		0	1	
Don't know/refuse	.20	.11	.06	.21		0	1	

* p < .05

*** p < .001, two-tailed two-sample t-tests.

† p < .010

** p < .01

TABLE 2

Survey Questions for Measures of Sex, Pregnancy, and Contraceptive Use

	Survey Question	Baseline Interview	Weekly Journal Interviews	Closeout Interview
Sex				
Age at first sex <17	How old were you the first time you had sexual intercourse?	X		
Number of sexual partners	With how many total partners have you had sexual intercourse?	X		
Ever had sex without birth control	Have you ever had sexual intercourse without using some method of birth control, such as condoms, pills, or another method?	X		
Pregnancy				
Number of prior pregnancies	Please think of all the times you have been pregnant, whether you are currently pregnant or the pregnancy ended in live birth, miscarriage, stillbirth, abortion, or ectopic pregnancy. How many times have you been pregnant in your life? (CASI)	X		X
Reported a new pregnancy in past 12 months	<ul style="list-style-type: none"> • Do you think there might be a chance that you are pregnant right now? • Has a pregnancy test indicated that you are currently pregnant? 		X	
Contraceptive Use				
Ever used contraception	Have you ever used anything that can help people avoid becoming pregnant?	X		X
Currently using any type of contraception	Are you currently using anything that can help people avoid becoming pregnant, even if you are not using it to keep from getting pregnant yourself?	X		X
Pills	• Are you taking birth control pills for any reason?	X		X
Patch/Ring	<ul style="list-style-type: none"> • Are you using the birth control patch for any reason? • Are you using the NuvaRing for any reason? 	X		X
Depo-Provera	• Do you have a current Depo-Provera shot or any other type of contraceptive shot?	X		X
IUD/Implanon	<ul style="list-style-type: none"> • Do you have an implant such as Norplant or another contraceptive implant? • Do you have an IUD? 	X		X
Other	• Do you avoid having sex because it might be a time of the month you could get pregnant?	X		X
Contraceptive use during past week	• Since ____, did you use or do anything that can help people avoid becoming pregnant, even if you did not use it to keep from getting pregnant yourself?		X	
Pills	• Did you use birth control pills (for any reason)?		X	

	Survey Question	Baseline Interview	Weekly Journal Interviews	Closeout Interview
Patch/Ring	<ul style="list-style-type: none"> • Did you use the birth control patch (for any reason)? • Did you use the NuvaRing? 		X	
Depo-Provera	<ul style="list-style-type: none"> • Did you use Depo-Provera or any other type of contraceptive shot? 		X	
IUD/Implanon	<ul style="list-style-type: none"> • Did you have an implant such as Norplant or another contraceptive implant? • Do you have an IUD? 		X	
Other	<ul style="list-style-type: none"> • Did you avoid having sex because you thought it was a time of the month you could get pregnant? 		X	

TABLE 3

Differences in Sex, Pregnancy, and Contraceptive Use at Baseline, Journal [J] versus Control [C] Group

Variable	Unweighted Proportions		[C]-[J]					
	[C] (n=100)	[J] (n=100)	Un-Weighted Difference	p	sig	Weighted Estimate of Difference	p	sig
Sex								
Age at first sex <17	.47	.54	-.07	.32		-.11	.14	
Number of sexual partners	2.45	3.30	-.85 (.48)	.08	†	-1.15 (.64)	.07	†
Ever had sex without birth control	.43	.46	-.03 (.07)	.67		-.05 (.07)	.46	
Pregnancy								
Number of prior pregnancies	.31	.48	-.17	.10	†	-.21	.04	*
Contraceptive Use								
Ever used contraception	.76	.77	-.01 (.06)	.87		-.05 (.06)	.44	
Currently using any type of contraception	.54	.56	-.02 (.07)	.78		-.06 (.08)	.44	
Contraceptive Method Type								
Pills	.26	.35	-.09 (.07)	.17		-.12 (.06)	.07	†
Other	.11	.10	.01 (.04)	.82		-.01 (.05)	.84	
Depo-Provera	.11	.07	.04 (.04)	.33		.04 (.05)	.38	
IUD/Implanon	.04	.03	.01 (.03)	.70		.01 (.02)	.58	
Patch/Ring	.02	.01	.01 (.02)	.56		.02 (.02)	.42	

Standard errors in parentheses. Unweighted differences tested for statistical significance with two-sample two-tailed t-tests. Weighted estimates computed using estimated propensity of being in journal (versus control) group, based on significant ($p < .10$) differences in Table 1 (enrolled in high school, living with a sexual partner, and biological mother < 20 at first birth). Weighted differences tested for statistical significance with teffects in Stata, two-tailed tests.

** $p < .01$

*** $p < .001$

† $p < .010$

* $p < .05$

TABLE 4

Changes in Pregnancy and Contraceptive Use Across the 12-Month Study Period, Journal versus Control Group (number of respondents in parentheses)

	Unweighted Proportions		[C]-[J]					
	Control [C] (n=94)	Journal [J] (n=92)	Un-Weighted Difference	p	sig	Weighted Estimate of Difference	p	sig
Pregnancy								
New pregnancy in past 12 months	.15	.20	-.05 (.07)	.49		-.06	.38	
Contraceptive Use								
Change in proportion who have <i>ever</i> used contraception	+.11	+.02	.08 (.05)	.07	†	.15 (.06)	.01	**
Change in proportion who are <i>currently</i> using contraception	+.19	+.09	.10 (.07)	.16		.15 (.07)	.02	*
Proportion of non-users who began using a contraceptive method during the study period ^a	+.51	+.36	.15 (.11)	.15		.16 (.15)	.28	
Proportion who switched from less to more effective method during the study period ^b	+.10	+.14	-.04 (.07)	.57		-.05 (.08)	.51	
Change in proportion who are using each specific method								
Pills	+.13	+.14	-.01 (.06)	.83		-.10 (.08)	.23	
Other	+.07	-.04	.12 (.06)	.04	*	.15 (.07)	.03	*
Depo-Provera	-.02	-.03	.01 (.04)	.79		.05 (.04)	.29	
IUD/Implanon	-.01	.03	-.04 (.03)	.16		-.02 (.03)	.55	
Patch/Ring	.02	-.01	.03	.18		.04 (.02)	.11	

Standard errors in parentheses.

Unweighted differences tested for statistical significance with two-sample two-tailed t-tests.

Weighted estimates computed using estimated propensity of being in journal (versus control) group, based on significant ($p < .10$) differences in Table 1 (enrolled in high school, living with a sexual partner, and biological mother < 20 at first birth), Table 3 (number of sexual partners, number of prior pregnancies), and the initial value of the change score (e.g., initial levels of pill use for models of change in pill use). Weighted differences tested for statistical significance with teffects psmatch in Stata, two-tailed tests.

*** $p < .001$

^a Among those who were not using contraception at the baseline interview; control group n=45, journal group n=42.

^b Among those who were using contraception at the baseline interview; control group n=49, journal group n=50.

† $p < .010$

* $p < .05$

** $p < .01$

TABLE 5

Mean Change Scores for Attitude Measures, Journal versus Control Group (higher numbers = agreement)

	Unweighted Mean Change		[C]-[J]					
	Control [C] (n=94)	Journal [J] (n=92)	Un-Weighted Difference	p	sig	Weighted Difference	p	sig
Condoms								
1 If a woman asks her partner to use a condom, he will think that she doesn't trust him. ^a	-.21	-.43	.22 (.21)	.28		.07 (.19)	.70	
Sex								
2 Young people should not have sex before marriage. ^a	.06	-.14	.21 (.19)	.28		.21 (.21)	.30	
Cohabitation								
3 It is alright for a couple to live together without planning to get married. ^a	-.37	.15	-.22 (.20)	.28		-.17 (.26)	.52	
Marriage								
4 Suppose that your life turned out so that you never married, how much would that bother you? ^b	.07	.16	-.09 (.24)	.72		.12 (.27)	.65	
Pregnancy								
5 Getting pregnant at this time in your life is one of the worst things that could happen to you. ^a	-.60	-.41	-.19 (.20)	.34		.05 (.27)	.84	
Childbearing								
6 Being a mother and raising children is the most fulfilling experience a woman can have. ^a	.32	-.09	.41 (.16)	.009	**	.28 (.20)	.16	
7 Relationships between men and women improve after they have a baby together. ^a	-.21	-.12	-.09 (.15)	.54		-.22 (.18)	.21	
8 If a woman waits for the perfect time to have a baby, she	.33	-.14	.47 (.16)	.003	**	.40 (.17)	.02	*

	Unweighted Mean Change		[C]-[J]					
	Control [C] (n=94)	Journal [J] (n=92)	Un-Weighted Difference	p	sig	Weighted Difference	p	sig
will probably have trouble getting pregnant. ^a								
9 It is alright for a woman to have a child without being married. ^a	.11	-.02	.13 (.16)	.43		.01 (.18)	.94	
10 Suppose your life turned out so that you never had children, how much would that bother you? ^b	.14	.12	.02	.90		-.08 (.19)	.65	
11 What do you think is the ideal number of children for the average American family?	-.04	-.14	.10 (.13)	.45		.18 (.16)	.28	

Standard errors in parentheses.

Unweighted differences tested for statistical significance with two-sample two-tailed t-tests.

Weighted estimates computed using estimated propensity of being in journal (versus control) group, based on significant ($p < .10$) differences in Table 1 (enrolled in high school, living with a sexual partner, and biological mother < 20 at first birth), Table 2 (number of sexual partners, number of prior pregnancies), and the initial value of the change score. Weighted differences tested for statistical significance with teffects psmatch in Stata, two-tailed tests.

*** $p < .001$, two-tailed two-sample t-tests.

^a 1 = strongly disagree, 2 = disagree, 3 = neutral/neither, 4 = agree, 5 = strongly agree

^b 0 = not at all, 5 = extremely

* $p < .05$

** $p < .01$