Comparing Web Panel Samples vs. Non-Panel Samples of Medical Doctors

LinChiat Chang[1] & Jeremy Brody[2]
[1]LinChiat Chang Consulting LLC, San Francisco, CA
[2]Kantar Health, New York, NY

Abstract

Although web panels are widely used in market research, there are persistent concerns about panel conditioning, panel attrition, and self-selection biases. We compared web panel samples to "fresh" random samples recruited via telephone across 3 medical specialties: neurology (n=167 web vs. n=97 phone), pulmonology (n=83 web vs. n=68 phone), and pediatrics (n=56 web vs. n=60 phone). All physicians, regardless of sample source, completed the surveys via the Internet. Comparative analyses were conducted on 163 measures of practice characteristics, treatment choices, attitudes and perceptions. Focus was placed on whether key research conclusions differed by sample source. Few significant differences emerged between the panel samples and the non-panel samples, and no systematic bias was manifested by the panel samples. Methodological limitations of this study are addressed.

Key Words: web surveys; physician web panels; healthcare market research.

Background & Objective

Although physician web panels are routinely used in healthcare market research, there are persistent concerns about the quality of panel samples. Concerns about panels of medical doctors tend to revolved around 3 main themes: panel conditioning, panel attrition, and self-selection biases. Panel conditioning occurs when prior surveys change respondents' behaviors or change the way respondents answer questions on subsequent surveys. This reactive effect of prior surveys on later responses can take different forms: repeated surveys may raise consciousness about specific domains and affect respondents actual choices and opinions; repeated surveys on the same topics may crystallize attitudes and/or result in more extreme attitudes; and unmotivated panel members become increasingly savvy about how to respond in order to finish the survey quickly to earn substantial cash honorariums in the least amount of time possible.

All panels lose members across time. Panel attrition is not a problem if attrition is random across all panel members. But a panel that is representative at the outset can deteriorate in quality when it loses panel members that disproportionately hold certain characteristics. For example, if doctors with higher patient loads are more likely to quit the panel than those with lower patient loads, or if panel members who are less satisfied with the level of cash honorariums are more likely to quit the panel than those who are happy with the current level of compensation. Finally, even if each panel sample is carefully shaped to match population characteristics, self-selection biases can undermine the representativeness of panel samples because doctors who agree to serve on panels may

have attitudes, preferences, lifestyles, and experiences that are different from doctors who are not on the panel.

Past Research on Panel Issues

Concerns about the quality of panel samples are grounded in decades of research. For example, panel effects have been documented on studies of consumer behavior. Longer time enrolled in the Consumer Expenditure Interview Survey panel was related to more purchases in house furnishing and apparel, but fewer purchases of vehicles, public transportation, and some utilities (Silberstein & Jacobs, 1989). Asking respondents once about their intent to buy any consumer product increased likelihood to buy, but repeatedly surveying about intent to buy significantly reduced purchasing rates (Morwitz, Johnson & Schmittlein, 1993). Respondents over-reported their household expenditures on the first day of the Consumer Expenditure Survey Diary panel, but under-reported expenditure on subsequent days (Pearl, 1979). Participation in the MSU Consumer Panel on food purchases over multiple years made respondents more price-conscious about food purchases, as well as more conscious about the variety and adequacy of different nutrition elements in their diet (Quackenbush & Shaffer, 1960).

Panel effects have also been demonstrated on studies of health and healthcare utilization. Respondents over-reported illnesses in initial interviews on the California Health Survey, and under-reported illnesses in subsequent interviews; when matched against independent physician records, subsequent waves drastically underestimated number of people with illnesses (Mooney, 1962). Respondents subjected to 4 rounds of data collection on the National Medical Care Expenditure Survey panel produced more coherent, accurate, and complete data than respondents subjected to 5 rounds of data collection in the same time period (Cohen & Burt, 1985). This finding demonstrates that once passed a threshold, over-surveying could significantly reduce data quality.

In the domain of political behavior, longer tenure on a panel survey of election behavior is associated increased likelihood to vote, as well as more accurate responses on voting behavior when matched against official voting records (Traugott & Katosh, 1979). In another study, respondents became more politicized after first interview on the British Social Attitudes Panel Survey, that is, they provided more extreme answers and fewer "don't know" answers on subsequent surveys. Across time, respondents also became more comfortable expressing less desirable attitudes, e.g. admitting that oneself is “very racially prejudiced” (Waterton & Lievesley, 1989).

Longer time enrolled in a national crime survey panel is related to declining crime victimization rates. The hypothesized reason for this trend is that panel members become more aware about crime and take more precautionary measures in their daily lives. Hence, they are less likely to become victims of crime (Lehnen & Reiss, 1978; Woltman & Bushery, 1977). In the monthly Current Population Survey where respondents are repeatedly surveyed on their employment status, unemployment rate was significantly higher in the first wave of survey than subsequent waves (Bailar, 1989). There are higher variances among panel respondents than non-panel respondents in reported marital satisfaction, because a marriage panel forces respondents to think through their marital experiences, thereby polarizing both positive and negative experiences (Veroff, Hatchett & Douvan, 1992). Longer tenure on panel survey of household savings was associated with higher data quality, due to attrition of unmotivated respondents and increased accuracy of those who remained (Ferber, 1964).

On the other hand, some research have found no or weak panel effects. For example, Cordell and Rahmel (1962) found that participating in Nielsen surveys on media use did not alter later reports of media use. Likewise, Himmelfarb and Norris (1987) found that being interviewed on a wide range of topics did not alter people's subsequent reports of mental health, physical health, self-esteem, social support, or life events experienced. Willson and Putnam (1982) found in a meta-analysis that answering questions caused attitudes toward objects to become slightly more positive, but these effects were quite small and inconsistent across studies. Even though some studies suggest that interviewing people on a particular topic may induce them to become more cognitively engaged in that topic (Bridge, Reeder, Kanouse, Kinder, Nagy, & Judd 1977; Granberg & Holmberg 1991), others did not replicate the same effect (Mann 2005). Investigations into panel attrition have also turned up little cause for concern. Most studies found little or no sample composition changes attributable to panel attrition (e.g., Zagorsky & Rhoton 1999; Fitzgerald, Gottschalk, & Moffitt 1998a, 1998b; Falaris & Peters 1998; Clinton 2001).

Interestingly, regular experience answering survey questions could enhance response quality, because motivated panel members could gain a better understanding of the response process and how to provide more accurate data, which could potentially improve overall data quality (Chang & Krosnick 2009; Donovan & Radosevich 1999; Smith, Branscombe, & Bormann 1988). Also, panel members may become especially self-aware and introspective about their thoughts, attitudes, emotions, and behaviors, further improving their ability to later report on those phenomena accurately (Menard 1991). Consistent with this reasoning, research on panel surveys has shown that people's answers to attitude questions become increasingly reliable as they gain more experience responding to them (Jagodzinski, Kuhnel, & Schmidt 1987).

Furthermore, it is possible that the published research literature is biased towards studies that uncovered panel issues; thus there could be other research studies in which no notable panel conditioning effects was found. Although panels are susceptible to conditioning effects, conditioning effects do not always exist. More importantly, these past research studies showing panel effects or lack thereof were all based on consumer panels, and do not speak to highly specialized physician panels. We should not generalize findings from consumer panels to physician panels, because physicians receive the same fundamental training and their treatment choices are restricted by the same clinical guidelines, thus the dynamics underlying variation in survey responses in physician samples may be different from consumer samples.

Current Investigation

In short, the available research literature is based almost entirely on consumer panels, and is silent on potential biases in more specialized panels such as a physician web panel. To this end, we conducted our own analyses of data from studies where we had drawn physician samples from both the TNSjstreet[1] web panel as well as telephone recruitment by independent telephone survey firms. All physicians, regardless of sample source, completed the surveys online. These dual-sample studies allowed us to determine

[1] The Jstreet web panel was founded in 1999 focusing only on physicians located in the United States. The panel was sold to NFO in August 2002, and NFO was subsequently acquired by TNS in 2003. In 2003, the panel built on its 23,000 U.S. physicians by expanding to Europe. It was merged with the AllGlobal physician panel in 2009; and the current panel size is estimated at 250,000 physicians worldwide. All research reported in this paper is based on data collected in 2003.

whether and to what extent web panel samples of medical doctors might differ from "fresh" non-panel samples. The main question underlying this investigation was whether research conclusions and recommendations would be different if the surveys were based on the web panel samples or based on the "fresh" non-panel samples.

Study One

This study was conducted to assess market share within a hypothetical scenario in the treatment of asthma and COPD (Chronic Obstructive Pulmonary Disease). The sampling frame was a target list of high-prescribing physicians provided by the pharmaceutical company that commissioned this research. It was necessary to recruit "fresh" samples over the telephone to supplement the web panel sample because the web panel did not yet have sufficient numbers of physicians to meet the sample size required for analyses. In other words, this research was conducted to meet specific business objectives; it was not conducted to test for differences in sample quality. Nonetheless, the data afforded us the opportunity to investigate whether these two sample sources would yield systematically different results. The panel sample consisted of 83 pulmonologists/allergists and 56 pediatricians; while the telephone-recruited fresh sample consisted of 68 pulmonologists/allergists and 60 pediatricians. Both samples completed the survey via the Internet. The survey length was 30 minutes on average, and the physicians from both sample sources were compensated at the same level of honorarium. Due to the proprietary nature of these research findings, all results reported below are stripped of substantive content that are irrelevant to methodology. All data were collected in 2003.

Results

As shown in Table 1, there was no difference between the two sample sources on most practice characteristics. The only difference that emerged as significant was in number of years in clinical practice post residency among Pulmonologists/Allergists: the non-panelists had a few more years in clinical practice than those who were already serving as TNSjstreet panelists, t=2.61, p<.01.

Table 1: Comparing Samples on Clinical Practice Characteristics

	Pulmonologists/ Allergists		*Pediatricians*	
	PANEL	*NON-PANEL*	*PANEL*	*NON-PANEL*
Percent time spent in direct patient care	95%	95%	96%	97%
Number of years in clinical practice	17*	14	15	16
Number of COPD/asthma patients per month	180	179	78	77

* Significant difference from estimate based on non-panel sample, t=2.61, p<.01.

As shown in Table 2, there was no difference between the two sample sources on all patient characteristics.

Table 2: Comparing Samples on Patient Characteristics

	Pulmonologists/ Allergists		*Pediatricians*	
	PANEL	*NON-PANEL*	*PANEL*	*NON-PANEL*
Percent on prescription drug for COPD/asthma	85%	88%	63%	67%
Percent of COPD patients with no asthma	23%	18%	96%	96%
Percent of asthma patients with no asthma	77%	82%	4%	4%

The main research objective of this study was to assess market share in response to a hypothetical scenario in the future where a new medication would become available for the treatment of asthma and COPD. Hence, it was critical to explore whether the market share estimates would differ significantly between the two sample sources. As shown in Table 3, there was no significant difference between the two sample sources on the market share projections of pulmonologists and allergists. However, a few significant differences emerged among pediatricians, such that panel pediatricians appeared to be more cautious with high dosing of Product A (t=2.99, p<.01) and more comfortable with low dosing of Product B (t=3.18, p<.01), compared to pediatricians in the "fresh" sample. Examining the projected prescription volumes within Product A alone, it is apparent that non-panelists expected to prescribe about the same amount of high dose (29%) vs. low dose (30%) of Product A, whereas the panelists expected to prescribe only about half the amount of high dose (19%) vs. low dose (37%) of Product A. The ratios of projected prescription volumes within Product B alone did not exhibit the same dramatic gap, but slight trends suggest that non-panel pediatricians were less conservative than panel pediatricians in Product B dosing as well.

Table 3: Comparing Samples on Market Share Projections
(Because combination therapy was permitted, columns do not sum to 100%)

	Pulmonologists/ Allergists		*Pediatricians*	
	PANEL	*NON-PANEL*	*PANEL*	*NON-PANEL*
Product A (low dose)	17%	13%	37%	30%
Product A (moderate dose)	46%	51%	17%	20%
Product A (high dose)	14%	18%	19%*	29%
Product B (low dose)	15%	14%	27%*	18%
Product B (moderate dose)	7%	6%	6%	4%
Product B (high dose)	17%	18%	5%	4%
Product C	28%	27%	27%	24%
Other inhaled corticosteroids	5%	6%	2%	1%
Other maintenance medications	16%	16%	5%*	2%

* Significant difference from estimates based on non-panel sample, t=2.99, p<.01 for Product A (high dose); t=3.18, p<.01 for Product B (low dose); t=2.55, p<.01 for other maintenance medications.

Summary

Only 4 significant differences emerged from the 40 measures that afforded comparison between the two sample sources. Although no systematic difference was found on

clinical practice characteristics or patient characteristics, one noteworthy finding was the general tendency of panel pediatricians to be more cautious with high dosing compared to non-panel pediatricians. On the other hand, there was no difference between panel and non-panel samples in terms of market share projections among pulmonologists and allergists.

Study Two

This study was conducted to provide data for a market segmentation among high prescribers treating multiple sclerosis. The sampling frame was again a target list of high-prescribing physicians provided by the pharmaceutical company that commissioned this research. Similar to the circumstances in Study One, it was necessary to recruit "fresh" samples over the telephone to supplement the web panel sample because the web panel did not yet have sufficient numbers of physicians to meet the sample size required for segmentation analyses. Thus this research was conducted to address specific business objectives; it was not conducted for this methodological comparison. Nonetheless, the data afforded us the opportunity to investigate whether these two sample sources would yield systematically different results. The panel sample consisted of 167 neurologists; while the telephone-recruited fresh sample consisted of 97 neurologists. Both samples completed the survey via the Internet. The survey length was 45 minutes on average, and the non-panel neurologists were compensated at a 50% higher level of honorarium than panel neurologists. As with Study One, all results reported below are stripped of substantive content that are irrelevant to methodology. All data were collected in 2003.

Results

As shown in Table 4, there was no difference between the two sample sources on most practice characteristics. The only difference that emerged as significant was in total number of patients treated per month, whereby non-panel neurologists treated fewer patients per month than their counterparts who were TNSjstreet panelists, t=2.20, p<.01. However, this difference was inconsequential because the business focus of this research was on the treatment of patients with severe, relapsing forms of multiple sclerosis, which exhibited no significant difference between the two sample sources.

Table 4: Comparing Samples on Clinical Practice Characteristics

	PANEL	*NON-PANEL*
Number of years in clinical practice post residency	15	15
Percent time spent in direct patient care	97%	96%
Percent time spent conducting research	1%	3%
Number of patients personally treated per month	312*	277
Number of Multiple Sclerosis diagnosis per year	98	95
Number of patients with relapsing Multiple Sclerosis per year	70	69
Number of relapsing patients on Multiple Sclerosis therapy	65	62

* Significant difference from estimate based on non-panel sample, t=2.20, p<.01.

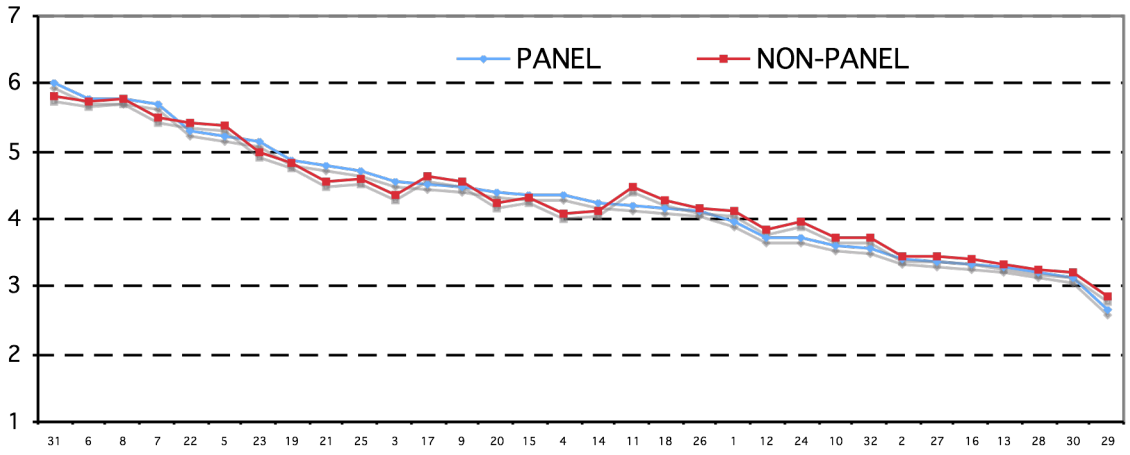
As shown in Table 5, there was no significant difference between the two sample sources on neurologists' choice of medications to treat less aggressive and more aggressive forms of multiple sclerosis.

Table 5: Comparing Samples on Treatment Choices

	PANEL	*NON-PANEL*
Treatment of less aggressive forms of multiple sclerosis		
• Product A	39%	34%
• Product B	13%	17%
• Product C	35%	31%
• Product D	13%	18%
Treatment of more aggressive forms of multiple sclerosis		
• Product A	18%	21%
• Product B	33%	31%
• Product C	12%	13%
• Product D	37%	35%

The neurologists were asked to indicate their likelihood to consider each one of 36 clinical risk factors when determining whether a newly diagnosed patient with a relapsing form of MS has a more aggressive form of the disease, using a 7-point likelihood scale. There was no significant difference between the two sample sources on all 36 clinical risk factors. In addition, the relative importance of those risk factors were comparable between sample sources, because the ratings were highly correlated with Pearson's r = .99, p<.001. This trend is exhibited in Figure 1.

Figure 1: Comparing Samples on Relative Impact of Clinical Risk Factors



The neurologists were also asked to rate their agreement with 76 attitudinal measures tapping their opinions on a wide range of domains such as drug dosing, drug administration, cost issues, long-term disease control, patient empowerment, aggressive vs. conservative therapy, interferon therapy, as well as their tendency to switch treatment regimens. Out of the 76 attitudes, only 5 yielded significant differences between the two sample sources. The Bonferroni correction is appropriate in this circumstance when we are conducting a series of so many tests of significance on variables that are not entirely independent of one another and based on data from the same set of respondents. Given that we have as many as 76 attitudinal variables, if we go on testing long enough we will

find a difference that reaches statistical significance simply by chance alone. Thus it is important to apply the family-wise error rate on this set of significance tests to ensure that we do not attach too much importance to a few significant results among a mass of non-significant ones (Bland, 2000; Holm, 1979). As expected, once the Bonferroni correction was applied, there was no significant difference between the two sample sources on all 76 attitudes. In addition, the relative ratings on those attitudes were highly correlated between the two sample sources, Pearson's $r = .98$, $p<.001$.

The primary business objective driving Study Two was to produce a market segmentation of high-prescribing neurologists treating multiple sclerosis. Thus it was important to ascertain whether segmentation results would be different between the two sample sources. Although it was not possible to undertake a full segmentation due to insufficient sample size within each sample source, it is possible to attempt the key steps of segmentation based on a subset of variables.

Table 6: Attitudinal Measures Sorted by High to Low Standard Deviation (Displaying only S.D.>1.50)

Attitude Item#	*PANEL*	*NON-PANEL*
2	1.93	1.98
11	1.81	1.88
20	1.79	1.66
62	1.73	1.69
7	1.72	1.71
50	1.65	1.61
73	1.64	1.57
18	1.59	1.66
69	1.58	1.60
23	1.57	1.58
76	1.53	1.56
61	1.53	1.50
19	1.51	1.56
46	1.50	1.43
4	1.50	1.41
30	1.50	1.32
39	1.49	1.42
49	1.48	1.50

Using the 76 attitudinal measures, we could identify the attitudes that produced the maximal differentiation among neurologists within each sample source, and then run some exploratory cluster analyses based on that subset of attitudes. Examination of the standard deviations of the 76 attitudinal measures within each sample source revealed a striking consistency – as shown in Table 6, the top dozen most differentiating attitudes with standard deviations above 1.50 were virtually identical between the two sample sources.

To reduce any redundancy between these attitudes, correlations and exploratory factor analyses were conducted within each sample to ascertain the need to combine certain variables into composite factors. Again, the results were strikingly consistent. In both samples, attitudes #2, 11, 62, and 50 clearly loaded on one factor, while attitudes #18 and 20 clearly loaded on another factor. Hence, these variables were combined to form their respective factors in both samples.

Finally, model-based exploratory cluster analysis was applied on these attitudinal measures within each sample. This technique was chosen because it can test for a variety of data models and apply maximum likelihood estimation and Bayesian criteria to identify the most optimal model and number of clusters. The best or most optimal model is selected according to BIC (Bayesian Information Criteria) for EM (Expectation Maximization) initialized by hierarchical clustering for parameterized Gaussian mixture models. In short, the model and number of clusters with the largest BIC would be identified as the best model. Again, the results were very consistent between the two sample sources – the best model produced by both sample sources was spherical with two equal volume clusters.

Summary

Only 1 significant difference emerged from the 123 measures that afforded comparison between the two sample sources. Ratings on long batteries of items were highly correlated between the two samples. The market segmentation results derived from exploratory cluster analysis were very consistent between the two samples. In short, there was no systematic difference that suggests different research conclusions would have been reached if one sample source had been used instead of the other.

Discussion

Comparative analyses were conducted on a total of 163 measures of practice characteristics, treatment choices, attitudes and perceptions. In particular, comparisons were made to ascertain whether the key research conclusions would have been different if one sample source was used instead of the other. Very few differences emerged from the comparison of panel vs. "fresh" random samples, and these few differences exhibited no convergent or systematic bias. Further, there was little indication that using one or the other sample source would have led researchers and their clients to different conclusions on the key research objectives. In short, evidence from this study demonstrates that survey estimates based on samples drawn from the TNSjstreet panel were very comparable to "fresh" non-panel samples.

One obvious limitation of the current investigation is that both studies were conducted in 2003. Hence, one might argue the panel was still relatively young in 2003, and panel conditioning or attrition effects or self-selection biases might have taken their toll since then. Analogous comparisons based on more recent samples are needed to address this concern.

Another limitation is the fact that both the panel samples and the non-panel samples probably share the same selection bias of medical doctors who are at least marginally interested in participating in market research in exchange for cash honoraria. The research conclusions that could be drawn from these two sample sources may not ultimately be representative of the entire population of physicians in these respective medical specialties. Nonetheless, this bias is consequential only if doctors who are willing to participate in market research are consistently and significantly different from doctors who would never participate in market research on the very specific measures included in each study. To date, we have not located any research that speaks to the extent of this bias and how it might impact generalizability of research conclusions. We look forward to future research that could inform this issue, and thereby aid in the development of weighting methodologies or other adjustment techniques that could alleviate this bias, if it does exist.

Acknowledgements

The authors wish to thank Patrick Foster for his assistance in locating all materials required for completion of this analysis, and Tony Burke, CEO of AllGlobal, for his support and approval to publish these research findings in the public domain. We also thank members of the Political Psychology Research Group (PPRG) at Stanford University for their helpful feedback and input on an earlier draft of this research paper.

References

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

Bland, Martin. 2000. *Introduction to Medical Statistics*. New York, NY: Oxford University Press, Inc.

Bridge, R. Gary, Leo G. Reeder, David Kanouse, Donald R. Kinder, Vivian T. Nagy, and Charles Judd. 1977. "Interviewing Changes Attitudes – Sometimes." *Public Opinion Quarterly* 41:57-64.

Chang, LinChiat and Jon A. Krosnick. 2009. "National Surveys via RDD Telephone Interviewing vs. the Internet: Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly 73*: 641-678.

Cordell, Warren N. and Henry A. Rahmel. 1962. "Are Nielson Ratings Affected by Non-cooperation, Conditioning, or Response Error?" *Journal of Advertising Research* 2:45-49.

Corder, Larry S. and Daniel G. Horvitz. 1989. "Panel effects in the National Medical Care Utilization and Expenditure Survey." In *Panel Surveys*, ed. Daniel Kasprzyk, Greg Duncan, Graham Kalton, and M.P. Singh. New York: John Wiley and Sons.

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

Donovan, John J. and David J. Radosevich. 1999. "A Meta-analystic Review of the Distribution of Practice Effect: Now You See It, Now You Don't." *Journal of Applied Psychology* 84:795-805.

Falaris, Evangelos M. and H. E. Peters. 1998. " Survey Attrition and Schooling Choices." *The Journal of Human Resources 33*:531-54.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998a. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *NBER Technical Working Papers.* National Bureau of Economic Research, Inc.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998b. "An Analysis of the Impact of Sample Attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics." *The Journal of Human Resources 33*:300-344.

Fitzsimons, Gavan J. and Vicki Morwitz. 1996. "The Effect of Measuring Intent on Brand Level Purchase Behavior." *Journal of Consumer Research 23*:1-11.

Granberg, Donald, and Soren Holmberg. 1991. "Self-reported Turnout and Voter Validation." *American Journal of Political Science 35*:448-459.

Greenwald, Anthony G., Catherine G. Carnot, Rebecca Beach, and Barbara Young. 1997. "Increasing Voting Behavior by Asking People if They Expect to Vote." *Journal of Applied Psychology* 72:315-318.

Himmelfarb, Samuel and Fran H. Norris. 1987. "An Examination of Testing Effects in a Panel Study of Older Persons." *Personality and Social Psychology Bulletin 13*:188-209.

Holm, Sture. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6:65-70.

Jagodzinski, Wolfgang, Steffen M. Kuhnel, and Peter Schmidt. 1987. "Is There a "Socratic Effect" in Nonexperimental Panel Studies? Consistency of an Attitude toward Guestworkers." *Sociological Methods and Research* 15:259-302.

Kraut, Robert E. and John B. McConahay. 1973. "How Being Interviewed Affects Voting: An Experiment." *Public Opinion Quarterly* 37:398-406.

Mann, Christopher B. 2005. "Unintentional Voter Mobilization: Does Participation in Pre-election Surveys Increase Voter Turnout?" *Annals of the American Academy of Political and Social Science 601*:155-168.

Menard, Scott. 1991. *Longitudinal Research*. Newbury Park, CA: Sage Publications

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

Silberstein, Adriana R. and Curtis A. Jacobs. 1989. "Symptoms of Repeated Interview Effects in the Consumer Expenditure Interview Survey." In *Panel Surveys*, ed. Daniel Kasprzyk, Greg Duncan, Graham Kalton, and M.P. Singh. New York: John Wiley and Sons.

Smith, Eliot R., Nyla Branscombe, and Carol Bormann. 1988. "Generality of the Effects of Practice on Social Judgement Tasks." *Journal of Personality and Social Psychology* 54:385-395.

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

Veroff, Joseph, Shirley Hatchett, and Elizabeth Douvan. 1992. "Consequences of Participating in a Longitudinal Study of Marriage." *Public Opinion Quarterly* 56:315-327.

Wilson, V.L. and R.R. Putnam (1982). "A Meta-analysis of Pretest Sensitization Effects in Experimental Design." *American Educational Research Journal* 19: 249-258.

Woltman, Henry, and John Bushery. 1977. "Update of the National Crime Survey Panel Bias Study." *Internal U.S. Census Bureau Report*.

Zagorsky, Jay and Pat Rhoton. 1999. *Attrition and the National Longitudinal Survey's Women Cohorts*. Center for Human Resource Research, Ohio State University.