Panel Conditioning in Longitudinal Social Science Surveys

John Robert Warren
Andrew Halpern-Manners

University of Minnesota
Department of Sociology
Minnesota Population Center

Version: September 2009

DRAFT: PLEASE DO NOT CITE OR QUOTE WITHOUT AUTHORS' PERMISSION

Paper prepared for presentation at the 2010 annual meetings of the Population Association of America. The National Science Foundation (SES-0647710) and the University of Minnesota's Life Course Center, Department of Sociology, College of Liberal Arts, and Minnesota Population Center have provided support for this project. We warmly thank Penny Edgell, Eric Grodsky, Conrad Hackett, Michael Hout, Ross Macmillan, Jeylan Mortimer, Lowell Taylor, Florencia Torche, and workshop participants at the University of Minnesota, the University of Texas, the University of Wisconsin-Madison, and New York University for their constructive criticism and helpful comments. However, all errors or omissions are the authors' responsibility.
Please direct correspondence to John Robert Warren, Department of Sociology, University of Minnesota, 909 Social Sciences, 267 ~ 19th Ave. South, Minneapolis, MN 55455 or email warre046@umn.edu.

Panel Conditioning in Longitudinal Social Science Surveys

ABSTRACT

Does participating in one wave of a longitudinal survey affect respondents' reports of their social and economic well-being in follow-up survey waves? If survey participation does alter respondents' subsequent answers to questions about their social and economic well-being, then this calls into question the validity of information derived from any number of widely used data resources. In this paper, we estimate the magnitude of what methodologists have called panel conditioning or "time-in-survey" effects. Previous efforts to estimate the magnitude of panel conditioning effects have utilized methodologically weak designs and have focused on consequences for a limited range of measures. We use a stronger research design and use large-scale survey data on a wide range of measures of individuals' social and economic well-being to test a series of theoretically derived propositions about the circumstances under which panel conditioning effects should be most severe.

Longitudinal surveys provide tremendous methodological advantage in making causal inferences and in understanding processes that unfold over time. Despite their great value, they also present a variety of uniquely challenging methodological problems. Among the least well understood of these problems is panel conditioning, or bias introduced when participating in one wave of a longitudinal survey changes respondents' attitudes, behaviors, and/or the quality of reports of attitudes or behaviors in subsequent survey waves.[1] The lack of social science attention to such potentially consequential biases is surprising. If partaking in a longitudinal survey alters participants' responses to subsequent attitudinal and/or behavioral survey questions, either because their actual attitudes or behaviors change or because the quality of their reports of them change, this calls into question the validity and reliability of information derived from any number of widely used longitudinal data resources.

We have three main objectives in this paper—one is conceptual and two are empirical. First, we seek to map the substantive and methodological situations in which users of longitudinal survey data should expect panel conditioning to arise and to be most consequential. Using theory and research from consumer marketing, cognitive psychology, and elsewhere, we develop seven propositions about the circumstances under which panel conditioning effects should be most pervasive. Second, we offer provisional evidence about the merits of these seven propositions. Our objective in this section is to document whether and under what circumstances having responded to questions in earlier survey waves affects respondents' subsequent answers

[1] We use the term "panel conditioning" synonymously with what has been called, among many other things, "time in survey effects" (Corder and Horvitz 1989), "mere measurement effects" (Godin et al. 2008; Morwitz et al. 1993), and "question-behavior effects" (Spangenberg et al. 2008; Sprott et al. 2006).

to parallel questions. Third, we investigate whether panel conditioning also biases the results of multivariate analyses. Beyond its implications for marginal distributions of responses to single survey items, does panel conditioning affect estimates of multivariate relationships?

Although our research sheds important new light on the nature and magnitude of panel conditioning biases in longitudinal surveys, we recognize that our work is only a starting point for what we hope will become a rigorous new line of methodological inquiry. For reasons described below, and despite our belief that they are the best that can be achieved using existing data, our empirical findings are necessarily provisional. We end by calling for new data collection that will more definitively speak to our empirical questions. In the meantime, our theoretical propositions and our empirical evidence should provide useful practical guidance to those who collect or analyze longitudinal data. If nothing else, we hope that our work will motivate and facilitate subsequent theoretical and empirical explorations that can provide as much understanding of panel conditioning biases as we currently have about other problems associated with the collection and analysis of longitudinal data (e.g., panel attrition).

In the sections that follow we (1) review and critique the existing empirical evidence on the nature and magnitude of panel conditioning biases; (2) use theory and evidence from a number of disciplines to develop a set of seven propositions about the circumstances under which panel conditioning may be most pronounced; (3) describe our empirical strategy for obtaining evidence regarding these propositions based on data from three major national surveys; (4) present and discuss our results with an eye toward providing initial practical guidance to users of longitudinal data; and (5) outline a research agenda that will allow for a more complete understanding of the nature and magnitude of panel conditioning biases.

PREVIOUS EVIDENCE

Methodologists have long understood that "testing" and "reactivity" can pose serious threats to the internal validity of experimental and non-experimental research (e.g., Campbell and Stanley 1966; Campbell 1957; Lazarsfeld 1940; Lazarsfeld 1941). For example, we know that test takers usually score higher on intelligence or cognitive assessments the second time they take them, mainly because they are more familiar with the format, style, and requirements of the testing instrument (Grindstaff et al. 2006; Hausknecht et al. 2007; Richardson and Robinson 1921). Less well understood is whether these biases also arise in larger-scale longitudinal surveys of the sort commonly used in the social, behavioral, and health sciences. Nonetheless, researchers in several disciplines have investigated the nature and magnitude of panel conditioning biases in what have been largely disconnected lines of inquiry.

Political Science & Public Opinion

Public opinion researchers have studied the impact of participating in opinion polls on respondents' attitudes. Early work by Crespi (1948), for example, suggested that such polls may have the desirable effects of educating and cognitively stimulating respondents, but the unwanted effect of biasing their opinions in a number of ways. Likewise, Waterton and Livesley (1989: 336) found that participating in the British Social Attitudes Panel politicizes respondents' attitudes and leads them to report more honestly over time.

Political scientists and others have frequently concluded that participation in political opinion polls increases voter turnout in the U.S. and elsewhere (Bartels 1999; Clausen 1968; Granberg and Holmberg 1992; Greenwald et al. 1987; Kraut and McConahay 1973; Simmons et al. 1993; Traugott and Katosh 1979; Voogt and Van Kempen 2002; Yalch 1976). For example, using an experimental design, Kraut and McConahay (1973) found that being interviewed before

a primary election doubled respondents' probability of voting in that election (as validated by precinct records). On the other hand, Mann (2005) recently disputed the methodological basis of most of these findings; see also Smith et al. (2003).

Consumer Marketing & Cognitive Psychology

Grounded in theory and research in cognitive psychology, consumer-marketing researchers have long studied the effect of measuring a person's *intended* or *forecasted* behaviors on their *actual* behaviors (e.g., Borle et al. 2007; Chandon et al. 2005; Dholakia and Morwitz 2002; Feldman and Lynch 1988; Fitzsimons and Williams 2000; Janiszewski and Chandon 2007; Morwitz 2005; Sherman 1980; Spangenberg et al. 2003; Williams et al. 2004). These scholars have demonstrated panel conditioning effects on automobile purchases (e.g., Morwitz et al. 1993), personal computer purchases (e.g., Chandon et al. 2005), recycling behaviors (e.g., Sprott et al. 1999), fitness club attendance (e.g., Spangenberg 1997), on-line grocery purchases (e.g., Chandon et al. 2004), a variety of delinquent behaviors (Fitzsimons et al. 2007; Spangenberg and Obermiller 1996), and many others. In each case, the act of measuring respondents' intended or predicted future behaviors appears to alter their actual subsequent behaviors. In most cases, these effects are inferred from experimental research designs.

This empirical finding, however, has not been universal. Other investigators in consumer marketing and cognitive psychology have reached a different conclusion, finding no evidence of panel conditioning with respect to feelings of economic security (Sobol 1959), consumer expenditures (Silberstein and Jacobs 1989), customer satisfaction (Swan et al. 1981), and various other outcomes (e.g., Clinton 2001; Wang et al. 2000).

Public Health and Medicine

Consumer marketing researchers and cognitive psychologists have occasionally considered panel conditioning effects on outcomes of relevance to public health. For example, Williams et al. (2006) find that asking college students about health behaviors can have the effect of increasing healthy behaviors (exercise) and of increasing unhealthy behaviors (illegal drug use). Likewise, in a 2008 debate in the *Journal of Consumer Psychology*, several authors weighed the ethical and public health implications of potentially increasing risky health behaviors among teens by simply asking about them in public health surveys (e.g., Fitzsimons and Moore 2008).

Public health scholars have also investigated the magnitude of panel conditioning biases for measures of specific public health outcomes. Battaglia, Zell, and Ching (1996), for example, found evidence that participating in a survey about childhood vaccinations increased the rate at which parents immunized their children. Wilson and Howell (2005) partly attribute apparent increases in the prevalence of arthritis in the Health and Retirement Survey (HRS) to panel conditioning, although there is some debate about the quality of their evidence (Weir and Smith 2007; Wilson and Howell 2007). Godin et al. (2008) found that asking about blood donation significantly increased respondents' frequency of blood donations. Likewise, O'Sullivan et al. (2004) found panel conditioning effects on respondents' rates of screening for colorectal cancer.

On the other hand, Sutton et al. (1994) found no such effects on rates of screening for breast cancer, Corder and Horvitz (1989) concluded that panel conditioning does not bias reports of medical expenditures, and Underwood et al. (2006) found no panel conditioning effects on older people's fear of falling, number of reported falls, or frequency of contact with health service professionals.

Economics and Sociology

Economists and sociologists have only occasionally been explicitly concerned about panel

conditioning biases. One notable exception involves the panel design of the Current Population Survey (CPS), which, some scholars suggest, may have the consequence of leading to some degree of panel conditioning bias in estimating unemployment rates (Bailar 1975; Bailar 1989; Hansen et al. 1955; Shack-Marquez 1986; Solon 1986). These authors have demonstrated that for any particular calendar month, unemployment rates for respondents in their initial month in the CPS rotation are higher than for respondents who have been in the panel for longer.[2] The implication is that CPS respondents are less likely to be unemployed, or at least less willing to report unemployment, after their first of eight CPS interviews.

Economists and other social scientists have also found evidence of panel conditioning in the Knowledge Network Panels (Clinton 2001; Nancarrow and Cartwright 2007), the German General Social Survey (Porst and Zeifang 1987), and in a large-scale panel run by AT&T (Toh et al. 2006). There do not appear to be similar biases, however, in the Survey of Income and Program Participation (McCormick et al. 1992; Pennell and Lepkowski 1992) or the National Medical Care Utilization and Expenditure Survey (Corder and Horvitz 1989).

LIMITATIONS OF PREVIOUS EVIDENCE

These efforts to identify and quantify biases from panel conditioning have generally suffered from three important limitations. First, theoretical perspectives on the mechanisms potentially giving rise to panel conditioning have not been widely applied outside of consumer marketing

[2] Indeed the issue has made its way into documentation about the design of the CPS (U.S. Bureau of Labor Statistics 2000), where Table 16-10 shows that in September 1995 the unemployment rate for CPS respondents in their first month in the sample was 8.6 percent higher than for CPS respondents as a whole in that month.

and cognitive psychology. Second, most prior researchers have utilized research designs that are weak with respect to their ability to yield valid estimates of panel conditioning effects and/or with respect to their ability to produce generalizable findings. Third, no prior research has examined the extent to which panel conditioning may alter the substantive conclusions researchers draw about relationships among two or more variables. In the remainder of this section, we discuss each of these issues in turn.

Theoretical Perspectives

Social scientists have generally not thought carefully about the circumstances under which panel conditioning effects might arise. When might we expect survey participation to change respondents' actual attitudes and behaviors? When might we expect survey participation to change merely the quality of their reports of those attitudes and behaviors? Without answers to questions like these, researchers are unable to anticipate situations in which panel conditioning is worth worrying about as they collect and analyze longitudinal survey data.

In contrast, researchers in consumer marketing and cognitive psychology have well-developed theories about the mechanisms that give rise to particular forms of panel conditioning. This work is generally informed by theoretical perspectives on the subtle cognitive processes that underlie attitude formation and change, decision making, and the relationship between attitudes and behaviors (e.g., Fazio 1989; Fazio et al. 1986; Feldman and Lynch 1988; Millar and Tesser 1986; Tesser 1978). Several of the seven propositions that we develop below are based on theory and research in these fields, but they are intended to be widely applicable to any substantive field that utilizes data from longitudinal surveys of individuals.

Research Design

Researchers have typically employed one of three research designs in their investigations of panel conditioning effects. First, they have considered changes over time in survey responses by a single group of respondents to a longitudinal survey (Sturgis et al. 2009; Toh et al. 2006). For example, Crespi (1948) described "commitment biases" that result from respondents being asked the same attitude questions on multiple occasions. His evidence for this form of bias was that respondents became less likely across survey waves to offer "no opinion" responses. One problem with this design is that it is hard to disentangle the effects of panel conditioning from endogenous change in the attitude or behavior being studied (Campbell and Stanley 1966).

Second, a number of studies—mostly in consumer marketing and cognitive psychology—have randomly assigned individuals to a treatment group (pre-test *and* post-test surveys) or a control group (post-test survey only) in order to estimate the impact of panel conditioning (e.g., Borle et al. 2007; Bridge et al. 1977; De Amici et al. 2000; Godin et al. 2008; O' Sullivan et al. 2004; Williams et al. 2006; Yalch 1976). Although these experimental designs are much stronger with respect to establishing causal effects of panel conditioning, the broader generalizability of their findings is unclear. Many of these studies are carried out among students in college classes, customers of particular businesses, or voters in specific precincts, and they generally focus on a narrow range of substantive topics (e.g., computer purchases or voting).

Third, other researchers have proceeded by comparing survey responses from members of a longitudinal panel to those from members of an independent cross-sectional sample drawn from the same population. Such a design has been utilized, for example, in research on panel conditioning effects on unemployment rates in the CPS and SIPP (e.g., Bailar 1975), consumer behaviors (e.g., Nancarrow and Cartwright 2007), public opinion (e.g., Voogt and Van Kempen

2002), and public health (e.g., Underwood et al. 2006). Details of the problems with this design are laid out elsewhere (e.g., Holt 1989; Sturgis et al. 2009; Williams and Mallows 1970). Most importantly, such a design potentially confounds biases from panel *conditioning* with those from panel *attrition*. Whereas the cross-sectional sample may be representative of some population, the panel sample may have suffered from non-random attrition over time. Unless researchers take steps to adjust for the resulting panel selectively, differences in responses between the two samples cannot be clearly attributed to panel conditioning. Although some of the research reviewed above makes efforts to make such adjustments, the success of those efforts is unclear.

Multivariate Relationships?

To our knowledge, no prior research has looked beyond the implications of panel conditioning for marginal distributions of responses to survey items. Even if it is true, for example, that participating in a survey about vaccinations increases the chances that parents will subsequently vaccinate their children (Battaglia et al. 1996), does it follow that panel conditioning also biases our assessment of socioeconomic differentials in rates of childhood vaccinations? In the present analysis, we go beyond estimating the effects of panel conditioning for distributions of survey measures to investigate the consequences of panel conditioning for multivariate analyses.

UNDER WHAT CIRCUMSTANCES DOES PANEL CONDITIONING ARISE?

The empirical evidence reviewed above suggests that participating in one wave of a longitudinal survey may alter, at least in some situations, participants' responses to questions on subsequent survey waves. In general, what are those situations? That is, when should researchers expect biases from panel conditioning? The seven propositions that we set forth below guide our subsequent empirical analyses, but we hope that they will also serve as an initial set of guidelines

for producers and users of longitudinal data who may be concerned about the consequences of panel conditioning for their research.

Proposition 1: *Real changes in an attitude will occur as the result of responding to questions about that attitude when respondents' initial attitudes are less crystallized and when the issue at hand is salient for them.*

Many people treat attitudes as "crystallized" or fixed entities in people's minds. Schuman and Presser (1981: 271) define "crystallized attitudes" as those "that exist independently of our measurement, and that when appropriately measured show high reliability." In contrast, social psychologists, marketing researchers, and others view attitudes as frequently varying in their degree of crystallization (Bridge et al. 1977; Millar and Tesser 1986; Tesser 1978; Tourangeau et al. 2000) and in their degree of cognitive accessibility to those who hold them (Fazio 1989; Fazio et al. 1986; Feldman and Lynch 1988; Powell and Fazio 1984). The implication is that respondents' attitudes may be quite fluid and malleable, particularly when the topic is one about which they have given relatively little thought.

With this conceptualization of attitudes in mind, research in survey methodology, experimental social psychology, and consumer marketing has shown that thinking about an attitude as a result of responding to a survey question has the potential to change that attitude (Millar and Tesser 1986; Morwitz and Fitzsimons 2004; Sudman et al. 1996; Tesser 1978; Zaller and Feldman 1992) and to make the attitude more accessible to the respondent in the future (Chandon et al. 2005; Feldman and Lynch 1988). Respondents who lack crystallized attitudes about a topic will nonetheless offer a response to a question about that attitude in a baseline survey (Bishop et al. 1986). However, the act of participating in the baseline survey may set in motion a series of thoughts or actions that change that attitude by the time of a follow-up survey (Sturgis et al. 2009; Waterton and Lievesley 1989; Wilson et al. 1996). As a result, when asked

about that attitude again in the follow-up survey, some people's responses may differ from their responses in the baseline survey (Kardes et al. 1993).

We expect that responding to attitude items in a survey will most strongly affect attitudes when respondents have not already developed strongly held or carefully considered views about that topic. Furthermore, this effect will only come about when the topic is perceived as important or salient for respondents (Bridge et al. 1977; Duncan and Kalton 1987).

Proposition 2: *Real changes in a behavior will occur as the result of responding to questions about that behavior when the questions serve to increase respondents' interest in, awareness of, or knowledge about that behavior.*

To explain why participating in a pre-election survey increased respondents' chances of voting, Clausen (1968) speculated that the survey "stimulated" respondents' interest in the election; this increase in interest in the election led more respondents to turn out to the polls. In replicating and re-evaluating Clausen's (1968) findings, Kraut and McConahay (1973) developed two alternative hypotheses: the "alienation reduction" hypothesis and the "self concept" hypothesis. In their own way, each of Clausen's (1968) and Kraut and McConahay's (1973) proposed cognitive mechanisms implies that responding to survey questions about a behavior alters respondents' motivation to engage in that behavior.

Although there is little empirical work on the topic, some have also suggested that survey questions may serve to convey new information to respondents. This new information may facilitate behaviors that would not have otherwise have been possible (Fitzsimons et al. 2007; Nancarrow and Cartwright 2007). For example, a survey of low-income individuals that focuses on their use of the State Children's Health Insurance Program (SCHIP) may inform many respondents that the SCHIP exists and is something for which their children may be eligible.

The extent to which survey items affect actual subsequent behaviors will depend on

whether respondents view the focal behavior as having some positive utility. A survey of respondents' cholesterol consumption may have the positive utility of making them better informed about which foods are high in cholesterol, thus potentially altering eating habits. Conversely, a survey of respondents' illicit drug use may make them better informed about varieties of illicit drugs, but is unlikely to motivate them to use illicit drugs more often.

Proposition 3: *Attitudes and behaviors will (at least appear to) change over time when survey questions require respondents to provide socially non-normative or stigmatized responses.*

Survey questions can force respondents to confront the reality that their attitudes and/or behaviors conflict with what mainstream society regards as normative or appropriate. Reflecting on this conflict can create "cognitive dissonance" in respondents' minds, and reporting non-normative or stigmatized attitudes and/or behaviors in a survey is a potentially embarrassing experience. In the context of research on the effects of answering questions about behavioral *intentions* or *expectations* on respondents' *actual* subsequent behaviors, cognitive psychologists note that one way for respondents to mitigate cognitive dissonance is to modify their subsequent behaviors (e.g., Fitzsimons and Moore 2008; Levav and Fitzsimons 2006; Spangenberg et al. 2008; Toh et al. 2006; Williams et al. 2006). We suggest that this process—which is well documented for questions regarding *predicted or intended behaviors*—may also operate in the context of questions regarding *attitudes* and regarding *previous or ongoing behaviors*.

Some respondents may react to survey questions by bringing their *actual* attitudes or behaviors into closer conformity with social norms. However, we suggest that many respondents may simply avoid cognitive dissonance and the embarrassment associated with offering socially non-normative or stigmatized responses by merely bringing their *reported* attitudes or behaviors

into closer conformity with social norms.[3] For example, the cognitive dissonance and possible embarrassment caused by answering questions about alcohol consumption on a baseline survey may lead heavy-drinking respondents to drink less or else to report inaccurately lower levels of alcohol consumption on follow-up surveys. In either case, apparent declines in levels of alcohol consumption across survey waves represent a form of panel conditioning (Williams et al. 2006).[4] In general, one or both of these forms of panel conditioning may arise in the context of any survey questions for which there are clear socially non-normative or stigmatized responses.

Proposition 4: *Attitudes and behaviors will appear to change across survey waves as the respondent becomes more comfortable with and trusting of the survey experience*

Survey research methodologists have found that respondents' judgments about the relative benefits and risks associated with answering survey questions are significantly related to the chances that they provide complete and accurate answers (Dillman 2000; Groves 1989; Willis et al. 1994). As respondents become more familiar with and trusting of the survey process and with interviewers and interviewing organizations, they may become less suspicious of interviewers and their confidence in the confidentiality of their responses grows (Fowler 1995).

---

[3] Lacking external validating information, it is hard to determine whether this form of panel conditioning changes respondents' actual attitudes and behaviors or just their reports of them (Duncan and Kalton 1987).

[4] The distinction between social desirability biases and panel conditioning is an important one. Social desirability may cause respondents to under-report socially stigmatized attitudes or behaviors. However, this under-reporting will happen on both baseline and follow-up surveys, and will not affect inferences about change over time. In contrast, panel conditioning will bias inferences about changes between baseline and follow-up surveys in attitudes and behaviors.

One consequence of this dynamic is that the tendency for survey respondents to give "socially desirable" answers may diminish as respondents' trust in interviewers and survey research projects grows (Nancarrow and Cartwright 2007; Tourangeau et al. 2000; Waterton and Lievesley 1989). Whereas Proposition 3 suggests that respondents will be *less* likely to offer socially non-normative or stigmatizing responses across survey waves, Proposition 4 suggests that they will be *more* likely to do so. For example, respondents may be more willing to report holding controversial or extreme views on social issues in follow-up surveys after they experience no negative consequences for sharing less extreme views in baseline surveys.

Proposition 5: *Attitudes and behaviors will appear to change across survey waves as respondents learn to manipulate a survey instrument in order to minimize their burden.*

Respondents sometimes find surveys to be tedious, cognitively demanding, excessively detailed, and/or longer than they might like. As a result, longitudinal survey respondents may learn to direct or manipulate the survey experience in such a way that minimizes its length and thus their burden (Bailar 1989; Hernandez et al. 1999; Kalton and Citro 2000; Nancarrow and Cartwright 2007; Wang et al. 2000). If true, this dynamic may lead to the false impression that attitudes or behaviors have changed across survey waves. For example, respondents may learn that numerous supplemental questions follow for each employee that they report supervising at the time of a baseline survey. As a result, and in order to reduce the duration of follow-up surveys, some respondents may subsequently report supervising fewer employees.

Proposition 6: *Attitudes and behaviors will appear to change across survey waves as respondents learn to provide more accurate and complete responses.*

Waterton and Livesly (1989: 324) note that repeated interviewing of the same respondents may lead "to improved understanding [by respondents] of the rules that govern the interview process." While this may translate into undesirable manipulation of the survey process

(as posited above), it may also lead to more accurate responses (Bailar 1989). Respondents' ability to provide accurate and complete answers may improve across waves of a longitudinal survey, resulting in evidence of change across survey waves when no such change has really occurred (Duncan and Kalton 1987; Kalton and Citro 2000; Nancarrow and Cartwright 2007; Sturgis et al. 2009; Wang et al. 2000).

Evidence concerning "practice effects" in cognitive and psychometric testing supports this proposition. This vast literature shows that scores can improve across administrations of a test simply as a result of respondents' improved familiarity with the testing instrument (Basso et al. 2002; Basso et al. 2001; Goldberg et al. 2007; Grindstaff et al. 2006; Hausknecht et al. 2007; Johnson et al. 1991; Kulick et al. 1984; Lemay et al. 2004; Richardson and Robinson 1921).

Proposition 7: *Panel conditioning is more likely when surveys occur more frequently*

In different ways, the six propositions above each imply mechanisms giving rise to panel conditioning that may be less powerful the longer the length of time between baseline and follow-up surveys. In each case, the longer the interval between surveys the more that intervening life events, subject maturation and change, historical events, forgetfulness, and other factors may overwhelm, counteract, or mute the effects of answering baseline questions on respondents' answers to follow-up questions. Although there have been no empirical tests of this proposition, it is generally true—as reviewed below—that the largest panel conditioning effects are found when baseline and follow-up surveys are separated by less than a year and that when surveys are separated by a year or more there are fewer reported effects.

RESEARCH DESIGN

We next describe our strategy for obtaining empirical evidence concerning the seven propositions developed above and concerning our question about whether panel conditioning

also biases the results of multivariate analyses. After outlining our empirical strategy and describing the three longitudinal surveys used in our analyses in this section, we go on to describe our focal measures and empirical results.

To estimate the magnitude of panel conditioning we compare members of a longitudinal panel (the "treatment" group) to *statistically equivalent* members of a fresh cross-sectional sample (the "control" group).[5] After making appropriate adjustments (described below) for panel attrition, any differences that obtain between the two groups may be solely attributed to differences in their exposure (or lack thereof) to the survey instrument. Our strategy for considering the implications of panel conditioning for multivariate analyses involves tests of the statistical significance of differences across treatment and control groups in the size of multiple regression coefficients.

Design components of three major longitudinal surveys—the U.S Current Population Survey (CPS), the German Socioeconomic Panel (GSOEP), and the 2008 U.S. General Social Survey (GSS)—facilitate this research design.[6] Although different in nature and content, all three are similar in one important respect—respondents are surveyed repeatedly and new members are periodically introduced or "rotated" into the panel. In the case of the CPS—which each month includes individuals in about 50,000 households who are representative of the civilian, household-based population of US—panel members (or rotation groups) are enumerated on eight separate occasions,

[5] Note that we use the terms "treatment group" and "control group" for heuristic purposes only. Although respondents are exogenously (and randomly) assigned to the different groups, we recognize that our research design and general analytic strategy does not technically amount to a controlled experiment.

[6] It is also possible to conduct these analyses using the Survey of Income and Program Participation.

spread across two equal time frames. After a rotation group enters the sample, its members are surveyed for four consecutive months, left un-enumerated during the subsequent eight months, and finally resurveyed for another four months. This design guarantees that at any point in time, 1/8 of the sample is in the first month of enumeration (rotation group 1), 1/8 is in the second month (rotation group 2), and so forth. It also ensures differences across rotation groups in respondents' familiarity with yearly CPS supplements like the March Demographic Supplement, such that in any given supplement 1/2 of the sample is responding to supplemental questions for the first time (rotation groups 1-4) but the other 1/2 (rotation groups 5-8) is responding to those items for the second time.[7]

In the case of the GSOEP—a longitudinal study of private households in the Federal Republic of Germany—original sample members have been interviewed on an annual basis since 1984 (Wagner et al. 2007). The survey instruments are comprised of wide array of social, economic, behavioral, and attitudinal items. In order to maintain sufficient statistical power and to ensure continued representativeness, new subsamples were added to the study in 1990, 1994, 1998, 2000, 2002, and, most recently, in 2006. Of these six additional subsamples, three are of particular interest for present analysis—the 1998 "E" subsample, the 2002 "F" subsample, and the 2006 "H" subsample, all of which were drawn using the same sampling scheme that was used to select members of the original longitudinal panel. This feature makes possible our key comparison: In 1998, 2002, and 2006 we are able to compare the responses of members of the longitudinal panel to the responses

[7] We restrict the CPS treatment and control groups to individuals who are household heads, who are not members of "replacement households," who were interviewed in person, and for whom information was not gathered by proxy. These restrictions limit our focus to people who answered questions about themselves and to people who experienced the same mode of interview.

of members of a fresh cross-sectional sample drawn at random from the same population.[8]

The GSS—which contains a standard "core" of demographic, behavioral, and attitudinal questions, plus topics of special interest—has traditionally consisted of annual or semi-annual cross-sectional surveys of about 3,000 people in American households. Beginning in 2006, a portion of the sample has been followed for longitudinal follow-up surveys. Those respondents participating in the GSS for the first time in 2008 constitute our control group; those who first participated in 2006 and who were then re-interviewed in 2008 constitute the treatment group.

Our key comparisons are only instructive insofar as the treatment and control groups differ *only* with respect to their exposure to items on the survey instrument. However, this important stipulation is violated if patterns of non-response vary substantially and systematically across treatment and control groups. Even if absolute non-response rates are identical for treatment and control group members, it is not necessarily true that treatment group non-responders will have the same distribution of attributes as control group non-responders. Recovering unbiased estimates of the magnitude of panel conditioning effects, therefore, depends in part on our ability to account for panel attrition in particular and differential response more generally.

To minimize the chances of conflating effects associated with panel conditioning with those attributable to panel attrition, we employ a post-stratification weighting technique (Wu and Sitter

[8] We restrict the GSOEP treatment and control groups to individuals who were household heads, who responded to the previous year's survey, who were interviewed by computer assisted personal interviewing, who were not members of the 2004 GSOEP oversample of high income individuals, and for whom information was not gathered by proxy. These restrictions limit our focus to people who answered questions about themselves and to people who experienced the same mode of interview.

2001). The methodology, which is often termed raking or sample-balancing (Deming and Stephan 1940; Little 1993; Stephan 1942), uses an iterative proportional fitting (IPF) algorithm to generate weights (Izrael et al. 2000; Izrael et al. 2004). These weights can, in turn, be used to correct for known discrepancies between a sample and a target population, which arise as a result of non-response or related coverage issues (see, e.g., Little 1993).

The raking algorithm proceeds as follows. First, we tabulate marginals for the control group on a specified set of raking variables (hereafter referred to as marginal control totals). Our choice of raking variables is informed by two main considerations: (1) they must be plausibly related to sample attrition; and (2) they cannot themselves be susceptible to panel conditioning. Each of our three data sets includes a wealth of sociodemographic indicators that satisfy these conditions.[9] Next,

[9] In the CPS, the raking variables include age (15-34, 35-44, 45-54, 55-64, or 65+), sex (male or female), race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, or other), marital status (currently married or not married), nativity (born in the US and both parents born in the US, born in the US but at least one parent born abroad, or born abroad), urbanicity (principal city, balance of city, nonmetropolitan, or not identified), and region (northeast, Midwest, south, or west). In the GSOEP, they consist of age (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, or 80+), sex (male or female), nationality (German national or non-German national), marital status (currently married, not currently married but married previously, or never married), and state of residence (one of 16 states or locales). In the GSS, the raking variables include age (18-34, 35-44, 34-54, 55-64, or 65+), sex (male or female), race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, or other), nativity (born in the US and both parents born in the US, born in the US but at least one parent born abroad, or born abroad), marital status, region of residence, number of siblings, and number of children.

the IPF algorithm uses the marginal control totals to compute the appropriate weights. The weights are first adjusted to be consistent with control totals from the marginal distribution of the first raking variable. The resulting weights are then recalibrated to the control totals for the second marginal distribution, a process that is repeated for each of the raking variables. One sequence of adjustments through all of the raking variables represents a single iteration. The algorithm iterates until the weighted marginals converge to the control totals (within a specified tolerance) for all of the marginal distributions simultaneously. After implementing the weights neither the treatment nor the control group is necessarily perfectly representative of any population, but that is relatively unimportant for our purposes. What *is* important is that the groups are effectively equated *to one another* so that comparison of responses across treatment and control groups can be more confidently attributed to panel conditioning effects.

In supplementary analyses, we find that after implementing the post-stratification weights, members of the treatment and control groups have nearly identical distributions of key variables that were *not* included among the raking variables, including educational attainment, occupational prestige, income, and length of residency in current dwelling. This gives us confidence that there are few unobserved, systematic differences between individuals in the treatment and control groups.

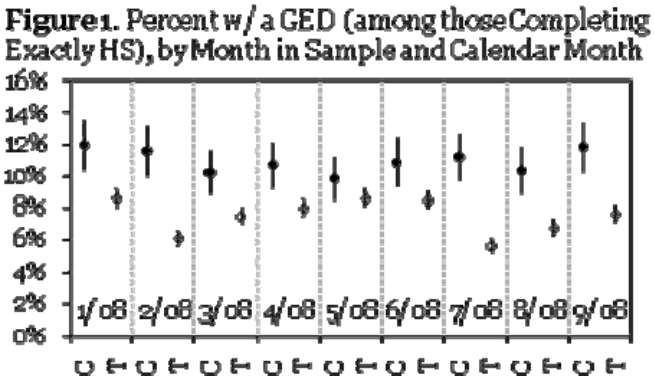
PRELIMINARY RESULTS

In this section, we present preliminary evidence regarding three of the propositions described above. This evidence is based only on data from the CPS, and includes just a limited number of survey items. Our goal is to demonstrate the plausibility of the propositions detailed above. By the time of the annual meetings, we will have fully developed results for all of the propositions using many more measures and all three data resources.

Proposition 3: *Attitudes and behaviors will (at least appear to) change over time when survey questions require respondents to provide socially non-normative or stigmatized responses.*

Proposition 4: *Attitudes and behaviors will appear to change across survey waves as the respondent becomes more comfortable with and trusting of the survey experience*

To provide preliminary evidence regarding these propositions we utilize two CPS items. First, for those who have completed exactly 12 years of formal schooling, the CPS basic monthly survey asks whether respondents completed high school by earning a regular diploma or by obtaining a General Educational Development (GED) credential. Because dropping out of high school is non-normative and potentially stigmatizing, these two propositions will suggest that members of the treatment group (those in "month in sample" two through eight) will report different rates of having earned GEDs than members of the control group (those in "month in sample" one). The two propositions differ, however, in the empirical expectation about the direction of these differences. Second, we use an item from the December Supplement to the CPS which collects information about food security. This item ascertains whether respondents have run short of money for food over the preceding 12 months. Again, these two propositions suggest that treatment group members (those in "month in sample" five through eight) and control group members (those in "month in sample" one through four) will differ with respect to how often they offer this potentially stigmatizing response.



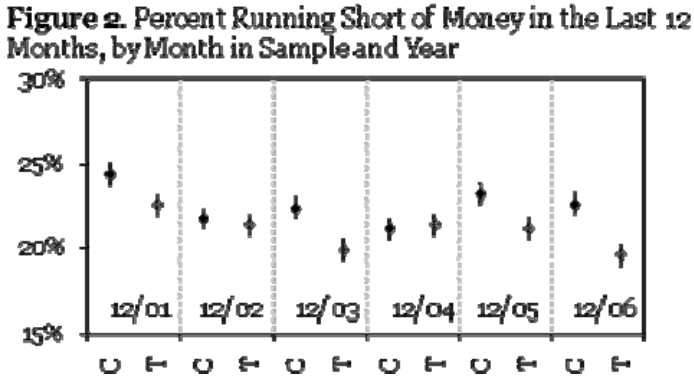


Figure 1 gives the percent completing high school via a GED (among those with exactly 12 years

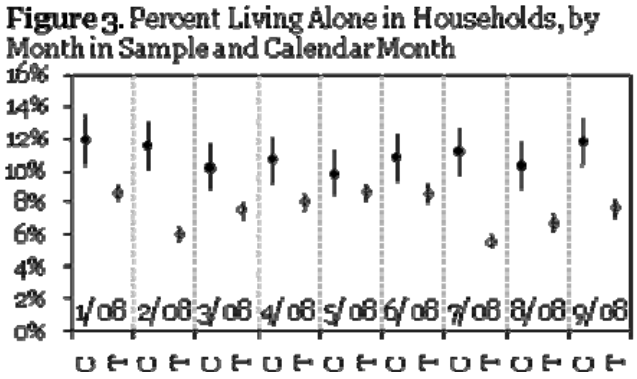
of schooling), disaggregated by treatment (T) vs. control group (C) and calendar month (Jan. 2008 through Sept. 2008). Dots and circles represent point estimates for the control and treatment groups, respectively, and the attached line segments give 95% confidence intervals. Figure 2 is similarly constructed, and gives the percent running short of money for food, disaggregated by treatment vs. control group and calendar month.

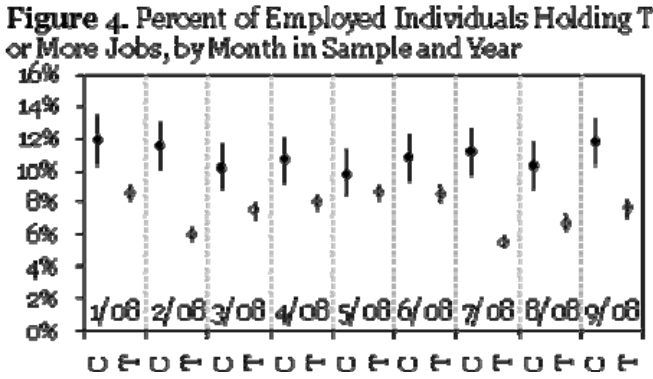
Both figures provide support for Proposition 3. In every calendar month, we observe that respondents are less likely to report having earned a GED if they previously participated in the CPS; in all but one month, this difference is statistically significant. In five of six Decembers, we observe that respondents are less likely to be running short of food money if they previously participated in the CPS; in four of the five months these differences are statistically significant. Beyond their statistical significance, the differences that we observe are typically substantial in magnitude. In January of 2008, for example, respondents in the treatment group were roughly 3.5 percentage points less likely to say that they completed high school via a GED, a difference that amounts to a nearly 30 percent decline relative to the control group.

Proposition 5: *Attitudes and behaviors will appear to change across survey waves as respondents learn to manipulate a survey instrument in order to minimize their burden.*

To provide preliminary evidence regarding this proposition we utilize two CPS items. First, we consider a measure of whether the respondent reports that anyone else lives in their household. Because respondents receive several additional questions about each member of their household, we hypothesize that panel conditioning will lead treatment group members (those in "month in sample" two through eight) to report having fewer people in their households than members of the control group (those in "month in sample" one). Second, we use an indicator of whether employed respondents hold more than one job. The basic month survey asks for details about each of respondents' jobs—hours worked, occupation, and so forth. Here, this proposition

implies that treatment group members will be less likely than control group members to hold two or more jobs.





Figures 3 and 4 give the percent living alone and the percent holding two or more jobs, respectively; both figures are disaggregated by treatment (T) vs. control (C) group and calendar month (Jan. 2008 through Sept. 2008). Both figures provide support for this proposition. In every calendar month, we observe that respondents are less likely to report that someone else lives in their household if they previously participated in the CPS; this difference is statistically significant in all nine months. In seven of nine months, we observe that employed respondents are less likely to hold two or more jobs if they previously participated in the CPS. Here again, the differences between treatment and control groups are usually substantial in magnitude. In January of 2008, for example, employed treatment group members were about 1/3 less likely to hold multiple jobs.

Again, by the annual meetings we intend to have developed results for all of the propositions, using data from the CPS, GSOEP, and GSS.

References

Bailar, B. A. 1975. "Effects of Rotation Group Bias on Estimates from Panel Surveys." *Journal of the American Statistical Association* 70:23-30.

Bailar, Barbara A. 1989. "Information Needs, Surveys, and Measurement Errors." Pp. 1-24 in *Panel Surveys*, edited by D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh. New York: Wiley.

Bartels, Larry M. 1999. "Panel Effects in the American National Election Studies." *Political Analysis* 8:1-20.

Basso, M. R., F. D. Carona, N. Lowery, and B. N. Axelrod. 2002. "Practice effects on the WAIS-III across 3-and 6-month intervals." *Clinical Neuropsychologist* 16:57-63.

Basso, M. R., N. Lowery, C. Ghormley, and R. A. Bornstein. 2001. "Practice effects on the Wisconsin Card Sorting Test - 64 card version across 12 months." *Clinical Neuropsychologist* 15:471-478.

Battaglia, Michael P., Elizabeth Zell, and Pam Ching. 1996. “Can Participating in a Panel Sample Introduce Bias into Trend Estimates?” National Immunization Survey Working Paper. Available at http://www.cdc.gov/nis/pdfs/estimation_weighting/battaglia1996b.pdf. Washington, D.C.: National Cenrter for Health Statistics.

Bishop, George F., Alfred J. Tuchfarber, and Robert W. Oldendick. 1986. "Opinions on fictitious issues: The pressure to answer survey questions." *Public Opinion Quarterly* 50:240-250.

Borle, S., U. M. Dholakia, S. S. Singh, and R. A. Westbrook. 2007. "The impact of survey participation on subsequent customer behavior: An empirical investigation." *Marketing Science* 26:711-726.

Bridge, R. G., L. G. Reeder, D. Kanouse, D. R. Kinder, V. T. Nagy, and C. M. Judd. 1977. "Interviewing Changes Attitudes - Sometimes." *Public Opinion Quarterly* 41:56-64.

Campbell, D. T. and J. C. Stanley. 1966. *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally.

Campbell, Donald T. 1957. "Factors Relevant to the Validity of Experiments in Social Settings." *Psychological Bulletin* 54:297-312.

Chandon, P., V. G. Morwitz, and W. J. Reinartz. 2004. "The short- and long-term effects of measuring intent to repurchase." *Journal of Consumer Research* 31:566-572.

—. 2005. "Do intentions really predict behavior? Self-generated validity effects in survey research." *Journal of Marketing* 69:1-14.

Clausen, Aage R. 1968. "Response Validity: Vote Report." *The Public Opinion Quarterly* 32:588-606.

Clinton, Joshua D. 2001. “Panel Bias from Attrition and Conditioning: A Case Study of the Knowledge Networks Panel.” Unpublished manuscript, retreived from http://www.princeton.edu/~clinton/WorkingPapers/C_WP2001.pdf on April 5, 2006. Department of Political Science, Stanford University.

Corder, Larry S. and Daniel G. Horvitz. 1989. "Panel Effects in the National Medical Care Utilization and Expenditure Survey." Pp. 304-318 in *Panel Surveys*, edited by D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh. New York: Wiley.

Crespi, Leo P. 1948. " The Interview Effect in Polling." *Public Opinion Quarterly* 12:99-111.

De Amici, Donatella , Catherine Klersy, Felice Ramajoli, Loretta Brustia, and Pierluigi Politi. 2000. "Impact of the Hawthorne Effect in a Longitudinal Clinical Study: The Case of Anesthesia." *Controlled Clinical Trials* 21:103-114.

Deming, W. Edwards and Frederick F. Stephan. 1940. "On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known." *The Annals of Mathematical Statistics* 11:427-444.

Dholakia, U. M. and V. G. Morwitz. 2002. "The scope and persistence of mere-measurement effects: Evidence from a field study of customer satisfaction measurement." *Journal of Consumer Research* 29:159-167.

Dillman, D. A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley.

Duncan, G. J. and G. Kalton. 1987. "Issues of Design and Analysis of Surveys Across Time." *International Statistical Review* 55:97-117.

Fazio, R. H. 1989. "On the Power and Functionality of Attitudes: The Role of Attitude Accessibility." Pp. 153-179 in *Attitude Structure and Function*, edited by A. R. Pratkanis, S. J. Breckler, and A. G. Greenwald.

Fazio, R. H., D. M. Sanbonmatsu, M. C. Powell, and E R. Kardes. 1986. "On the Automatic Activation of Attitudes." *Journal of Personality and Social Psychology* 50:229-238.

Feldman, J. M. and J. G. Lynch. 1988. "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior." *Journal of Applied Psychology* 73:421-435.

Fitzsimons, G. J. and S. G. Moore. 2008. "Should we ask our children about sex, drugs and rock & roll? Potentially harmful effects of asking questions about risky behaviors." *Journal of Consumer Psychology* 18:82-95.

Fitzsimons, G. J., J. C. Nunes, and P. Williams. 2007. "License to sin: The liberating role of reporting expectations." *Journal of Consumer Research* 34:22-31.

Fitzsimons, G. J. and P. Williams. 2000. "Asking questions can change choice behavior: Does it do so automatically or effortfully?" *Journal of Experimental Psychology-Applied* 6:195-206.

Fowler, Floyd J. Jr. 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks, CA: Sage Publications.

Godin, G., P. Sheeran, M. Conner, and M. Germain. 2008. "Asking questions changes behavior: Mere measurement effects on frequency of blood donation." *Health Psychology* 27:179-184.

Goldberg, T. E., R. S. Goldman, K. E. Burdick, A. K. Malhotra, T. Lencz, R. C. Patel, M. G. Woerner, N. R. Schooler, J. M. Kane, and D. G. Robinson. 2007. "Cognitive improvement after treatment with second-generation antipsychotic medications in first-episode schizophrenia - Is it a practice effect?" *Archives of General Psychiatry* 64:1115-1122.

Granberg, D. and S. Holmberg. 1992. "The Hawthorne Effect in Election Studies - The Impact of Survey Participation on Voting." *British Journal of Political Science* 22:240-247.

Greenwald, A. G., C. G. Carnot, R. Beach, and B. Young. 1987. "Increasing Voting-Behavior by Asking People if They Expect to Vote." *Journal of Applied Psychology* 72:315-318.

Grindstaff, T. L., K. E. J. Christiano, A. M. Broos, D. A. Straub, N. S. Darr, and K. A. Westphal. 2006. "Assessment of a practice effect in serial sensory organization testing scores of healthy adults." *Perceptual and Motor Skills* 102:379-386.

Groves, R. M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley.
Hansen, Morris H., William N. Hurwitz, Harold Nisselson, and Joseph Steinberg. 1955. "The Redesign of the Census Current Population Survey." *Journal of the American Statistical Association* 50:701-719.
Hausknecht, J. P., J. A. Halpert, N. T. Di Paolo, and M. O. M. Gerrard. 2007. "Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability." *Journal of Applied Psychology* 92:373-385.
Hernandez, Lyla M., Jane S. Durch, Dan G. Blazer II, and Isabel V. Hoverman. 1999. *Gulf War Veterans: Measuring Health*. Committee on Measuring the Health of Gulf War Veterans, Division of Health Promotion and Disease Prevention Institute of Medicine. Washington, D.C.: National Academies Press.
Holt, D. 1989. "Panel Conditioning: Discussion." Pp. 340-347 in *Panel Surveys*, edited by D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh. New York: Wiley.
Izrael, David, David C. Hoaglin, and Michael P. Battaglia. 2000. "A SAS Macro for Balancing a Weighted Sample." in *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference*. Cary: SAS Institute, Inc.
—. 2004. "To Rake or Note to Rake is Not the Question Anymore with the Enhanced Raking Macro." in *Proceedings of the Twenty-Ninth Annual SAS Users Group International Conference*. Cary: SAS Institute, Inc.
Janiszewski, C. and E. Chandon. 2007. "Transfer-appropriate processing response fluency, and the mere measurement effect." *Journal of Marketing Research* 44:309-323.
Johnson, B. F., K. Hoch, and J. Johnson. 1991. "Variability in Psychometric Test-Scores - The Importance of the Practice Effect in Patient Study Design." *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 15:625-635.
Kalton, Graham and Constance F. Citro. 2000. "Panel Surveys: Adding the Fourth Dimension." Pp. 36-53 in *Researching Social and Economic Change*, edited by D. Rose. London: Routledge.
Kardes, F. R., C. T. Allen, and M. J. Pontes. 1993. "Effects of Multiple Measurement Operations on Consumer Judgment - Measurement Reliability or Reactivity." *Advances in Consumer Research* 20:280-283.
Kraut, Robert E. and John B. McConahay. 1973. "How Being Interviewed Affects Voting: An Experiment." *The Public Opinion Quarterly* 37:398-406.
Kulick, James A., Chen-Lin Kulick, and Robert L. Bangert. 1984. "Effects of Practice on Aptitude and Achievement Test Scores." *American Educational Research Journal* 21:435-447.
Lazarsfeld, Paul F. 1940. ""Panel" Studies." *The Public Opinion Quarterly* 4:122-128.
—. 1941. "Repeated Interviews as a Tool for Studying Changes in Opinion and Their Causes." *American Statistical Association Bulletin* 2:3-7.
Lemay, S., M. A. Bedard, I. Rouleau, and P. L. G. Tremblay. 2004. "Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects." *Clinical Neuropsychologist* 18:284-302.
Levav, J. and G. J. Fitzsimons. 2006. "When questions change behavior - The role of ease of representation." *Psychological Science* 17:207-213.
Little, R. J. A. 1993. "Post-Stratification: A Modeler's Perspective." *Journal of the American Statistical Association* 88:1001-1012.

Mann, C. B. 2005. "Unintentional voter mobilization: Does participation in preelection surveys increase voter turnout?" *Annals of the American Academy of Political and Social Science* 601:155-168.

McCormick, Maryann K., Derrick M. Butler, and Rajendra P. Singh. 1992. "Investigating Time in Sample Effect for the Survey of Income and Program Participation." Pp. 554-559 in *Proceedings of the Survey Research Methods Section of the American Statistical Association*.

Millar, M. G. and A. Tesser. 1986. "Thought-Induced Attitude-Change - The Effects of Schema Structure and Commitment." *Journal of Personality and Social Psychology* 51:259-269.

Morwitz, V. G. 2005. "The effect of survey measurement on respondent behaviour." *Applied Stochastic Models in Business and Industry* 21:451-455.

Morwitz, V. G. and G. J. Fitzsimons. 2004. "The mere-measurement effect: Why does measuring intentions change actual behavior?" *Journal of Consumer Psychology* 14:64-74.

Morwitz, V. G., E. Johnson, and D. Schmittlein. 1993. "Does Measuring Intent Change Behavior?" *Journal of Consumer Research* 20:46-61.

Nancarrow, C. and T. Cartwright. 2007. "Online access panels and tracking research: The conditioning issue." *International Journal of Market Research* 49:573-594.

O' Sullivan, I., S. Orbell, T. Rakow, and R. Parker. 2004. "Prospective research in health service settings: Health psychology, science and the 'Hawthorne' effect." *Journal of Health Psychology* 9:355-359.

Pennell, Steven G. and James N. Lepkowski. 1992. "Panel Conditioning Effects in the Survey of Income and Program Participation." Pp. 566-571 in *Proceedings of the Survey Research Methods Section of the American Statistical Association*.

Porst, R. and K. Zeifang. 1987. "A Description of the German General Social Survey Test-Retest Study and a Report on the Stabilities of the Sociodemographic Variables." *Sociological Methods & Research* 15:177-218.

Powell, M. C. and R. H. Fazio. 1984. "Attitude Accessibility as a Function of Repeated Attitudinal Expression." *Personality and Social Psychology Bulletin* 10:139-148.

Richardson, F. and E. Robinson. 1921. "Effects of Practice Upon the Scores and Predictive Value of the Alpha Intelligence Examination." *Journal of Experimental Psychology* 4:300-317.

Schuman, Howard and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.

Shack-Marquez, J. 1986. "Effects of Repeated Interviewing on Estimation of Labor-Force Status." *Journal of Economic and Social Measurement* 14:379-398.

Sherman, S. J. 1980. "On the Self-Erasing Nature of Errors of Prediction." *Journal of Personality and Social Psychology* 39:211-221.

Silberstein, Adriana R. and Curtis A. Jacobs. 1989. "Symptoms of Repeated Interview Effects in the Consumer Expenditure Survey." Pp. 289-303 in *Panel Surveys*, edited by D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh. New York: Wiley.

Simmons, C. J., B. A. Bickart, and J. G. Lynch. 1993. "Capturing and Creating Public-Opinion in Survey-Research." *Journal of Consumer Research* 20:316-329.

Smith, Jennifer K., Alan S. Gerber, and Anton Orlich. 2003. "Self-Prophecy Effects and Voter Turnout: An Experimental Replication." *Political Psychology* 24:593-604.

Sobol, Marion. 1959. "Panel Mortality and Panel Bias." *Journal of the American Statistical Association* 54:52-68.

Solon, G. 1986. "Effects of Rotation Group Bias on Estimation of Unemployment." *Journal of Business & Economic Statistics* 4:105-109.
Spangenberg, E. R., A. G. Greenwald, and D. E. Sprott. 2008. "Will you read this article's abstract? Theories of the question-behavior effect." *Journal of Consumer Psychology* 18:102-106.
Spangenberg, E. R. and C. Obermiller. 1996. "To Cheat or Not to Cheat: Reducing Cheating by Requesting Self-Prophecy." *Marketing Education Review* 6:95-103.
Spangenberg, E. R., D. E. Sprott, B. Grohmann, and R. J. Smith. 2003. "Mass-communicated prediction requests: Practical application and a cognitive dissonance explanation for self-prophecy." *Journal of Marketing* 67:47-62.
Spangenberg, Eric R. 1997. "Increasing Health Club Attendance Through Self-Prophecy." *Marketing Letters* 8:23-31.
Sprott, D. E., E. R. Spangenberg, and A.W. Perkins. 1999. "Two More Self-Prophecy Experiments." Pp. 621-626 in *Advances in Consumer Research*, edited by L. Scott and E. J. Amould. Provo, UT: Association for Consumer Research.
Sprott, David E., Eric R. Spangenberg, Lauren G. Block, Gavan J. Fitzsimons, Vicki G. Morwitz, and Patti Williams. 2006. "The question-behavior effect: What we know and where we go from here." *Social Influence* 1:128-137.
Stephan, Frederick F. 1942. "An Iterative Method of Adjusting Sample Frequency Tables When Expected Marginal Totals are Known." *Annals of Mathematical Statistics* 13:166-178.
Sturgis, P., N. Allum, and I. Brunton-Smith. 2009. "Attitudes Over Time: The Psychology of Panel Conditioning " Pp. 113-126 in *Methodology of Longitudinal Surveys*, edited by P. Lynn. New York: Wiley.
Sudman, Seymour, Norman M. Bradburn, and Norbert Schwarz. 1996. *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publications.
Sutton, S, G Bickler, J Sancho-Aldridge, and G Saidi. 1994. "Prospective study of predictors of attendance for breast screening in inner London." *Journal of Epidemiology and Community Health* 48:65-73.
Swan, J. E., I. F. Trawick, and M. G. Carroll. 1981. "Effect of Participation in Marketing-Research on Consumer Attitudes toward Research and Satisfaction with a Service." *Journal of Marketing Research* 18:356-363.
Tesser, A. 1978. "Self-Generated Attitude Change." Pp. 289-338 in *Advances in Experiemental Social Psychology*, vol. 11, edited by L. Berkowitz. New York: Academic Press.
Toh, R. S., E. Lee, and M. Y. Hu. 2006. "Social desirability bias in diary panels is evident in panelists' behavioral frequency." *Psychological Reports* 99:322-334.
Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinsk. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
Traugott, M. W. and J. P. Katosh. 1979. "Response Validity in Surveys of Voting-Behavior." *Public Opinion Quarterly* 43:359-377.
U.S. Bureau of Labor Statistics. 2000. *Current Population Survey: Design Methodology*. Technical Paper 63. Washington, D.C.: U.S. Bureau of the Census.
Underwood, M. R., S. Parsons, S. M. Eldridge, A. E. Spencer, and G. S. Feder. 2006. "Asking older people about fear of falling did not have a negative effect." *Journal of Clinical Epidemiology* 59:629-634.

Voogt, R. J. J. and H. Van Kempen. 2002. "Nonresponse bias and stimulus effects in the Dutch National Election Study." *Quality & Quantity* 36:325-345.

Wagner, Gert G., Joachim R. Frick, and Jürgen Schupp. 2007. "The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements." *Schmollers Jahrbuch* 127:139-169.

Wang, Kevin, David Cantor, and Adam Safir. 2000. "Panel Conditioning in a Random Digit Dial Survey." Pp. 822-827 in *Proceedings of the Survey Research Methods Section of the American Statistical Association*.

Waterton, Jennifer and Denise Lievesley. 1989. "Evidence of Conditioning Effects in the British Social Attitudes Panel Survey." Pp. 319-339 in *Panel Surveys*, edited by D. Kasprzyk, G. J. Duncan, G. Kalton, and M. P. Singh. New York: Wiley.

Weir, D. R. and J. P. Smith. 2007. "Do panel surveys really make people sick? A commentary on Wilson and Howell (60:11,2005,2623-2627)." *Social Science & Medicine* 65:1071-1077.

Williams, P., G. J. Fitzsimons, and L. G. Block. 2004. "When consumers do not recognize "benign" intention questions as persuasion attempts." *Journal of Consumer Research* 31:540-550.

Williams, Patti, Lauren G. Block, and Gavan J. Fitzsimons. 2006. "Simply asking questions about health behaviors increases both healthy and unhealthy behaviors." *Social Influence* 1:117 - 127.

Williams, W. H. and C. L. Mallows. 1970. "Systematic Biases in Panel Surveys Due to Differential Nonresponse." *Journal of the American Statistical Association* 65:1338-1349.

Willis, G., M. Sirken, and G. Nathan. 1994. *The Cognitive Aspects of Responses to Sensistive Survey Questions*. Working Paper Series, No. 9. Hyattsville, MD: National Center for Health Statistics.

Wilson, S. E. and B. L. Howell. 2005. "Do panel surveys make people sick? US arthritis trends in the Health and Retirement Study." *Social Science & Medicine* 60:2623-2627.

Wilson, S. and B. L. Howell. 2007. "Disease prevalence and survey design effects: A response to Weir and Smith." *Social Science & Medicine* 65:1078-1081.

Wilson, Timothy D., Suzanne J. LaFleur, and D. Eric Anderson. 1996. "The Validity and Consequences of Verbal Reports About Attitudes." Pp. 91-114 in *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, edited by N. Schwarz and S. Sudman. New York: Jossey-Bass Publishers.

Wu, C. and R.R. Sitter. 2001. "A Model-Calibration Approach to Using Complete Auxiliary Information from Survey Data." *Journal of the American Statistical Association* 96:185-193.

Yalch, R. F. 1976. "Pre-Election Interview Effects on Voter Turnout." *Public Opinion Quarterly* 40:331-336.

Zaller, J. and S. Feldman. 1992. "A Simple Theory of The Survey Response - Answering Questions Versus Revealing Preferences." *American Journal of Political Science* 36:579-616.