Nonresponse Bias in a Mail Survey of Physicians

Emily McFarlane, Murrey G. Olmsted, Joe Murphy, & Craig A. Hill
RTI International, [1] Research Triangle Park, North Carolina

Paper presented at the 61st Annual Conference of the American Association for Public Opinion Research (AAPOR)
Montreal, Quebec
May 20, 2006

Abstract

In the absence of additional information, response rates are often used alone as a proxy measure for survey quality. However, total survey error is comprised of many factors, including both sampling and nonsampling errors; nonsampling errors may arise from responders and nonresponders alike (Biemer and Lyberg, 2003). Despite low response rates common in physician surveys, it is unclear to what extent response bias exists among studies with physicians. The effects of nonresponse bias in a brief mail survey of physicians were examined. In a national sample of board-certified physicians, a short survey was mailed asking physicians to nominate the five best hospitals in their specialty regardless or cost or location. Up to three follow-ups were mailed to nonresponders to gain participation. The final response rate was 47.3%.

As more than half of the sample were nonresponders, there was potential for bias if nonresponders differed significantly from responders. Willingness to respond, measured by overall response and timing of response, was analyzed with respect to several demographic variables including gender, region, specialty, urbanicity, and survey length. Next, respondent outcome measures were analyzed with respect to the demographic variables and timing of response. The outcome measures of interest were (1) nominating a top hospital in their specialty, (2) nominating two top hospitals in their specialty, and (3) nominating only hospitals in their region. This paper will describe the factors associated with nonresponse in physician surveys, the direction of nonresponse bias, and the relationship between response bias and response rates.

Keywords: nonresponse bias, physicians

1. Introduction

Physicians are an elite population with valuable knowledge and experience. Their opinions are commonly sought in surveys dealing with medical issues where input from the general population is not sufficient. However, researchers are becoming concerned about the low response rates typically found in physician surveys and the consequences for data quality.

An article reviewing 178 manuscripts published in medical journals in one year showed that response rates for surveys with physicians were on average more than 10% lower than surveys with non-physicians (Asch, Jedrziewski, & Christakis 1997). While the average response rate for a mailed physician questionnaire was relatively constant from 1985 – 1995 at about 52% for surveys with samples over 1,000 (Cummings, Savitz, & Konrad, 2002), new evidence suggests that response rates for mail surveys with physicians have declined over the past decade (Cull, Karen, O'Connor, Sharp, and Tang, 2005).

In the absence of additional information, response rates are often used alone as a proxy measure for survey quality. However, total survey error is comprised of many factors, including both sampling and nonsampling errors. Nonsampling errors may arise from responders and nonresponders alike (Biemer and Lyberg, 2003). Despite low response rates in physician surveys, it is unclear to what extent response bias exists among studies with physicians. Low response rates are not necessarily an indicator of response bias.

A recent review of the literature estimated that only 18 percent of surveys with physicians analyzed differences between responders and nonresponders (Cummings, Savitz, and Konrad, 2002). However, the results from the several published studies on response bias show that it may be less of concern for physician surveys compared to surveys with the general population (Kellerman and Herold, 2001). Studies of response bias for physicians and other health care professionals generally found no or only minimal amounts of response bias (Barton et al. 1980, McCarthy, Loval, and MacDonald 1997; Thomsen 2000). An examination of response bias in 50 surveys of pediatricians showed only modest amounts of response bias regardless of the response rate (Cull, Karen, O'Connor, Sharp and Tang 2005). When bias was present, the direction showed that female pediatricians, young pediatricians, and nonspecialty members were more likely to be respondents.

In a review of physician responses to surveys from 1967 to 1999, seven surveys used late responders as a proxy for nonresponders to evaluate potential response bias effects (Kellerman and Herold, 2001). Five of the seven surveys found no differences in demographic variables such as income, area, type of practice, and gender between early responders and late responders. When differences were observed, early responders were more likely to live in suburban areas and have higher annual incomes than late responders.

The present study further investigates response bias by examining demographic differences in physicians' willingness to respond to a short mail questionnaire and how this could impact survey results. We then explored the direction of nonresponse bias for these characteristics and its consequences for the survey estimates. Finally, we examined the relationship between response bias and response rates.

1. *RTI International is a trade name of Research Triangle Institute.*

2. Methods

Since 1990, *U.S. News & World Report* has assessed the quality of hospitals in the United States annually in the form of lists collectively titled "America's Best Hospitals." Each year, the magazine identifies hospitals of exceptional quality from over 6,000 hospitals in the United States across 17 medical specialties. Hospitals are assigned a composite score and ranked based on data from multiple sources. One of the primary sources of data is a survey of board-certified physicians asking them to nominate the "best hospitals" in their medical specialty.

2.1 Specialty

The sample for the 2005 physician survey consisted of 3,400 board-certified physicians selected from the American Medical Association (AMA) Physician Masterfile. Stratifying by region of the country (Midwest, Northeast, South, West) and medical specialty, we selected a random probability sample of 200 physicians (50 from each region) from each of the 17 specialty areas. The 17 medical specialties represented in the sample included the following: Cancer; Digestive Disorders; Ear, Nose, and Throat; Geriatrics; Gynecology; Cardiology; Hormonal Disorders; Kidney Disease and Nephrology; Neurology and Neurosurgery; Ophthalmology; Orthopedics; Pediatrics; Psychiatry; Rehabilitation; Respiratory Disorders; Rheumatology; and Urology.

2.2 Materials

150 sampled physicians per specialty were mailed a one-page, single-sided questionnaire containing a single hospital nomination item, asking them to identify as many as five hospitals in their specialty that provide the best care to patients regardless of location or expense. This survey is referred to as the "short form." The specific language of the question is as follows:

> *"Please list in the spaces below, the five hospitals (and/or affiliated medical schools) in the United States that you believe provide the best care for patients with the most serious or difficult medical problems associated with <<medical specialty>>, regardless of location or expense:"*

An additional 50 physicians in each specialty were mailed a one-page, double-sided questionnaire. The second side of the questionnaire contained questions asking the physicians to rate the importance of various factors they used as a basis for the nominations they provided on the first page. Along with the first questionnaire, physicians were sent a cover letter, a business reply envelope, and a token incentive in the form of a $2 bill. This survey is referred to as the "long form." Up to three follow-ups in two-week intervals were sent to physicians who did not respond to the previous mailings.

2.3 Analysis

The five demographic variables (gender, region, urbanicity, specialty, and the length of the survey) were chosen because they were available on the sampling frame and they are relevant to the survey estimates. Three categories of urbanicity were used: Primary Metropolitan Statistical Area (PMSA), Metropolitan Statistical Area (MSA), and non-metropolitan.

The survey estimates of interest for this study were the hospitals nominated by the physicians. Because hundreds of hospitals were nominated and because the nominations differed by specialty, we categorized the nominations into three dichotomous outcome variables to use for the analysis. The first outcome variable is whether the physician nominated at least one of the top five hospitals (from the previous year's rankings) in his/her specialty, *OneTopHospital*. The second measure was whether the physician nominated at least two of the top five hospitals in his/her specialty, *TwoTopHospitals*. The third measure was whether the physician nominated only hospitals in his/her region, *HospInRegion*.

We employed logistic regression to assess the effect of each physician characteristic, controlling for the effects of the other variables, on both response status and timing of response. Relative bias was calculated for characteristics that differed by response status. If the estimates were based only on the respondents, the bias tells you by what percentage the characteristic will be overestimated. Relative Bias is a measure of magnitude of the bias and is defined as:

$$\text{Rel } B(\bar{y}_r) = B(\bar{y}_r) / \bar{y}_r$$

Where:

$B(\bar{y}_r) = \bar{y}_r - \bar{y}_t$,

$\bar{y}_r$ *is the mean for respondents (using base weights), and*

$\bar{y}_t$ *is the mean for the full sample (using base weights).*

For any variables exhibiting response bias, we then employed logistic regression to assess the effects of the characteristics on the three survey estimates of interest: *OneTopHospital, TwoTopHospitals, and HospInRegion*.

Finally in order to determine if response bias was associated with lower response rates, we treated each specialty as an individual survey. We therefore had 17 different surveys with varying response rates. We assessed bias for the physician characteristics by comparing the characteristics of the full sample to that of the respondents for each specialty. Bias was defined above as the mean value on a characteristic for the full sample subtracted from the mean value of the characteristic for the respondents.
Bias values of 0 demonstrate that there is no difference between the full sample and the respondents on the given characteristics. Positive bias values indicate that the respondents show more of that characteristic, such as more males, than the full sample. In order to show whether the relative bias was statistically significant from 0, one-sample t-tests were conducted for each of the characteristics.

Bias distributions for each characteristic were then plotted as a function of response rate. Pearson's correlations were calculated to determine whether surveys with higher response rates were associated with lower response bias for a given characteristic.

3. Results

3.1 Response Rates

Surveys were mailed to 3,400 board-certified physicians across the United States. A total of 31 physicians were deemed ineligible because they had retired or passed away. The first mailing produced responses from 16.8% of the physicians. The second mailing produced responses from an additional 9.6% of the physicians. The third mailing generated an additional 10.5% of the physicians. The final mailing generated a response from another 10.4% of the physicians. The final response rate using AAPOR standard response rate formula number 2 was 47.3%.

Table 1 compares the physician characteristics of those who responded to those who did not respond, using base weights. Although there were differences in response by specialty, it was not included in these analyses because the survey estimates are dependent on specialty (i.e. nominating a top hospital in the physician's specialty). The effects of specialty on response were examined in a separate analysis below. Logistic regression was employed to assess the effects of each of the physician characteristics on response, while controlling for the other variables. No significant differences were found in response at the 0.05 level based on region (F = 1.70, p = .167) or urbanicity (F = 2.97, p = .052). Post-hoc analyses for urbanicity showed that the means for the three categories were not significantly different from one another. A significant difference did occur between males and females. Males were significantly more likely to respond than females ($F = 8.01$, $p < .005$). An estimate of the relative bias for males due to nonresponse was 3%. There was also a significant difference in survey length. Physicians receiving the short form were more likely to respond compared to physicians receiving the long form ($F = 15.35$, $p < .001$). An estimate of the relative bias for physicians receiving the short form due to nonresponse was 4%.

To compare early versus later responders, ordinal logistic regression was used to assess the effects of the physician characteristics on timing of response (mailing 1, mailing 2, mailing 3, or mailing 4). No significant differences were present in timing of response based on region ($F = 0.87$, $p = .457$), urbanicity ($F = 1.77$, $p = .171$), or survey length ($F = 0.60$, $p = .441$). One significant difference did occur for sex, with males being more likely to be early responders compared to females ($F = 10.44$, $p = .001$).

3.2 Consequences on Survey Estimates

Next we wanted to assess the consequences of response bias in the demographic variables on the survey estimates. If males and females nominate hospitals in the same way, then the fact that males are more likely to respond than females will not affect the quality of the survey estimates. However if males do nominate hospitals differently than females, we want to assess the impact of this difference. We used the variables exhibiting response bias for the analyses. To assess the consequences of response bias on the survey estimates, logistic regression was performed with respect to sex and survey length. Table 2 shows the percentage of physicians indicating one of the three outcome variables by sex and survey length.

Table 1: Physician Demographic Characteristics by Response Status

VARIABLE	Responders n_r=1592	Nonresponders n_m=1777
Region		
Midwest	19.5%	21.8%
Northeast	27.2%	25.4%
South	33.9%	32.1%
West	19.4%	20.7%
Sex*		
Male	80.0%	75.3%
Female	20.0%	24.7%
Urbanicity		
PMSA	7.9%	7.2%
MSA	42.6%	46.4%
Non-metropolitan	49.5%	46.4%
Survey Length**		
Short	78.2%	72.0%
Long	11.8%	18.0%

*p<.05, **p<.0001

Men were significantly more likely to nominate one or two top hospitals in their specialty ($F = 17.21$, $p = <.001$; $F = 23.62$, $p = <.001$). In addition, women were significantly more likely to nominate hospitals only in their region ($F = 26.17$, $p = <.001$). Survey length had no significant correlation with the percentage of physicians nominating at least one top hospital in their specialty or nominating hospitals only in their region. However, physicians receiving the short form were slightly more likely to nominate at least 2 top hospitals in their specialty compared to physicians receiving the long form; this difference was significant at the 0.10 level ($F = 2.71$, $p = .10$).

Table 2: Physician Demographics by Survey Estimates

	% nominating *OneTopHospital*	% nominating *TwoTopHospitals*	% nominating *HospInRegion*
Sex			
Male	83.2%	59.4%	13.2%
Female	73.3%	44.4%	24.9%
Survey Length			
Short	81.1%	57.4%	15.9%
Long	81.6%	52.7%	14.6%

3.3 Response Bias by Response Rate

Next, we wanted to assess whether response bias was associated with lower response rates. To do this, we treated the 17 specialties as 17 different surveys. Response rates varied from 35.5% in Respiratory Disorders to 61.2% in Ear, Nose and Throat. Figure 1 shows the response rates for each of the 17 specialties. There was a significant difference in response between medical specialties (F = 3.96, p < .001). Sex and Survey Length, which were both found to affect response rates, were examined individually to test for response bias across the 17 surveys.

For 11 out of 17 specialties, the bias for males was a positive value indicating that male physicians were more likely to respond than female physicians (see Figure 2). The mean response bias across the 17 surveys was 2.3 percentage points more males. This difference was statistically significant (t=2.904, p=.01). As response rates increased, the amount of positive bias decreased as can be seen by the trend line that was fit to the data. However, this correlation was not statistically significant (R = -.3656, p=.15).

For 12 out of 17 specialties, short form bias was a positive value indicating that physicians receiving the short form were more likely to respond. The mean response bias for long form across the 17 specialties was 3.4 percentage points more for physicians receiving the short form (see figure 3). This difference was statistically significant (t=-2.951, p=.01). Larger response rates were not associated with reducing the negative bias (R = -.2769, p=.28).

Figure 1: Response Rate by Specialty

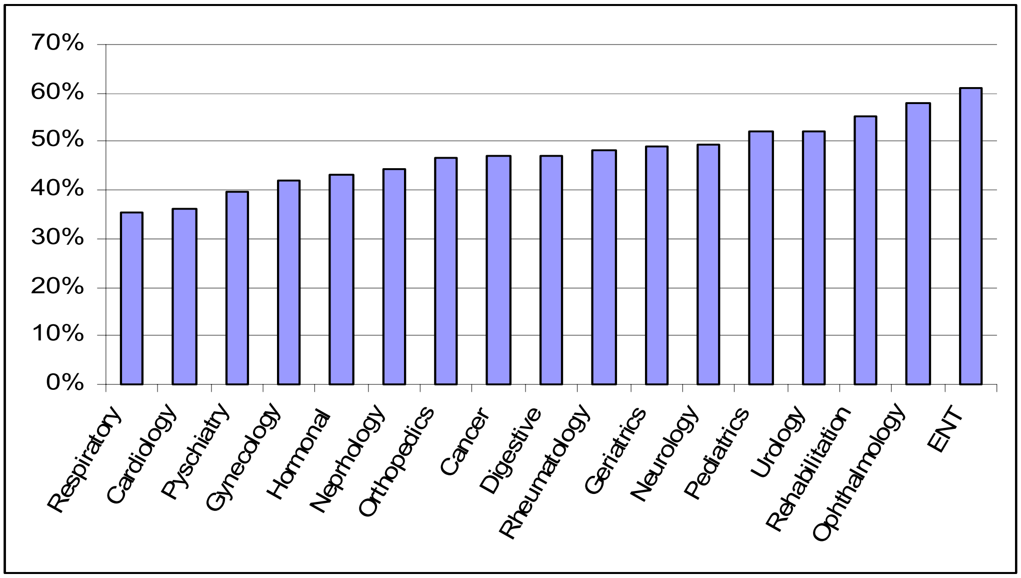


Figure 2: Gender Response Bias for Males by Response Rate

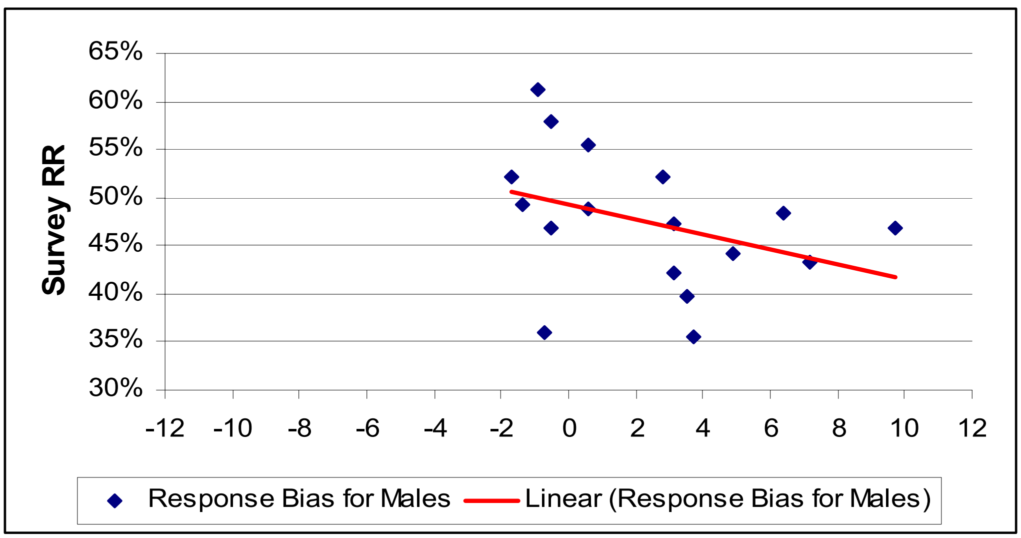
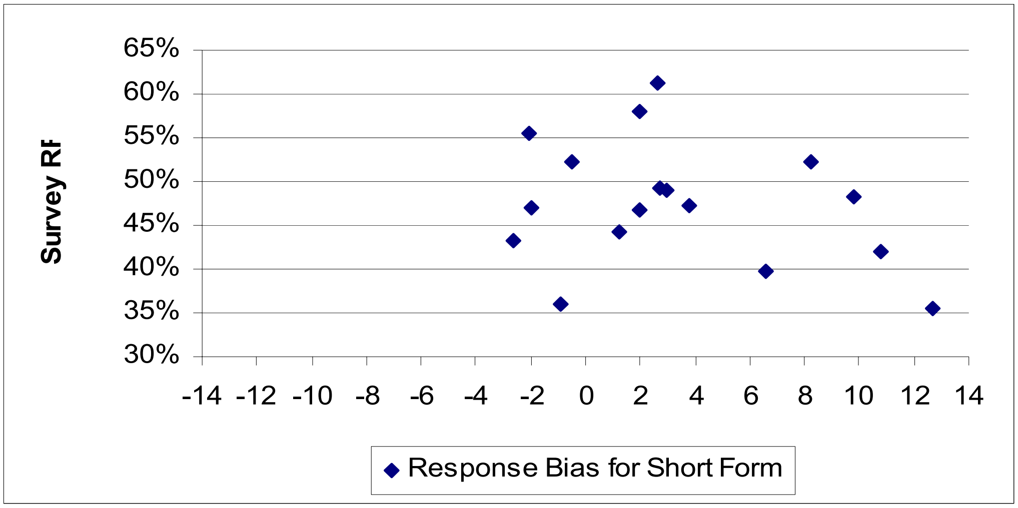


Figure 3: Survey Length Response Bias for Short Form by Response Rate



4. Summary and Discussion

In a mail survey of physicians, modest amounts of response bias were found for gender and length of survey, but not for urbanicity or region. Specifically, male physicians were significantly more likely to respond compared to female physicians. Male physicians were also significantly more likely to be early responders. The follow-up mailings helped to reduce, but not eliminate gender response bias. Male physicians were also more likely to nominate at least one or two of the five top hospitals from the previous year's rankings. In addition, women were more likely to nominate hospitals only in their region. As a result the physician nominations may be potentially biased towards the top hospitals from previous years and against hospitals in physicians' region due to gender nonresponse bias. The fact that males were more likely to respond to the survey than females is interesting and contrary to what is found in most surveys with the general population and with physicians. One explanation for this unusual finding is that rankings, in general, are more of a male construct than a female construct. Therefore the topic of the survey was considerably more appealing to males than females, influencing their decision to participate.

Survey length also had a significant impact on response rates. Physicians who received the short questionnaire (one page, one-sided survey) were significantly more likely to respond than physicians receiving the slightly longer questionnaire (one page, two-sided survey). The additional questions included on the "long" questionnaire asked physicians what they used as a basis for the nominations they provided on the first page. When asked to rate the factors that influenced their response, these factors became more salient to the physicians. Making physicians cognitively aware of what sources they were using to base their nominations may have marginally affected how they nominated hospitals. Physicians receiving the long form were slightly less likely to nominate at least two hospitals that had ranked in the top 5 in the previous year's ranking. There was no difference between early and late responders with regard to survey length.

Correlations were conducted to determine whether higher response rates were associated with lower response bias. As response rates increased, the amount of gender bias for males decreased, but this correlation was not statistically significant. There was no significant association between response bias for survey length and response rates. While response rates by specialty ranged between 36% and 61%, modest amounts of response bias were found for gender and survey length across the surveys. This indicates that response bias is a potential problem in physician surveys regardless of response rate.

Several limitations of these results should be noted. The physician characteristics used in this study were limited to those available on the sampling frame. There might be additional factors associated with response that we did not assess, such as age of physician and whether the physician works at a teaching hospital. Another limitation is that the correlation analysis between response rates and response bias was limited by having only 17 observations (specialties), which might not have been enough to determine small differences. In addition, as the highest response rate observed was only 61%, it is unclear how surveys with significantly higher response rates would perform. The biggest limitation is that while some of the characteristics associated with response were shown to be related to the survey estimates, we do not know how nonresponders would have answered.

In conclusion, this study evaluated response bias in a short survey of physicians. Modest amounts of response bias were found for gender and survey length. Follow-up waves of data collection helped to reduce gender response bias, but did not eliminate it. However higher response rates across different medical specialties were not always associated with lower response bias. While increasing response rates can reduce or eliminate response bias for some variables, high response rates are not a guarantee that response bias is not present.

Future research on this topic should include additional known characteristics of physicians in the sample frame to better characterize respondents vs. nonrespondents.

Additionally, more research is needed to determine the reasons for response differences between specialties.

References

Asch, David A., M. Kathryn, Jedrziewski, and Nicholas A. Christakis. 1997. "Response Rates to Mail Surveys Published in Medical Journals." *Journal of Clinical Epidemiology* 50:1129-1136.

Barton, J., C. Bain, C. H. Hennekens, B. Rosner, C. Belanger, A. Roth, and F. E. Speizer. 1980. "Characteristics of Respondents and Nonrespondents to a Mailed Questionnaire." *American Journal of Public Health* 70:823-825.

Biemer, Paul, and Lars Lyberg. 2003. *Introduction to Survey Quality*. New York: Wiley.

Brennan, Mike, and Janet Hoek. 1992. "The Behavior of Respondents, Nonrespondents, and Refusers Across Mail Surveys." *The Public Opinion Quaterly* 56:530-535.

Cull, William, Karen G. O'Connor, Sanford Sharp, and Suk-fong S. Tang. 2005. "Response Rates and Response Bias for 50 Surveys of Pediatricians." *Health Services Research* 40: 213-225.

Cummings, Simone M., Lucy A Savitz, and Thomas Konrad. 2001. "Reported Response Rates to Mailed Questionnaires." *Health Services Research* 35:1347-1355.

Gough, Harrison G., and Wallace B. Hall. 1977. "A Comparison of Physicians Who Did or Did Not Respond to a Postal Questionnaire." *Journal of Applied Psychology* 62:777-780.

Guadagnoli, Edward, and Susan Cunningham. 1989. "The Effects of Nonresponse and Late Response on a Survey of Physician Attitudes." *Evaluation and the Health Professions* 12:318-328.

Kellerman, Scott E., and Joan Herold. 2001. "Physician Response to Surveys: A Review of the Literature." *American Journal of Preventive Medicine* 20:61-67.

McCarthy, G.M., J. J. Koval, and J. K. MacDonald. 1997. "Nonresponse Bias in a Survey of Ontario Dentists' Infection Control and Attitudes concerning HIV." *Journal of Public Health Dentistry* 57:59-62.

Schoenman, Julie A., Marc L. Berk, Jacob J. Feldman, and Andrew Singer. 2003. "Impact of Differential Response Rates on the Quality of Data Collected in CTS Physician Survey." *Evaluation and the Health Professions*, 26:23-42.

Sobal, Jeffery, and Kevin Scott Ferentz. 1989. "Comparing Physicians' Responses to the First and Second Mailings of a Questionnaire." *Evaluation and the Health Professions* 12:329-339.

Solberg, Leif I., Mary Beth Plane, Roger L. Brown, Gail Underbakke, and Patrick E. McBride. 2001. "Nonresponse Bias: Does it Affect Measurement of Clinician Behavior?" *Medical Care* 40:347-352.

Thomsen, S. 2000. "An Examination of Nonresponse in a Work Environment Questionnaire Mailed to Psychiatric Health Care Personnel." *Journal of Occupational Health Psychology* 5:204-210.