The Effects of Respondent Commitment and Feedback on Response Quality in Online Surveys

by

Kristen L. Cibelli

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Survey Methodology)
in the University of Michigan
2017

Doctoral Committee:

Professor Frederick G. Conrad, Chair
Research Scientist Steven G. Heeringa
Senior Research Fellow Floyd J. Fowler, Jr., University of Massachusetts, Boston
Adjunct Research Assistant Professor Ting Yan, Westat

Kristen L. Cibelli

kcibelli@umich.edu

ORCID iD: 0000-0002-0374-6209

Acknowledgments

It was having the opportunity to thank the many people who helped make this research possible that kept me going when times were tough. Special thanks go to my excellent and steadfast advisor, Dr. Fred Conrad. I am also grateful to the other members of my committee – Dr. Steve Heeringa, Dr. Ting Yan, and Dr. Jack Fowler. Thank you all for believing in the value of this research and my ability to successfully carry it out. This dissertation not only came to fruition but also is better because of your contributions, guidance and support.

I would like to express my gratitude toward Beth-Ellen Pennell, the Director of the International Survey Operations Division of the Survey Research Center of the Institute for Social Research (ISR), who offered me fellowship support as well as unwavering moral support and encouragement throughout the PhD process. Much appreciation also to those I have worked with at SRO for creating such a warm and enriching work environment.

Many thanks to Dr. Wayne McCollough and Dr. Roland Blackwood for helping me with introductions to key people at the Medical School who ended up being instrumental in helping me launch the study at the University of Michigan. Special thanks in particular to Dr. Beth Tarini, (formerly Assistant Professor in the Department of Pediatrics, and member of the Child Health Evaluation Research Unit (CHEAR) at the University of Michigan and current Division Director of General Pediatrics, University of Iowa), Co-Principal Investigator for the study at the University of Michigan, for her insights into the study population, thoughtful feedback and suggestions on the study design, help navigating the IRB process, and being an unexpected but very welcome source of encouragement and support throughout the process.

I am grateful to Dr. Frauke Kreuter who facilitated the embedding of the study in Germany and to Dr. Barbara Felderer who has collaborated with me on its

implementation and the data analysis. Thanks also to Dr. Brady West and Dr. Daniela Hochfellner for help understanding the administrative data and navigating the permissions process. Thanks also to Alexandra Schmucker and Malte Schierholz for preparing the data files and helping with data access for the analysis.

This research would not have been possible without the generous support of the Rackham School of Graduate Studies at the University of Michigan, the Charles Cannell Fund in Survey Methodology, the Rensis Likert Fund, the Michigan Program in Survey Methodology (MPSM), Dr. Steve Heeringa, and the University of Michigan Department of Pediatrics and in-kind support from Survey Research Operations (SRO) at ISR.

Many thanks also to the other faculty and staff at MPSM and to current and past students of the program many of whom provided camaraderie and lifelines at key moments – Thanks in particular to Chris Antoun and Mingnan Liu (members of my cohort), Chan Zhang, Sonja Ziniel, Jessica Broome, Matt Jans, Raphael Nishimura, Zeina Mneimneh, Tuba Suzer Gurtekin, Yanna Yan, Mengyao Hu, Brian Wells, Felicitas Mittereder, and Sharan Sharma.

I'm grateful to many people at the ISR who contributed to the technical work and other operational aspects of the study at the University of Michigan. Special thanks in particular to Patty Maher, Gregg Peterson, Hueichun Peng, Andrew Hupp, Robert Fenton, Ann Vernier, and Andrew Piskorowski at SRO and to Steve Pennell who offered important insights and advice into the IRB process.

I am eternally grateful to my husband, Kent Hibben, who married me along with my PhD and sacrificed a great deal to help me see it through – and to Kent and my family for always believing in me, even when I found it difficult to believe in myself – and to the wonderful friends I shared time with in Ann Arbor.

Thanks also to Diane Willimack, Brian Harris-Kojetin, Katherine Wallman, and Mario Callegaro for their mentorship and encouragement, particularly in the early stages of the process.

To all of those who have helped, in different ways and at different times along the way, including those who I may not have mentioned by name, my deepest and heartfelt thanks.

Table of Contents

List of Tables

List of Figures

Abstract

Answering questions completely, accurately and honestly is not always the top priority for survey respondents. In their pioneering work in the 1970s and 80s, Charles Cannell and colleagues at the University of Michigan demonstrated the promise of directly asking respondents to commit to providing complete and accurate answers. While promising, these studies were conducted decades ago, in interviewer administered modes, with limited data quality measures. This dissertation consists of two experimental studies investigating the effectiveness of commitment as well as automatic feedback in improving data quality in online surveys. Study 1 measures the effect of commitment – "yes" or "no" – in an online labor force survey. Study 2 measures the effect of asking respondents to commit to engaging in several response behaviors that seem likely to promote data quality, such as reading the questions carefully, and trying to be as precise as possible, in an survey of parents about their child's health and healthcare. Study 2 also examines the effect of providing feedback in response to behaviors that are associated with reduced data quality such as speeding and reporting an incomplete date. Both studies verify the accuracy of select responses using administrative records, in contrast to the indirect quality measures in previous evaluations of commitment and feedback.

In Study 1, results were stronger for those who committed versus those who were invited to commit but did not, and in Study 2 for those who committed to all of the requested response behaviors versus those who committed to engage in a few. Study 2 also found that feedback substantially reduced unwanted respondent behaviors and improved data quality by some measures over and above the effect of commitment. Overall, in both studies, commitment had a positive effect on data quality including response accuracy, much as Cannell and his colleagues would have expected, although in Study 2, some negative consequences were also evident. For example, it appeared that committed respondents were more likely to skip a question rather than answer it

inadequately. On balance, practitioners will likely find the trade-offs produced by asking for commitment to be favorable and worth considering in production surveys.

Introduction

Answering questions completely, accurately and honestly is not always the top priority for survey respondents. To the extent that the inaccuracy in survey responses is due to insufficient effort by respondents, it might help to directly ask respondents to try harder and elicit an explicit agreement from them to do so. The rationale for this technique is that agreeing or stating one's intention to behave in a certain way *commits* a person to carry out the terms of the agreement. Charles Cannell and colleagues at the University of Michigan pioneered this technique in the 1970s and 80s. Their research showed positive effects of asking survey respondents to commit to working hard to provide complete and accurate information on increasing the amount and quality of reporting (Cannell, Marquis, & Laurent, 1977; Cannell, Miller, & Oksenberg, 1981; Oksenberg, Vinokur, & Cannell, 1977b). However, these techniques have not been examined experimentally since the 1970s and 80s, nor evaluated widely outside the context of interviewer-administered interviews.

Early studies, based on classic studies and well-accepted psychological and sociological theory, discussed below, addressed the issue of informing respondents about the survey goals and expected response process. Key findings indicate that asking respondents to make a commitment and providing feedback increases respondent motivation to respond thoroughly and provide high quality answers to health-related surveys (Cannell et al., 1977, 1981; Oksenberg et al., 1977b). The quality of reporting on response tasks designed to be demanding in terms of recall, cognitive effort or self-disclosure increased (e.g. the number of reported health conditions, mentions to open-ended questions, number of symptoms reported in the pelvic region) with asking respondents to commit to providing complete and accurate information (Miller & Cannell, 1977, 1982; Oksenberg, Vinokur, & Cannell, 1977a; Oksenberg et al., 1977b;

Vinokur, Oksenberg, & Cannell, 1977) and the use of interviewer feedback tailored to the respondent's effort on the response process (e.g. "Thanks for taking your time."; "You answered that a bit quickly.").

Research on these techniques, commitment in particular, has focused on face-to-face or telephone interviews. Research to improve response quality in web surveys has explored the effect of feedback or "interventions" triggered by undesirable respondent behavior such as responding too quickly (Conrad, Couper, Tourangeau, & Zhang, in press) and skipping questions (DeRouvray and Couper 2002), as well as increasing answers to open-ended questions (Holland & Christian, 2009). For example, Conrad and colleagues (in press) examined the effect of feedback messages triggered by responding too quickly (or "speeding" based on a certain time threshold). One of Conrad et al.'s (in press) experiments included asking for respondent commitment.

The overall idea behind commitment is that the use of such an agreement or stating one's intention to behave in a certain way commits a person to carry out the terms of the agreement. Asking respondents to make an overt agreement to work hard to provide complete and accurate information is based on well-established theory examined in classic studies in sociology (Becker, 1960; Johnson, 1973) and social psychology (Lewin (1951) Bennet (1955) cited in Cannell et al., 1981). In sociology, commitment has been used to explain the way in which individuals become committed to particular courses of action for socially grounded reasons and not just personal needs or drives (Layder, 2005). Studies in social psychology suggest that it is necessary to obtain an individual's acceptance if one wants an individual to perform a certain task. Further, the studies suggest that the implicit or explicit agreement and decision to carry out a task is motivating in and of itself (Lewin (1951) Bennet (1955) cited in Cannell et al., 1981).

The concept of commitment remains prominent in social psychology. Commitment is central to Cialdini's "principle of consistency", which states that once people clearly commit to an action or position, they tend to behave in ways that are consistent with the commitment (2001, p. 76). Psychologists view the desire for consistency as a central motivator of behavior (Cialdini, 2009, Chapter 3). This research offers theoretical grounding for the effect of commitment in increasing respondent effort and the quality of the resulting data.

Existing studies show strong effects for respondent commitment in improving response quality. Oksenberg et al. (1977a) found that respondents in the commitment condition (compared to a control) showed the following increases in reporting: significantly more mentions to open-ended items, number of health conditions, amount reported for food and drink consumed, higher mean score on reported precise-to-day index for health events, checking outside sources, and number of symptoms reported in the pelvic region. Symptoms on the mental health scale increased but were not significant. Oksenberg et al. (1977b) found similar results. Compared with a control condition, commitment significantly increased the number of items reported to open-ended questions (low education only), checking outside sources, date precision of medical event dates, number of doctor visits, precision of reported food amounts and symptoms for the pelvic area. Similar results for commitment were observed in a telephone survey (Miller & Cannell, 1982). In an experimental web survey, Conrad et al. (in press) found commitment to improve response accuracy particularly among respondents with a college education or more (results for the lower education groups were not significant) and that only a very small percentage of respondents refused to make the commitment (1%). In another online survey experiment, Vannette (2016) examined the effect of several attention filters, including asking respondents to commit to providing their best answers. Commitment group respondents in this study were less likely to break-off and provided higher quality responses for some measures including longer responses to open-ended questions. Similar to Conrad et al. (in press), very few respondents refused to make the commitment.

It is important to examine the effectiveness of the commitment and feedback techniques in the current survey environment. Today's population is reluctant to respond to surveys in general and gaining and maintaining cooperation to self-administered surveys on the web is particularly challenging because there is no interviewer to maintain motivation and keep respondents on task. Further, research suggests that people do not expect to work hard when they are online (e.g. scanning instead of reading online (Nielsen & Loranger, 2006)). As mentioned above, classic studies and well-accepted psychological and sociological theory support the idea that both commitment and feedback will increase respondent motivation and performance. Contemporary theory and

existing research, discussed further in Chapter 4, support the idea that feedback tailored to the respondent's performance should lead to more conscientious responding.

The question is whether or not the effect of these treatments in today's survey environment – and applied to the web - will be sufficiently strong to yield the types of effects observed in the 1970s and 80s. The effects may be even stronger today because without something like a commitment statement or feedback interventions, it would probably not occur to many web survey respondents to take their time and try to be as accurate as possible. However, increasing reluctance to cooperate and expend the necessary effort among today's respondents may make these techniques less effective than when they were originally examined.

On the other hand, these techniques may offer a valuable opportunity to improve respondent motivation by establishing the importance of the survey and communicating its purpose and goals. The proliferation of data requests makes it difficult for respondents to distinguish between different parties asking for their time and information (Schober & Conrad, 2008). The techniques examined here offer a means of distinguishing a survey from marketing surveys or polls and establishing credibility. Commitment may effectively communicate and engage respondents in working hard to provide more accurate answers, a notion, which is unlikely to occur to many present day web respondents on their own. This may be more important in current survey practice, resulting in greater effects than when these techniques were originally examined.

Operationally, respondent commitment and feedback lend themselves easily to the online environment. Applying these techniques to a web survey seizes an important opportunity to draw on the interactive features of the web and enhance online survey practice. Increased communication and interactivity with respondents through these techniques may overcome some of the drawbacks of not having live interviewers while maintaining the privacy and convenience for respondents to choose when they complete the survey afforded by the self-administered nature of web surveys.

This dissertation consists of two experimental studies investigating the effectiveness of commitment as well as automatic feedback in promoting better data quality in online questionnaires. The first study measures the effect of commitment – "yes" or "no" – in an online labor force survey. The experiment was embedded in a survey conducted by the

Institute for Employment Research (Institut für Arbeitsmarkt und Berufsforschung (IAB)) in Germany fielded in December 2014 – January 2015. This study examines the effect of commitment on improving response accuracy (by validating responses with administrative record information) and other indicators of data quality. The design, methods, and results of this study are detailed in Chapter 1.

Cannell and his associates (Cannell, Miller, & Oksenberg, 1981; Miller & Cannell, 1982; Oksenberg, Vinokur, & Cannell, 1977b, 1977a) proposed asking respondents to commit to providing complete and accurate responses in a global, binary way – respondents could choose either to commit or not to commit. The second study measures the effect of asking respondents to commit to engaging in several specific response behaviors that seem likely promote data quality, such as reading the questions carefully, and trying to be as precise as possible, in an online survey of the parents of child patients at University of Michigan (UM) Health System. It was fielded in March – May 2016. Response accuracy to several health service utilization questions will be validated using medical record data from the University of Michigan Health Service. The second study also examines the effect of providing feedback in response to behaviors that are associated with reduced data quality such as speeding, reporting an incomplete date, and leaving open-ended response fields blank. The design, methods, and results for the main effect of commitment in this study are detailed in Chapter 2. Chapter 3 concerns the effects of *commitment level*, i.e., how many and which of the particular response behaviors respondents pledged to carry out, extending the findings reported in the previous chapter concerning the overall effects of commitment. Finally, Chapter 4 details the design, methods, and additional effects of feedback in the context of commitment.

References

Becker, H. S. (1960). Notes on the concept of commitment. *American Journal of Sociology*, 32–40.

Cannell, C. F., Marquis, K. H., & Laurent, A. (1977). *A summary of studies of interviewing methodology*. US Government Printing Office, Washington, DC 20402.

Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology*, *12*, 389–437.

Cialdini, R. B. (2001). Harnessing the science of persuasion. *Harvard Business Review*, *79*(9), 72–81.

Cialdini, R. B. (2009). Commitment and consistency. In *Influence: Science and Practice* (5th ed.). Boston, MA: Pearson.

Conrad, F. G., Couper, M. P., Tourangeau, R., & Zhang, C. (in press). Reducing speeding in web surveys by providing immediate feedback. *Cited with Permission.*

Holland, J. L., & Christian, L. M. (2009). The influence of topic interest and interactive probing on responses to open-ended questions in web surveys. *Social Science Computer Review*, *27*(2), 196–212.

Johnson, M. P. (1973). Commitment: A conceptual structure and empirical application. *The Sociological Quarterly*, *14*(3), 395–406.

Kessler, R. C., & Üstün, T. B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93–121.

Layder, D. (2005). *Understanding Social Theory*. SAGE.

Miller, P. V., & Cannell, C. F. (1977). Communicating measurement objectives in the survey interview. In *Strategies for communication research* (Vol. 6, pp. 127–151). Beverly Hills, CA: Sage Publications, Inc.

Miller, P. V., & Cannell, C. F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, *46*(2), 250–269.

Nielsen, J., & Loranger, H. (2006). *Prioritizing web usability*. Berkeley, CA: Pearson Education.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977a). Effects of commitment to being a good respondent on interview performance. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 74–108). Department of Health, Education and Welfare.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977b). The effects of instructions, commitment and feedback on reporting in personal interviews. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 133–199). Department of Health, Education and Welfare.

Schober, M. F., & Conrad, F. C. (2008). Survey interviews and new communication technologies. *Envisioning the Survey Interview of the Future*, 1–30.

Vannette, D. L. (2016). The Effects of Asking Attention Filter Questions: Evidence from a 14 Country Survey Experiment. Presented at the Annual Conference of the American Association for Public Opinion Research, Austin, TX.

Vinokur, A., Oksenberg, L., & Cannell, C. F. (1977). Effects of feedback and reinforcement on the report of health information. In *Experiments in interviewing*

techniques: Field experiments in health reporting, 1971-1977 (pp. 109–132). Department of Health, Education and Welfare.

Chapter 1: The Effect of Respondent Commitment on Response Quality in an Online Labor Force Survey (Study 1)

1.1. Introduction

Labor force economics is a high impact domain in which the quality of survey responses has been called into question – and which therefore might benefit from more committed respondents (Moore, Stinson, & Welniak, 1999). Questions about personal and household income are routinely asked on important national surveys. Yet challenges to the validity of income-related questions due to measurement error and item non-response, in particular, are well established (Moore et al.,1999; Yan, Curtin, & Jans, 2010). This chapter details the design, methods, and results of a study on the effect of asking respondents to make a commitment to providing complete, accurate and honest information in an online labor force survey. This study extends the existing research by examining the effect of asking respondents to commit to providing complete, accurate, and honest answers on improving response accuracy (by validating responses with administrative record information) and other indicators of data quality in an online survey conducted by the Institute for Employment Research (Institut für Arbeitsmarkt und Berurfsforschung (IAB)) in Germany.

1.2. Study Design and Methods

Half of the sample for this study was randomly assigned to the treatment condition and asked to commit to providing complete, accurate and honest information, as discussed in more detail below. The remaining sample individuals were assigned to a control condition. The following commitment statement appeared following the introductory screen – the German version is available in Appendix 1:

"You have been selected to represent a portion of the country's workforce. The results from the survey can influence political decisions and thus affect the lives of many people. In order for the information from this research to be the most helpful it is important that you try to be as accurate, complete, and honest as possible with your answers. To do this, it is important to think carefully about each question, search your memory, and take time in answering. Are you willing to do this?"

Respondents could choose "Yes - I agree" or "No - but I will proceed anyway"

The questionnaire for the survey was developed by researchers at IAB and included the following:

- Labor force participation questions – e.g. current employment status, gross annual income for 2013, and the number of months the respondent or respondent's household received unemployment benefits (out of the last 12 months)
- Workplace experience and professional motivation items – 22 questions using a 5-point scale from strongly agree to strongly disagree (e.g. "I have the opportunity to do what I do best in my work")
- Internet and social media use
- Feelings in everyday life – 8 questions using a 9-point scale from strongly agree to strongly disagree (e.g. "I often have fears and anxieties about my future")
- Volunteering – e.g. active volunteered in the last 12 months (Yes/No), average time volunteered in last 12 months, reason for volunteering.
- Attitudes toward data protection – e.g. how important, likelihood of sharing information with the Federal Statistics Office, etc.
- Demographics – age, education, marital status, number of people in household
- Consent to link survey responses with official record data (Yes/No)
- Consent to contact again for a follow-up survey

Also included were several debriefing items including whether the respondent checked appropriate documents and under what circumstances, how burdensome they found the survey, and any comments or suggestions for the survey.

Respondents for the survey were part of a previous IAB experiment examining the effect of different types of communication (via letters) regarding the sharing of their contact information between agencies in Germany (for more details, see Sakshaug et al., 2016). In the earlier experiment, respondents were randomly assigned across three treatment groups – 1) Received a letter where they had to opt-in to have their information shared (opt-in group); 2) Received a letter where they had to opt-out (opt-out group) of having their information shared; 3) Control – received no letter (control group). People who either opted-in in the opt-in group to have their information shared or did not opt-out of having their information shared in the opt-out group or were part of the control group were then contacted to respond to a telephone survey.

The current survey was conducted by the IAB itself, and did not involve sharing information with any other agency. The sample could therefore be selected from among those who were not part of the telephone survey: those who did not opt-in (in the opt-in group) or opted-out (in the opt-out group) of having their information shared. Additional sample was selected from another list of mostly public sector employees. 3,812 were selected from group 1, 1621 from group 2, 4,952 from group 3, and 9,986 from the additional list for a total of 20,731. These respondents were contacted and invited to participate in the web survey by postal mail in December 2014.

One reminder letter was sent approximately one week after the initial letter of invitation. The field period concluded on February 13, 2015. A total of 2,958 respondents completed the survey with an estimated response rate of 14.5%. Respondents were randomly assigned either to the commitment condition (n=1,499) or the control condition (n=1,459).

The data analyzed in the current study are used with the permission of the Research Data Centre (FDZ) of the Federal Employment Agency at the Institute for Employment Research in Germany. The administrative data (Integrierte erwerbsbiographie (IEB)) used for validation includes employer reporting about their employees including wages to the German social security system (vom Berge, Burghardt, & Trenkle, 2013). The

employment data includes the entire German workforce except for the self-employed, some civil servants (in German, "*beamte*"), and the unemployed. The employment and income data come from mandatory reports made by the employer to the social security agencies, which are used to calculate pension claims in the statutory public pension system. Every year, employers report on all employees in the past year. These reports include begin and end date of the employment, job characteristics, and income. The data may include multiple observations per individual with every employment episode constituting one spell. Income and job characteristics are provided on a spell basis as total income over the employment episode (which last at most 365 days from one notification to the next). Integrierte Erwerbsbiographien (IEB) data has been found to be a reliable source of employment status, wages, and transfer payments (Jacobebbinghaus & Seth, 2007).

As all sample cases for the study were selected from the administrative records (Sakshaug, Schmucker, Kreuter, Couper, & Singer, 2016), survey responses could easily be linked back to information in the administrative records for those who indicated their consent to the data linkage in a question asked in the survey. Thus, the accuracy measure reported below – the comparison of survey response and administrative values – can only be derived for respondents consenting to data linkage.

The consent rate for data linkage was 75 %, which is within the normal range for IAB surveys (see for example Sakshaug and Kreuter (2014); Sakshaug, Tutz and Kreuter (2013); Sakshaug, Wolter and Kreuter (2015)). We do not expect non-consent to lead to serious bias in the analysis. In their analyses of consent bias in the personal IAB PASS interview, Sakshaug and Kreuter (2012) found consent bias only for age and foreign citizenship, which is very small compared to other sources of bias. Beste (2011) found no serious non-consent bias and only finds respondents having a foreign citizenship and respondents receiving no income at all to be underrepresented in the linked survey administrative data set.

It is not possible to obtain frequencies by demographic variables such as age, sex, education, and income on the composition of the sample due to restrictions for data protection purposes. However, regression analysis indicates that there are no significant

differences between the control group and the treatment group by respondent age, sex, and education.

1.3. Outcome Measures and Hypotheses

The proposed study examines the effect of the proposed treatment on two types of outcome measures: (1) accuracy measure and (2) indirect measures of data quality.

Accuracy measure: Accuracy measures will be derived for reported annual gross income by comparing respondents' answers to the administrative records.

Indirect measures: The study also examines the effect of respondent commitment on several indirect data quality indicators including straightlining, acquiescence, socially desirable reporting, item nonresponse, respondents reporting that they checked outside records while answering questions, and consent to link their records and responses and to be contacted for a follow-up survey. The rationale behind these data quality measures is discussed further below.

Straightlining or item nondifferentiation occurs when respondents fail to differentiate between the items in a battery of questions by giving identical or nearly identical responses to all items (Krosnick, 1991; Yan, 2008). Straightlining is considered a form of satisficing behavior. Satisficing refers to a set of response strategies employed by respondents when fully answering a survey question would require substantial cognitive effort (Krosnick, 1991). Straightlining is considered a strong form of respondent satisficing because it is presumed that respondents expend very little cognitive effort when responding (Krosnick, 1991). They do not retrieve any information from memory or integrate information to make a judgment or estimation to inform their response. Instead, the respondent superficially chooses a place on the response scale and provides the same answer throughout the battery of questions (Yan, 2008). A lack of motivation is one reason that respondents may satisfice and straightline. Further, Krosnick's (1991) theory of satisficing also incorporates respondent ability (education and cognitive ability) and task difficulty as factors that, in addition to motivation, may lead respondents to satisfice. To the extent that a lack of motivation contributes to

straightlining, the hypothesis is that respondents in the commitment group will exhibit less of this behavior.

Acquiescence is the tendency of respondents to endorse "…an assertion made in a question, regardless of the assertion's content" Krosnick and Presser (2008, p. 275). Several theories have been posited as possible explanations for acquiescence. One theory is that it is personality trait or tendency for some individuals to be "agreeable" and polite in interpersonal interactions (Leech, 1983). A second theory proposes that acquiescence results from the desire for those of lower social standing to defer to people of higher social standing (Lenski and Leggat, 1960). The respondent may view the interviewer or researcher as having higher social standing. These explanations may be more likely to apply in face-to-face or telephone interviews and may be less likely to apply in web survey interview. A third theory claims that acquiescence results from the inclination of some respondents to "satisfice" when responding to a survey questionnaire (Yan, 2008). Acquiescence can be viewed as weak satisficing, when respondents only think about reasons why a statement might be true causing them to agree more often than disagree (Krosnick, 1991). Regardless of the theory, acquiescence involves answers that do not reflect the respondent's honest views. The expectation is that commitment encourages respondents to provide honest answers thereby reducing acquiescence. Commitment may also help to reduce acquiescence that is due to lack of respondent motivation. It is important to note that, in some cases, straightlining or agreement to all items in a battery of questions may reflect the respondent's true beliefs.

Disclosure bias is a key concern when collecting information of a sensitive nature. Several questions related to volunteering will be examined for social desirability on the assumption that reporting no volunteering or infrequent volunteering is socially undesirable. Respondents were asked whether they have volunteered in the last 12 months (Yes/No). Respondents who reported that they volunteered were asked how often they volunteered on average in the last 12 months (Several times a week, once per week, one or more times per month, or rarely), and the reasons for volunteering. For the reasons for volunteering, respondents were asked to rate the importance of the following reasons from 0 to 10 (0=not at all important, 10=very important): to improve something or help people; to meet new people; to improve own skills; to advance my career, and; it was

important for my resume.

The number of socially undesirable reports to these questions indicates respondents' willingness to disclose information that may cast them in a negative light (e.g. not volunteering, rarely volunteering and self-interested reasons for volunteering). Theory and past studies suggest that respondent commitment can encourage honest reporting (Oksenberg, Vinokur, & Cannell, 1977a). The hypothesis, therefore, is that commitment will decrease socially desirable reporting and increase socially undesirable reporting.

Item nonresponse is a widely used indicator of survey data quality. While respondents may not elect to provide answers to questions for a number of reasons, failing to provide an answer is often viewed as a form of satisficing. Respondents may not choose to spend the effort necessary to remember or report information even though they technically could do so (Krosnick, 1991). If commitment effectively engages respondents in providing complete and accurate information, we would expect to see less item nonresponse among respondent in the treatment group.

Respondents reporting that they checked outside records while answering questions and consenting to link their records and responses and to be contacted for a follow-up survey indicate respondent motivation and engagement in the survey. As noted above, commitment is expected to improve respondent motivation by establishing the credibility of the survey, communicating its purpose and goals, and committing respondents to doing their part to fulfill the goals. This should make it more likely that respondents who have committed will respond affirmatively to these measures.

Of particular concern with web surveys is the tendency for respondents to break-off, that is, starting but not completing the survey. If commitment successfully engages respondents in adhering to the terms of the commitment statement, it follows that they would be less likely to break-off.

To summarize, the hypotheses for this study are as follows:

Hypothesis 1: Commitment will decrease the amount of item nonresponse compared to the control condition.

Hypothesis 2: Commitment will increase the accuracy of reported income, based on values in the administrative records, compared to the control condition.

Hypothesis 3: Commitment will decrease the amount of straightlining and acquiescence compared to the control condition

Hypothesis 4: Commitment will decrease the number of socially desirable responses and increase the number of socially undesirable responses compared to the control condition.

Hypothesis 5: Commitment will increase respondent engagement in the survey process – increasing the number of respondents reporting that they checked outside records, granting consent to link their records, and willing to be contacted for a follow-up survey.

Hypothesis 6: Commitment will decrease the number of break-offs.

1.4. Analytical Methods

To examine the effect of commitment on the outcome measures discussed in this chapter, regression models are constructed with the outcome measure as the dependent variable and commitment as an independent variable, as in the example for linear regression shown in Equation 1.1. Covariates and interaction terms were entered subsequently into the models.

Equation 1.1. Regression Equation for Estimating the Effect of Commitment

$$\gamma = \beta_0 + \beta_1 + \varepsilon$$

γ = the dependent/outcome variable
β_0 = intercept
β_1 = Commitment

A sizeable number of respondents assigned to the treatment group did not agree to the request for commitment, as discussed further below. To examine potential differences in the response behaviors of those who refused to commit and those who did commit, regression models are constructed with the outcome measure as the dependent variable and binary indicators for "Committed" and "Not committed" are entered as independent

variables, as in the example for linear regression shown in Equation 1.2. The reference category for both "Committed" and "Not committed" is the control group.

Equation 1.2. Regression Equation for Estimating the Effect of Committed and Not Committed

$$\gamma = \beta_0 + \beta_1 + \beta_2 + \varepsilon$$

γ = the dependent variable/outcome variable
β_0 = intercept
β_1 = Committed
β_2 = Not Committed

Income from the administrative records is entered as a control variable for the models assessing the accuracy of reported income and reported income item nonresponse to control for the effect of income level on reporting accuracy and willingness to report income.

Respondent education level (high versus low) is entered subsequently into all models and examined for possible interactions with commitment. There is reason to believe that the effect of commitment may vary by respondent education level. Commitment could be more effective for low education respondents if it increases motivation or encourages them to apply more effort than they would have otherwise. Higher education respondents may have a higher level of motivation, regardless of the treatment, due to increased familiarity with thinking about questions, or an increased "need for cognition", which Krosnick (1991) has identified as a possible determinant of respondent motivation, that may have prompted them to seek or succeed in higher education. A finding from Oksenberg et al. (1977) (in the analysis presented in Cannell et al. (1981)) that commitment increased reporting to open-ended questions among low education respondents only supports this idea. On the other hand, one can also see how commitment might be more effective among higher education respondents. For example, higher education respondents may have more cognitive resources to draw on than low education respondents for certain response tasks. Conrad et al.'s (in press) finding that higher education respondents answered numeracy questions more accurately when they had made a commitment (and were prompted for answering too quickly), whereas lower

education respondents did not improve with commitment, supports this theory. Other studies found no differential effects of commitment by education level (Miller & Cannell, 1982; Oksenberg et al., 1977a; Vinokur et al., 1977). These results are only reported if the interaction is significant.

High versus low education is defined based on the educational system in Germany: Secondary education in Germany can be obtained in one of three school tracks (see for example Schneider, 2008): after primary education students are either directed to the Hauptschule, Realschule or Gymnasium. People graduate from Hauptschule after approximately 9 years of total schooling, from Realschule after 10 years and from Gymnasium after 12 to 13 years depending on the German Federal State. Students graduating from Gymnasium are eligible to proceed with tertiary education (University). We code education as being high for respondents who finished Gymnasium.

For the accuracy of reported income, several metrics are used to assess the accuracy and direction of error by comparing the reported value to the value in administrative records as follows:

- The **absolute difference** between reported income and income in the records
- A **ratio metric** of reported income relative to the income reported in the reference group while controlling for income in the records
- A **tri-variate measure** – 0 for a match (within a certain amount of error) between the reported income and income in the records, 1 if the difference between reported income and the records is greater than zero (i.e. overreporting), and 2 if the difference is less than zero (i.e. underreporting)

Each of these measures has strengths and limitations. The absolute value of the difference captures the magnitude of error but lacks the directionality of the error and conflates under and overreporting – treating under-reported income the same way as an over-report. A ratio measure of the reported income relative to the income reported in the reference group quantifies the extent of and direction of reporting on average, controlling for the actual income in the records. The tri-variate measure using multinomial regression makes it possible to examine the components of error – overreporting and underreporting. Results from a combination of these measures are discussed below.

The ratio metric may detect broad differences in response behavior that may be associated with more or less accuracy, but it needs to be interpreted in the context of the other measures. For example, a significant negative difference in the ratio metric could indicate 1) underreporting compared to the reference group or, 2) more accurate reporting of income and that the reference group is overreporting its true income, or 3) more actual income in the group of interest compared to the reference group. The third explanation is unlikely due to random assignment and can be ruled out if there is no significant difference in income in the records between the groups. Evidence of underreporting in the tri-variate metric and greater absolute error (i.e., less accuracy) would support the first explanation while evidence of less overreporting and less absolute error (i.e., more accuracy) would support the second explanation.

All analyses were conducted using Stata (*Stata Statistical Software*, 2015).

1.5. Results

Respondents receiving the commitment treatment were presented with a commitment statement following the introductory screen. Respondents could either select "Yes – I agree" or "No – but I will proceed anyway". Since respondents were not required to select either of these options to proceed, they could also click "Next" and proceed without answering. Looking at how people responded to the request for commitment shown in Table 1.1, we see that 95% of respondents clicked "Yes – I agree" and can be considered "Committed". Thirty-seven clicked "No - but I will proceed anyway" and 35 did not answer.

Table 1.1. Response to the Request for Commitment

Commitment	Response to the Request for Commitment	n	%
"Committed"	Yes – I agree	1427	95.2%
"Not committed"	No – but I will proceed anyway	37	2.5%
	No answer	35	2.3%
Total		1,499	100%

While treatment group respondents agreed overwhelmingly to the commitment statement, 72 respondents or 4.8% were "not committed" and could be considered "not committed". Though in the treatment condition, these respondents, in a sense, refused the treatment. It is reasonable to assume that these respondents may have behaved differently

than their "committed" counterparts. Therefore, the remaining results section is organized into two subsections. The first subsection, 1.5.1, examines the overall effect of the commitment treatment, regardless of whether respondents agreed or did not agree, on the outcome measures and hypotheses outlined above – an "intent to treat analysis". The second section, 1.5.2, examines differences between those who agreed to the commitment ("Committed") and those who refused to commit ("Not committed") on the same outcome measures.

1.5.1. The Overall Effect of Commitment

1.5.1.1. Item Nonresponse

This section examines the effect of commitment on item nonresponse overall and specifically to the question about reported income, which could be considered particularly burdensome compared to the other questions in the survey. The outcome variable for overall item nonresponse is the proportion of questions answered out of the number of questions asked, thereby taking into account skip patterns as well as the additional question associated with the commitment treatment – i.e. the request for commitment itself. The outcome variable for reported income is a simple binary measure – i.e. whether or not the respondent provided a response to the question asking for total gross income for the previous year.

The regression results shown in Table 1.2 show a significant negative effect of commitment on the proportion of item nonresponse overall and for item nonresponse for reported income. As is evident from the small regression coefficients, there was not much item nonresponse to the survey, in general. This is probably because, apart from a few questions about income, the items on the survey were not particularly sensitive or unduly burdensome. The R-squared values for these models are also low, suggesting that commitment may not explain much of the variation in item nonresponse. Nonetheless, we do see a significant effect for commitment, in support of Hypothesis 1, this suggests that commitment group respondents skipped a smaller proportion of questions overall, and were less likely to skip the income question.

Table 1.2. The Effect of Commitment on Item Nonresponse

VARIABLES	Overall Item Nonresponse	Income Item Nonresponse
Commitment	-0.008**	-0.243**
	(0.003)	(0.082)
Constant	0.115***	-0.210
	(0.002)	(0.057)
Observations	2,927	2,449
R-squared	0.002	

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.1.2. Accuracy of Reported Income

Commitment group respondents were more likely to respond to the income question. We now explore the extent to which these respondents were more or less accurate in their responses to the reported income question than the control group. It is possible to evaluate the accuracy of reported income for those who reported their income and consented to having their survey responses linked to administrative records; some respondents, including a particular type of seasonal/manual workers and the self-employed were not asked to report their income because it is not available in the administrative records to verify, and some respondents elected not to answer the question (item nonresponse). Administrative income is censored at a certain income limit (Beitragsbemessungsgrenze) that differs for East and West Germany and by year. In the year 2013, the income limit was 69,600 Euros for West and 58,800 Euros for East Germany. Because we know that the censored income does not match the "true" income, all cases with income in the administrative records that is equal to the income limit are excluded.

Descriptive statistics, shown in Table 1.3, indicate that overall reported income is lower than the income in the administrative records. The positive values for skewness and kurtosis for reported income suggest a right-skewed distribution, whereas, the administrative income follows a nearly normal distribution. This is also reflected in histograms provided in Appendix 1. Log transformations are therefore used for reported income and administrative income in the measures reported below.

Table 1.3. Descriptive Statistics for Reported Income and Income in the Administrative Records

	Reported Income	Administrative Income
	n=1,112	n=1,112
Mean	34,994	38,097
Standard Deviation	23,613	15,297
Median	35,000	38,059
Skewness	8.39	-0.10
Kurtosis	164.58	2.46

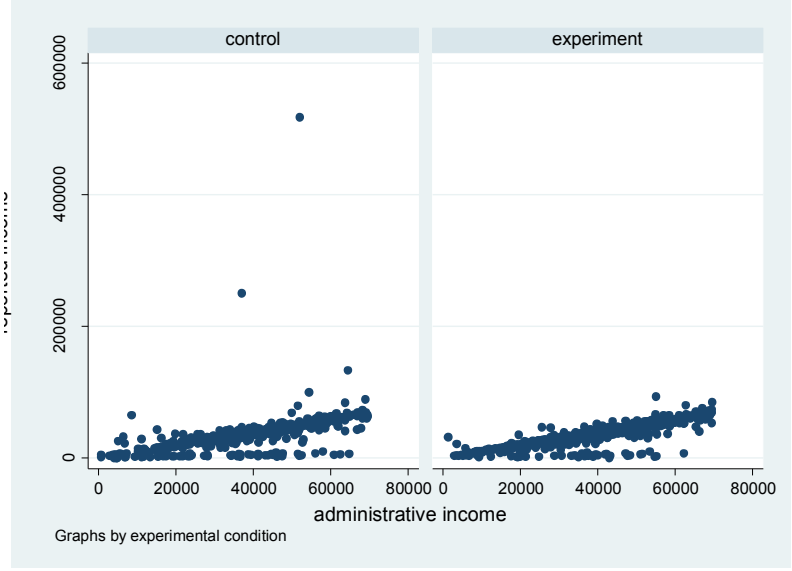
Comparing income in the administrative records for the commitment and control group, shown in Table 1.4, we see that the income reported in the records is quite similar between the treatment and control group. We also see that mean and median reported income between the commitment and control group are very similar. However, the large standard deviation and high level of skewness and kurtosis indicate a right-skewed distribution, indicating higher reported values, including some potential extreme observations, for the control group, whereas, these values for the commitment group are much lower. This is reflected in the plots for reported income by income in the records for the commitment and control groups shown in Figure 1.1, which clearly shows two extreme observations for the control group.

If commitment increased the accuracy of reported income, we would expect to see more cohesion around the line between reported income and income in the administrative records. This is evident in Figure 1.1, where there is a slightly tighter clustering between reported income and income in the records for the commitment (experiment) group compared to the control group. If commitment increased the accuracy of reported income, we would also expect to see a smaller absolute difference between median reported income and income in the records as well as a higher percentage of a match between reported income and income in the records (see the row labeled "Absolute difference (median)"). For the percentage of a match, we examine the percentages of a match within 1% of the records. As expected, in Table 1.4, we see a smaller absolute difference between median reported income and income in the administrative records for the commitment group compared to the control group. We also see slightly higher percentage of a match for the commitment group.

Table 1.4. Descriptive Statistics for Reported Income and Income in the Administrative Records for the Commitment and Control Group (Income in Euros)

	Control		Commitment	
	n=538		n=574	
	Reported Income	Income in Records	Reported Income	Income in Records
Mean	35,221	37,738	34,782	38,433
Standard Deviation	29,222	15,719	16,753	14,896
Absolute difference (median)	7,890		5,447	
	n	%	n	%
Overreporting	137	25.5	140	24.4
Matching (within 1%)	52	9.7	67	11.7
Underreporting	349	64.9	367	63.9

Figure 1.1. Total Reported Income by Income in the Administrative Records for the Commitment and Control Group



The regression results shown in Table 1.5 show that commitment group respondents had a smaller absolute difference between reported median income and income in the records compared to the control group. This result supports Hypothesis 2. While these results indicate that commitment group respondents reported their income more accurately in terms of less absolute error, there does not appear to be any significant differences in the direction of reporting based on results for the ratio metric – also shown in Table 1.5 – or the results for the tri-variate measure – (0=match, 1=overreport,

2=underreport) – using multinomial regression, shown in Table 1.6, both of which are not statistically significant.

Commitment group respondents may have been more accurate by looking up income information in records, which as reported below, they were significantly more likely to report doing than the control group. That commitment group respondents were significantly less likely to report a rounded answer than their control group counterparts ($p < 0.01$) supporting the idea that they looked up information and were thereby more likely to report a more precise, unrounded value.

Table 1.5. The Effect of Commitment on the Accuracy of Reported Income

VARIABLES	Log Absolute Difference	Total Log Reported Income (Ratio Metric)
Commitment	-0.020*	-0.006
	(0.008)	(0.007)
Log income in records	-0.033***	0.988***
	(0.007)	(0.006)
Constant	0.425***	0.085
	(0.078)	(0.063)
Observations	1,112	1,112

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Table 1.6. The Effect of Commitment on Overreporting and Underreporting

VARIABLES	Total Reported Income Model 1
logit 2: Overreporting versus zero difference	
Commitment	-0.200
	(0.200)
Log income in records	-0.120
	(0.199)
Constant	3.158
	(2.093)
logit 3: Underreporting versus zero difference	
Commitment	-0.208
	(0.221)
Log income in records	-0.521
	(0.210)
Constant	6.381**
	(2.196)
Observations	1,112

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.1.3. Straightlining and Acquiescence

Answering the same way to eight out of eight items in a grid (Yes/No) of self-perception questions using a 9-point Likert scale is defined as straightlining. For acquiescence, the dependent variable is the number of "agree" or "strongly agree" answers in the same battery of questions. A parameter estimate is not available for acquiescence due to the very low incidence of acquiescence in either the commitment or control group.

The regression results in Table 1.7 show no significant effect of commitment on straightlining, which fails to support Hypothesis 3. However, this could be because, due to the nature of the questions, it is unlikely that respondents would answer the same way to all of the questions if they were answering thoughtfully.

Table 1.7. The Effect of Commitment on Straightlining

VARIABLES	Straightlining
Commitment	-0.415
	(0.285)
Constant	3.735***
	(0.182)
Observations	2,683

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.1.4. Socially Desirable Reporting

Several items related to volunteering are used to examine the effect of commitment on socially desirable reporting. As a socially desirable activity, we would expect lower reports for reported volunteering. If commitment group respondents are more honest, we would also expect lower reports of altruistic reasons for volunteering - such as wanting to help others - and higher reports of selfish or self-centered reasons - such as wanting to advance one's career.

As shown in Table 1.8, results were in the expected direction for 5 out of 7 of these measures. For one of these measures, the result is marginally significant – with commitment group respondents marginally more likely to report that they volunteer

because it is important for their resume. These results offer partial support for Hypothesis 4.

Table 1.8. The Effect of Commitment on Socially Desirable Reporting

	Altruistic measures – expect negative			Selfish measures – expect positive			
VARIABLES	Volunteering in last year	Volunteering frequency	Help people	Meet new people	Improve skills	Advance career	Important for resume
Commitment	-0.040	-0.115	0.136	-0.068	0.143	0.157	0.311^
	(0.078)	(0.127)	(0.146)	(0.181)	(0.196)	(0.137)	(0.163)
Constant	0.338***	-0.950***	8.627***	5.123***	5.019***	2.292***	2.621***
	(0.055)	(0.088)	(0.102)	(0.127)	(0.137)	(0.096)	(0.113)
Observations	2,723	1,285	1,243	1,214	1,208	1,205	1,207

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.1.5. Checking Records and Consent

The results in Table 1.9 show that commitment group respondents were significantly more likely to report that they checked records, which suggests the request to look up information in records, as needed, in the commitment statement was effective. On the other hand, the commitment group was no more likely than control group respondents to consent to having their survey responses linked with administrative records or to be recontacted for a follow-up interview. Perhaps commitment group respondents would have been more likely to grant consent if this had been encouraged or requested as part of the commitment statement. These results offer partial support for Hypothesis 5.

Table 1.9. The Effect of Commitment on Checking Records, Consent to Records Linkage, and to Recontact

VARIABLES	Checked Records	Consent to Record Linkage	Consent to Recontact
Commitment	0.394***	-0.029	0.015
	(0.090)	(0.094)	(0.075)
Constant	-1.251***	1.292***	0.242***
	(0.067)	(0.066)	(0.053)
Observations	2,609	2,771	2,912

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.1.6. Break-offs

This section examines the effect of commitment on the number of break-offs, which is starting but not completing the survey. Contrary to expectation based on the regression results shown in Table 1.10, commitment is associated with a marginally higher chance of breaking off, which fails to support for Hypothesis 6. As a potential downside to commitment, more commitment group respondents may have opted not to complete the interview if they felt that they were unable to provide complete and accurate answers, as requested in the commitment statement.

Table 1.10. The Effect of Commitment on Breaking Off

VARIABLES	Break-off
Commitment	0.245^
	(0.132)
Constant	2.487***
	(0.098)
Observations	2,958

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.2. The Effect of Being Committed and Not Committed

The following section examines the effect of being committed and not committed on the outcome measures above with the exception of acquiescence because as with commitment overall, there were too few occurrences of acquiesce to estimate parameters for this measure, and social desirability as no significant differences were observed.

1.5.2.1. Item Nonresponse

Looking first at the results for income item nonresponse shown in Table 1.11, we see that committed respondents were significantly less likely to skip the income question while not committed respondents were more likely to skip it, at least directionally. Looking at the proportion of item nonresponse overall, we see a significant effect for those who committed in reducing the proportion of overall item nonresponse while not committed respondents skipped significantly more items.

Table 1.11. The Effect of Committed and Not Committed on Item Nonresponse

VARIABLES	Overall Item Nonresponse	Income Item Nonresponse
Committed	-0.013***	-0.263**
	(0.003)	(0.083)
Not committed	0.075***	0.254
	(0.010)	(0.304)
Constant	0.115***	-0.210
	(0.003)	(0.057)
Observations	2,927	2,449
R-squared	0.026	

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.2.2. Accuracy of Reported Income

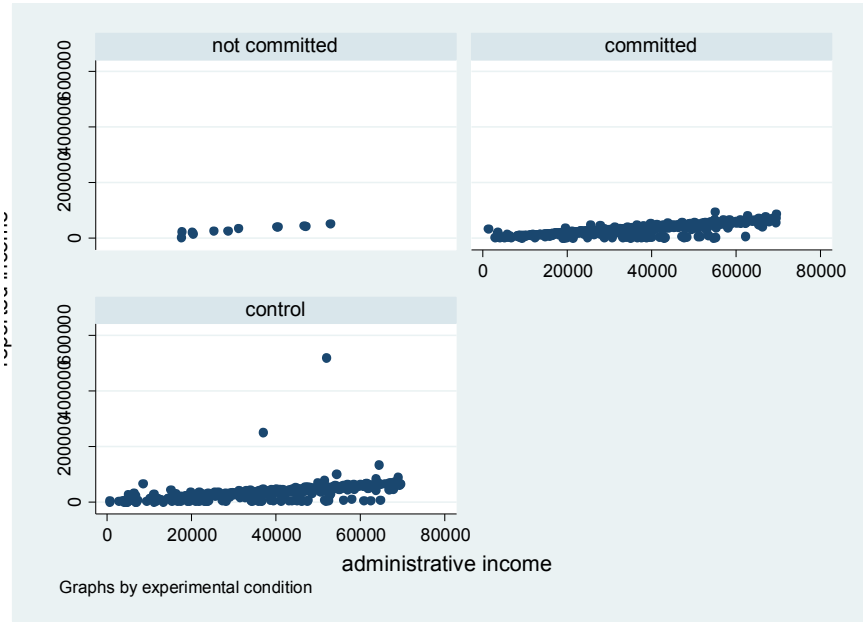
The previous subsection found that not committed respondents were less likely to respond to the income question. Accordingly, we see in the descriptive statistics presented in Table 1.13, that only 12 out of 72 not committed respondents provided an answer to this question. Also apparent in Table 1.12, is the larger absolute difference in the median reported income for the not committed compared to committed respondents indicating more error in reported income. It is still smaller than that of the control group. This is reflected in Figure 1.2, which shows reported income by income in the administrative records for the committed, not committed and the control group, and slightly greater dispersion for the not committed compared to the committed. We also see in Table 1.12 that not committed respondents reported a higher number of matches (within 1%) than the committed and control group but this based on very few reporting cases overall for the not committed.

Table 1.12. Descriptive Statistics for Reported Income and Income in the Administrative Records for Control Group, Committed and Not Committed

	Control		Committed		Not Committed	
	n=538		n=562		n=12	
	Reported Income	Income in Records	Reported Income	Income in Records	Reported Income	Income in Records
Mean	35,221	37,738	34,885	38,562	29,924	32,406
Standard Deviation	29,222	15,719	16,803	14,924	13,937	12,614
Absolute difference (median)	3,110		2,308		3,047	

	n	%	n	%	n	%
Overreporting	137	25.5	138	24.6	2	16.7
Matching (within 1%)	52	9.7	65	11.6	2	16.7
Underreporting	349	64.9	359	63.9	8	66.7

Figure 1.2. Total Reported Income by Income in the Records for Committed, Not Committed and the Control Group



The regression results for the median absolute difference between reported income and income in the records in Table 1.13 show a significantly smaller absolute difference for committed respondents. The coefficient for the not committed is also negative, suggesting that these respondents had a smaller absolute difference than the control group, but this result is not significant. The coefficient for the committed is slightly larger than for commitment overall reported above suggesting slightly improved accuracy for the committed in excluding the not committed. But overall, the results for the accuracy of reported income are essentially unchanged. The results for the ratio measure for overreporting and underreporting remained not significant from the results for commitment overall and are therefore not reported.

Committed respondents were significantly less likely than control group respondents to give a rounded answer to reported income ($p < .01$). This was not significant for not committed respondents.

Table 1.13. The Effect of Committed and Not Committed on the Accuracy of Reported Income

VARIABLES	Total Income Log Absolute Difference
Committed	-0.066^
	(0.040)
Not committed	-0.105
	(0.191)
Log income in records	-0.196***
	(0.035)
Constant	2.355***
	(0.364)
Observations	1,110

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.2.3. Straightlining

In section 1.5.1.3., we saw no effect for commitment on reducing straightlining. However, in the results shown in Table 1.14, we see that committed respondents were marginally less likely to straightline (providing the same answer to each question in the battery) while not committed respondents were marginally more likely to straightline. These results suggest that the additional straightlining among the not committed cancelled out the reduced of straightlining among the committed respondents in the analysis of commitment overall.

We are not able to report results for acquiescence because parameters could not be estimated due to the low incidence of acquiescence in either group.

Table 1.14. The Effect of Committed and Not Committed on Straightlining

VARIABLES	Straightlining
Committed	-0.534^
	(0.300)
Not committed	1.027^
	(0.623)
Constant	3.735***
	(0.182)
Observations	2,683

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.2.4. Checking Records and Consent

In section 1.5.1.5., we saw that commitment group respondents overall, were significantly more likely than the control group to report checking records but that they were no more likely to consent to having their survey responses linked with records or to be recontacted for a follow-up interview. As is evident in Table 1.15, the effect for checking records for commitment overall was driven by the committed respondents. Not committed respondents were not any more likely than control group respondents to report that they checked records. Conversely, while committed respondents were no more likely than controls to consent to having their survey responses linked or to being recontacted, not committed respondents were significantly less likely to consent to either request.

Table 1.15. The Effect of Committed and Not Committed on Checking Records, Consent to Records Linkage, and to Recontact

VARIABLES	Checked Records	Consent to Record Linkage	Consent to Recontact
Committed	0.411***	0.039	0.094
	(0.091)	(0.094)	(0.076)
Not committed	-0.084	-1.370***	-1.738***
	(0.362)	(0.285)	(0.311)
Constant	-1.251***	1.292***	0.242***
	(0.067)	(0.066)	(0.053)
Observations	2,609	2,771	2,912

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.5.2.5. Break-offs

In section 1.5.1.6. , we saw that commitment group respondents overall, were marginally more likely to break-off than the control group respondents. The results in Table 1.16 show that that effect was driven by the not committed, who were significantly more likely to break-off, while no significant effect is observed for the committed.

Table 1.16. The Effect of Committed and Not Committed on Breaking Off

VARIABLES	Break-off
Committed	0.153
	(0.136)
Not committed	1.389***
	(0.289)
Constant	2.487***

	(0.098)
Observations	2,958

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

1.6. Discussion

This study produced a number of promising effects for asking respondents to commit to providing complete, accurate, and honest answers. Results were particularly promising for those who committed versus those who were invited to commit but did not.

For the effect of commitment overall, regardless of whether respondents agreed or did not agree to the commitment, there was a significant reduction in item nonresponse overall and for reported income, specifically, which is a perpetual struggle for surveys that ask about personal or household income. Commitment group respondents were also significantly more accurate with their reported income in terms of reduced absolute error between reported income and income in the administrative records compared to the control group. It seems likely that checking records helped, which commitment group respondents were much more likely to report having done. They were also less likely to report a rounded answer than their control group counterparts, supporting the idea that committed respondents thought more carefully about their numerical answer or that they looked up information and were thereby more likely to report precise, unrounded answers. There was also some evidence of more socially undesirable reporting in the commitment group.

While the response of the treatment group to the request for commitment was mostly positive, 4.8% did not agree and could be considered “not committed”. Analyses examining the response behavior of the committed and not committed separately, compared to the control group, revealed striking differences. While committed respondents had significantly less item nonresponse, not committed respondents had significantly more compared to the control group. There was no significant effect for the not committed on the accuracy of reported income, while committed respondents were significantly more accurate in terms of the absolute difference between reported income and income in the administrative records. Not committed respondents were also marginally more likely to straightline, while committed respondents were marginally less

likely. Further, the not committed were also significantly more likely to break-off, and less likely to report checking records or to consent to having their survey responses linked with administrative records or to being recontacted for a follow-up interview. Because the control group includes many respondents who probably would have committed if asked, it is perhaps not surprising that the strongest contrasts are between the not committed and the control group compared to the committed and the control group.

In the original studies by Cannell and his associates (Miller & Cannell, 1982; Oksenberg, Vinokur, & Cannell, 1977b; Oksenberg et al., 1977a) very few sample members refused the commitment request in interviewer-administered (face-to-face and telephone) data collection. This was also the case in Conrad et al.'s (in press) and Vannette's (2016) web-based studies. The sizeable number of respondents who did not commit in the current study may be related to the web mode and the lack of social presence of an interviewer to not only increase the number of respondents who would commit but also motivate them to adhere to the terms of the commitment. Or it may have to do with attitudes toward survey participation and effort fifty or more years after the original studies.[1] Either way, the results from this study suggest that, while most respondents are likely to commit and provide better quality responses, a portion are likely not to commit and may provide inferior quality data. However, the relatively poor response behavior of the not committed – provided they are relatively few – may not diminish the overall effect of commitment. In this study, the results for the overall effect of commitment were only affected by the poorer response behavior of the not committed for two measures. First, there were marginally more break-offs for commitment overall compared to control, which could be attributed to the significantly higher proportion of break-offs for not committed. Second, there was no effect in reducing straightlining for commitment overall compared to control because gains made from the committed, who were marginally less likely to straightline, were cancelled out by marginally more straightlining among the not committed. Further, commitment could be used to identify those who are not willing to be conscientious - making it possible to filter out or treat

[1] The recent web studies involved paid volunteer samples (Conrad, Couper, Tourangeau, & Zhang, in press; Vannette, 2016) so may not be comparable to the current, probability web sample.

data from these respondents separately at the analysis stage. With very little operational cost, commitment appears to offer a simple yet effective approach to improving web survey quality.

References

Beste, J. (2011). Selektivitätsprozesse bei der Verknüpfung von Befragungs-mitProzessdaten: Record Linkage mit Daten des Panels „Arbeitsmarkt und soziale Sicherung "und administrativen Daten der Bundesagentur für Arbeit. Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg [Institute for Employment Research, Nuremberg, Germany].

Conrad, F. G., Couper, M. P., Tourangeau, R., & Zhang, C. (in press). Reducing speeding in web surveys by providing immediate feedback. *Cited with Permission.*

Jacobebbinghaus, P., & Seth, S. (2007). The German integrated employment biographies sample IEBS. *Schmollers Jahrbuch*, *127*(2), 335–342.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Miller, P. V., & Cannell, C. F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, *46*(2), 250–269.

Moore, J. C., Stinson, L. L., & Welniak, E. (1999). Income reporting in surveys: Cognitive issues and measurement error. *Cognition and Survey Research*, 155–174.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977a). Effects of commitment to being a good respondent on interview performance. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 74–108). Department of Health, Education and Welfare.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977b). The effects of instructions, commitment and feedback on reporting in personal interviews. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 133–199). Department of Health, Education and Welfare.

Sakshaug, J. W., & Kreuter, F. (2012). Assessing the magnitude of non-consent biases in linked survey and administrative data. In *Survey Research Methods* (Vol. 6, pp. 113–122). Citeseer.

Sakshaug, J. W., & Kreuter, F. (2014). The effect of benefit wording on consent to link survey and administrative records in a web survey. *Public Opinion Quarterly*, *78*(1), 166–176.

Sakshaug, J. W., Schmucker, A., Kreuter, F., Couper, M. P., & Singer, E. (2016). Evaluating active (opt-in) and passive (opt-out) consent bias in the transfer of federal contact data to a third-party survey agency. *Journal of Survey Statistics and Methodology*, smw020.

Sakshaug, J. W., Tutz, V., & Kreuter, F. (2013). Placement, wording, and interviewers: Identifying correlates of consent to link survey and administrative data. In *Survey Research Methods* (Vol. 7, pp. 133–144). Citeseer.

Sakshaug, J. W., Wolter, S., & Kreuter, F. (2015). Obtaining Record Linkage Consent: Results from a Wording Experiment in Germany. *Survey Methods: Insights from the Field (SMIF)*. Retrieved from http://surveyinsights.org/?p=7288

Stata Statistical Software. (2015). (Version Release 14). College Station, TX: StataCorp LP.

Vannette, D. L. (2016). The Effects of Asking Attention Filter Questions: Evidence from a 14 Country Survey Experiment. Presented at the Annual Conference of the American Association for Public Opinion Research, Austin, TX.

vom Berge, P., Burghardt, A., & Trenkle, S. (2013). *Sample-of-Integrated-Labour-Market-Biographies Regional-File 1975-2010 (SIAB-R 7510)* (FDZ-Datenreport No. 9). Nuremberg, Germany: Research Data Centre (FDZ).

Yan, T. (2008). Nondifferentiation. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (Vol. 2, pp. 520–521). Thousand Oaks, CA: Sage.

Yan, T., Curtin, R., & Jans, M. (2010). Trends in income nonresponse over two decades. *Journal of Official Statistics*, *26*(1), 145.

Appendix 1.1. Text of the request for commitment in German

"Sie wurden in einem Zufallsverfahren für diese Befragung ausgewählt, um einen Teil der Bevölkerung zu repräsentieren. Die Ergebnisse dieser Befragung können Einfluss auf politische Entscheidungen nehmen und damit das Leben vieler Menschen verändern. Ihre Teilnahme ist dann am wertvollsten, wenn Sie versuchen, genaue, vollständige und ehrliche Antworten zu geben. Es ist dabei wichtig, dass Sie in Ruhe über jede Frage nachdenken und sich die nötige Zeit zur Beantwortung der Frage nehmen. Manchmal fallen einem Antworten nicht sofort ein und es kann hilfreich sein, zur Beantwortung in geeigneten Unterlagen nachzusehen. Sind Sie bereit sich zu bemühen, bestmögliche Antworten zu geben?"

Ja, ich bin bereit
Nein, aber ich setzte die Befragung in jedem Fall fort

Appendix 1.2. Distribution of Total Reported Income and Income in the Administrative Records

Figure 1.3. Total Reported Income

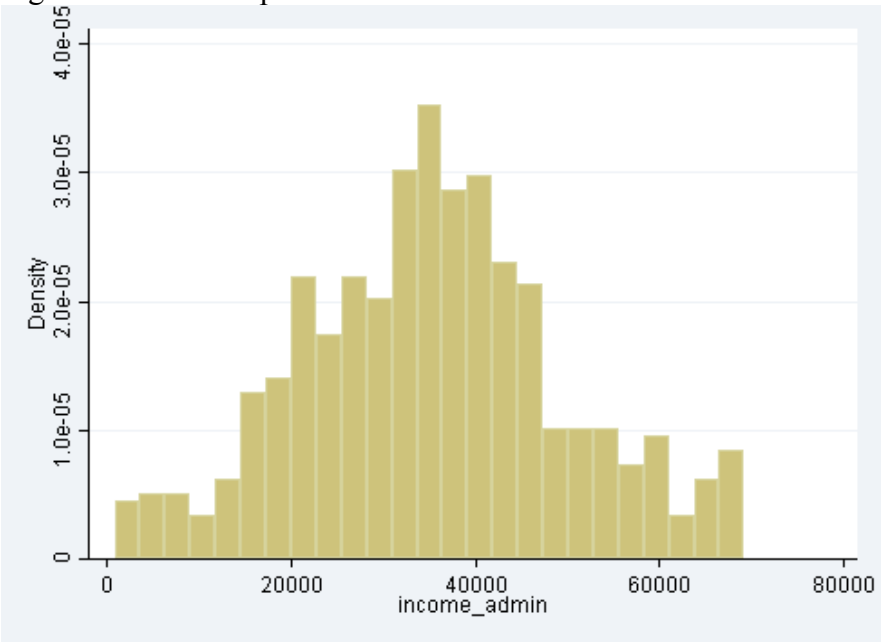
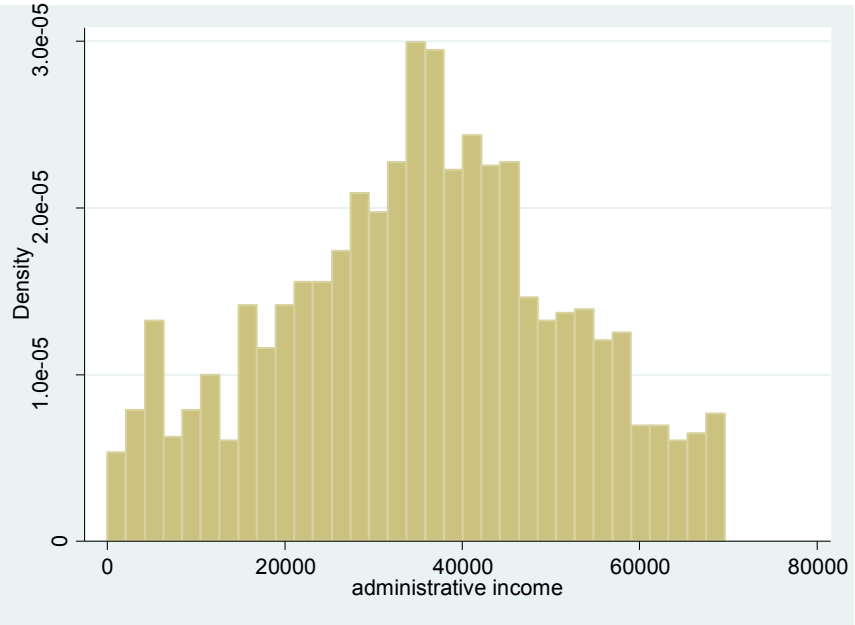


Figure 1.4. Total Income in the Administrative Records



Chapter 2: The Effect of Commitment in a Survey of Parents about their Child's Health and Healthcare (Study 2)

2.1. Introduction

This chapter details the design, methods, and results of a study on the effect of asking respondents to commit to certain desired respondent behaviors in a survey of parents about their child's health and healthcare. Building on Study 1: The Effect of Respondent Commitment on Response Quality in an Online Labor Force Survey discussed in the previous chapter, Study 2 examines two additional treatments – feedback and contextual recall cues (CRCs) – in addition to respondent commitment, on improving the accuracy of responses to health care service utilization questions. It also examines the effect of respondent commitment and feedback on several indicators of data quality. Further, in Study 2, the request for commitment is carried out differently than in Study 1 and in previous research examining this technique. The request for commitment in the current study includes a statement about the importance of complete, accurate and honest answers and then asks respondents to agree to perform each behavior in a list of five behaviors that should increase the quality of their answers including reading all of the questions carefully, trying to be as precise as possible, and looking up information in records or on a calendar if needed, by checking checkboxes for each behavior. The idea was that asking respondents to check boxes next to the requested behaviors would underscore what is meant by commitment and make it more likely that respondents would do what was being asked, thereby strengthening the treatment.

2.2. Study Design and Methods

Invitations to the survey for this study were sent via postal mail because this was the mode of contact available in the medical records, from which the study sample was drawn. Due to limited resources, uncertainty about what the response rate would be, and to ensure that there

would be a sufficient number of responding cases to detect an effect of commitment, the primary treatment of interest, the study employed a nested design, with the additional experimental factors, feedback and CRCs, nested within the principal treatment of commitment, as follows:

Group 1: Commitment

Group 2: Commitment+Feedback

Group 3: Commitment+CRCs

Group 4: Control

This chapter focuses on the effect of commitment. Feedback and CRCs, and their effect on responses, in the context of the commitment treatment, are discussed in chapter 4.

The following introductory statement appeared to all respondents:

> *"Welcome! We are conducting a study to learn more about the children we serve so that we can do a better job meeting their and your family's needs. We will be asking questions about your experiences for one of your children at our facilities, the past visits and the dates. There will also be questions about your child's health and lifestyle."*

Respondents randomly assigned to one of the treatment groups were presented with the commitment statement shown in Figure 2.1 following the introductory message shown above and two introductory survey questions. Respondents could select one or more of the checkboxes or none of the above. Respondents who selected none of the above were still able to proceed with the survey.

Figure 2.1. Commitment Statement

> You have been selected to represent families whose children receive care at the University of Michigan. In order for the information from this research to be the most helpful it is important that you try to be as accurate, complete, and honest as possible with your answers.
>
> I commit to the following [please check all that apply]:
>
> □ Reading all the questions carefully

□ Trying to be as precise as possible with my answers
□ Looking up information in records or on a calendar, if needed
□ Providing as much information as possible
□ Answering honestly

○ None of the above – but I will proceed anyway

The questionnaire for the survey included the following types of questions:

- Introductory items – how long has child been a patient at UM and general rating of child's health
- Health service utilization – child's visits to the pediatrician, specialists, the ER, and the hospital in the last 12 months as well as the dates of the most recent visits
- Open-ended health practices questions (3 questions)
- Behavioral frequency questions (which are be moderately sensitive and could be susceptible to socially desirable reporting) (9 questions)
- Likert-scale items on satisfaction with care at the University of Michigan (6 questions, one reverse coded)
- Demographics – respondent age, relationship to the child, education, marital status, employment status, household income
- Debriefing items – what respondent recalls about what they were asked to commit to (treatment group respondents only), whether the CRCs were helpful (CRC respondents only), level of confidence about the accuracy of answers to number of visit and date questions, whether checked records, level of effort put in, how burdensome was the survey, any suggestions for the survey

The full questionnaire is provided in Appendix 1.

The behavioral frequency questions (e.g., children's sleep adequacy, exercise, family meals, time spent in front of the TV, time spent in front of a computer, smoking in presence of child, allowing others to smoke in presence of child) are based on questions in the National Survey of Children's Health (2011-2012)[2]. Additional behavioral frequency questions (how often do you raise your voice or yell at child, how often do you spank your child) are based on questions from the 2000 National Survey of Early

[2] http://www.childhealthdata.org/learn/methods

Childhood Health (NSECH).[3] In the analyses reported below, answers to questions about the number of visits and the dates of recent visits are validated against medical records from the University of Michigan Honest Broker Office. Only the date of the most recent visit was asked because asking for the dates of all reported visits would unreasonably increase respondent burden and would lead to unequal numbers of observations across respondents. The remaining items in the questionnaire provide additional dependent variables through which data quality can be assessed.

Question-specific instructions outlined below in Table 2.1 were also provided as part of the commitment treatment.

Table 2.1. Question-specific Instructions Provided as Part of the Commitment Treatment

Questions	Instructions
Medical visits – e.g. During the past 12 months, how many times did [CHILD] see a primary care doctor or nurse practitioner at [his/her] University of Michigan pediatrician's office? Date of the most recent visit – e.g. What was the date of [CHILD]'s most recent visit to [his/her] University of Michigan pediatrician's office?	For the next set of questions, we'd like you to be as exact as you can about the number of visits and dates of the most recent visits. To be the most accurate, you may need to take time to think carefully before you answer.
Open-ended – e.g. What type of foods, vitamins and/or supplements do you give [CHILD] to stay healthy?	For the next items, we'd like you to provide as much information as you can, even things that may not seem important to you.
Behavioral frequency – e.g. How often do you smoke in the presence of your child? Satisfaction with care – e.g. The University of Michigan Health System delivers on its promises. [Five-point response scale - strongly agree to strongly disagree]	For the next few questions, you may need to take time to think carefully before you answer to be the most accurate.

A pilot was conducted between November 2014 and January 2015 before the main study, which was fielded between March and May 2015. The sample for the pilot was relatively small and so statistical power for the analysis was limited with few of the estimates reaching statistical significance. Because the results from the pilot were

[3] http://www.cdc.gov/nchs/data/slaits/survey_sech00.pdf

generally in the same direction as the main study, data from the pilot and the main study were pooled for the analysis presented here. For both the pilot and the main study, a list of child patients between the ages of 4 and 12 who had seen a doctor or nurse practitioner at the pediatrician's office at least once in the previous 12-months was obtained from the Honest Broker Office along with the name and mailing address of the child's parent or guardian. The Honest Broker Office (HBO) is a unit of the Office of Research at the UM Medical School that provides access to electronic health record data for research purposes. A sample was randomly selected for the pilot (n=300). Because of the limited sample for the pilot, patients who had experienced four out of four or three out of four of the types of visits asked about in the survey were oversampled. For the main study, all patients who had experienced four, three or two out of the four types of visits asked about in the survey (visits to the pediatrician, a specialist, the ER, and hospital) were sampled with certainty. The remainder of the sample was then randomly selected from the patients on the list who had experienced only one out of the four visits types to achieve the desired sample size (n=4,700).

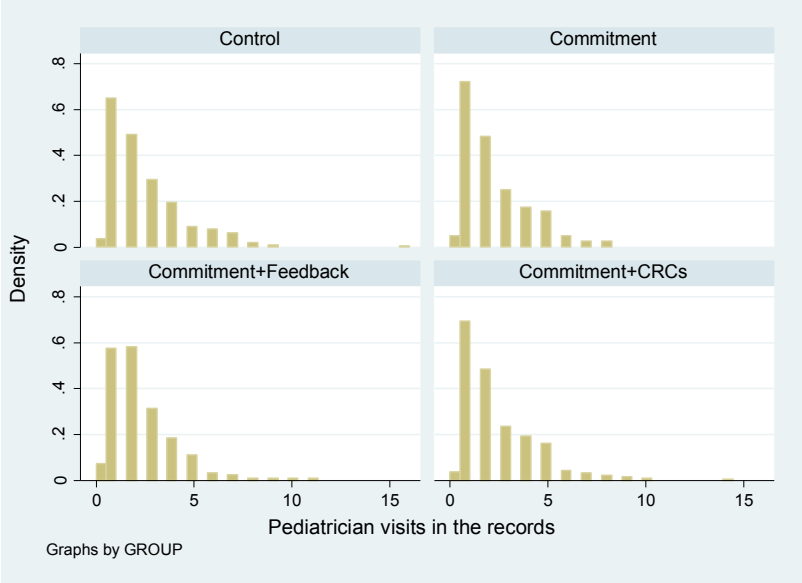
After sorting the list of child patients by the number of visits of each type and by child age, treatment group was assigned systematically in order to ensure a relatively even distribution in the number of visits and child age across the four treatment groups. Table 2.2 provides details about the selected sample composition. As indicated by the percentage of visits by treatment group for each visit type and the distribution of the number of pediatrician visits shown in Figure 2.2 (graphs for the other visits types are provided in the appendix), the frequency of visits in the records for each type are well balanced across the treatment groups and by child age.

Table 2.2. Composition of the Selected Sample

	Commitment only		Commitment and Feedback		Commitment and CRCs		Control			
	n	%	n	%	n	%	n	%		
	1,251	25.0	1,190	23.8	1,313	26.2	1,250	25.0		
Pediatrician visits									p-value of χ^2	χ^2 df
1	435	37.0	460	35.0	450	36.1	439	35.1		
2 – 4	587	49.3	685	52.2	634	50.7	650	52.0	0.834	6
5+	168	14.1	168	12.8	165	13.2	162	13.0		
Specialist visits										
0	162	13.6	151	11.5	158	12.6	170	13.6	0.621	9

1	473	39.8	535	40.8	498	39.8	521	41.7		
2 – 3	366	30.8	395	30.1	381	30.5	348	27.8		
4+	189	15.9	232	17.7	213	17.0	212	17.0		
ER visits										
0	839	70.5	931	70.9	885	70.8	864	69.1		
1	273	22.9	315	24.0	290	23.2	306	24.5	0.677	6
2+	78	6.6	67	5.1	75	6.0	81	6.5		
Hospital visits										
0	1,140	95.8	1,261	96.0	1,199	95.2	1,201	96.0	0.991	3
1+	50	4.2	52	4.0	51	4.1	50	4.0		
Mean child age	8.3		8.3		8.2		8.4			

Figure 2.2. Distribution of Pediatrician Visits in the Records by Treatment Group



Respondents were invited to complete the survey online in a letter sent via postal mail. Dr. Beth Tarini, a practicing Pediatrician at the University of Michigan Health System, Assistant Professor in the Department of Pediatrics, and member of the Child Health Evaluation Research Unit (CHEAR) served as a co-Principle Investigator and co-signed the invitation letters with me to reinforce the legitimacy and value of the study in the eyes of parents. The letters of invitation included a $2 cash pre-paid incentive. One reminder letter was sent to those who had not responded within a week of the initial invitation letter. Respondents were also offered entry into a lottery to win one of 20 $50 gift cards contingent upon completing the survey. The response rate was 27% (RR2 By AAPOR standards).

Table 2.3 shows the composition of the respondent sample including the response rate for each treatment group. The response rate for the Commitment+Feedback group is slightly higher at 34% than that of the other groups, which was 29% for both the Commitment only and Control groups and 27% for the Commitment+CRCs group, but results from a chi-square test indicate that the difference is not significant ($p = 0.807$). The respondent sample is relatively well balanced in the frequency of visits for each type, the relationship to the child, child age, parent age, and parent education. The sample includes a particularly high proportion of respondents with a college or graduate degree. Chi-square tests were conducted on the frequency of each visit type across experimental groups. For visits to a specialist, the ER and the hospital, there was no significant difference across the treatment groups. For visits to the pediatrician, results from the chi-square test indicate that there is a marginal difference ($p = 0.076$). This most likely reflects the slightly lower number of respondents with 5 of more visits in the Commitment+Feedback group and slightly higher number of respondents with between 2 and 4 visits. However, as discussed in more detail below, the number of visits in the records will be entered as a covariate in models to control for differences in the number of visits in the records across treatment groups and respondents within treatment groups.

Table 2.3. Composition of the Respondent Sample

	Commitment only		Commitment and Feedback		Commitment and CRCs		Control		p-value of χ^2	χ^2 df
Response rate	29%		34%		27%		29%		0.807	3
	n	%	n	%	n	%	n	%		
	357	24.3	401	27.0	361	24.3	367	24.7		
Pediatrician visits										
1	139	38.9	133	33.2	137	38.0	129	35.2		
2 – 4	167	46.8	225	56.1	166	46.0	184	50.1	0.076	6
5 or more	51	14.3	43	10.7	58	16.1	54	14.7		
Specialist visits										
0	46	12.5	54	15.1	36	9.0	36	10.0		
1	152	41.4	132	37.0	167	41.7	145	40.2	0.234	9
2 – 3	104	28.3	113	31.6	137	34.2	117	32.4		
4 or more	65	17.7	58	16.3	61	15.2	63	17.5		
ER visits										
0	252	70.6	299	74.6	275	76.2	264	71.9		
1	88	24.7	86	21.5	73	20.2	85	23.2	0.709	6
2 or more	17	4.8	16	4.0	13	3.6	18	4.9		
Hospital visits										

0	342	95.8	391	97.5	349	96.7	356	97.0	0.605	3
1 or more	15	4.2	10	2.5	12	3.3	11	3.0		
Other than child's mother	56	17.2	55	14.9	47	13.9	61	17.4		
Mean child age	8.1		8.2		8.2		8.4			
Mean parent age	40.6		40.8		41.4		41.6			
Parent education:										
High school or less	12.3%		8.7%		9.5%		10.0%		0.443	9
Some college	21.9%		20.4%		22.2%		22.5%			
College degree	28.9%		36.5%		29.0%		29.3%			
Graduate degree	36.9%		34.3%		39.4%		38.2%			

2.3. Outcome Measures and Hypotheses

The study draws on two types of outcome measures to evaluate the effect of commitment: (1) accuracy measures and (2) indirect measures of data quality. The accuracy measures are based on comparisons of answers to medical records and represent an advance in the evaluation commitment and feedback; previous studies relied exclusively on indirect measures, assuming that increased reporting was more accurate (Oksenberg, Vinokur, & Cannell, 1977b, 1977a; Vinokur, Oksenberg, & Cannell, 1977). Validation through medical records makes it possible to detect and better understand the effect of commitment on both underreporting and overreporting.

Accuracy measures: Accuracy is assessed by comparing the reported number of visits as well as the dates of the most recent visit to the pediatrician, a specialist, the ER and the hospital to the information in the child's medical records in several different ways, as detailed below.

Indirect measures: The proposed study also examines the effect of commitment on several indirect measures of data quality. These measures include item nonresponse, the number of mentions and length of responses to open-ended questions, straightlining, acquiescence, socially desirable reporting, break-offs and response time. The rationale behind these data quality measures, most of which were also used for Study 1, was discussed in the previous chapter. In addition to the outcome measures used for Study 1, Study 2 adds the length of open-ended questions and response time. Open-ended

questions are cognitively challenging because they require respondents to search their memory for relevant information, organize and articulate a response in their own words, and judge the adequacy of the response (without response options to use as a guide). Further effort is required for respondents to identify and report more than one response (or mention) and to provide a lengthy response to an open-ended question. The number of mentions and words to open-ended questions is thus a useful indicator of respondent effort and data quality (Krosnick, 1991). One would also expect that committed respondents, if they were upholding the terms of the agreement and reading all of the questions thoroughly and providing complete and accurate answers, would take more time to complete the survey than their control group counterparts.

To summarize, the hypotheses for this study are as follows:

Hypothesis 1: Commitment will decrease the amount of item nonresponse compared to the control condition.

Hypothesis 2: Commitment will increase the accuracy of reported medical visits and the date of the most recent visit, compared to values in the medical records, compared to the control condition.

Hypothesis 3: Commitment will increase the number of mentions and length of responses to open-ended questions compared to the control condition.

Hypothesis 4: Commitment will decrease the amount of straightlining and acquiescence compared to the control condition.

Hypothesis 5: Commitment will decrease the number of socially desirable responses and increases the number of socially undesirable responses compared to the control condition.

Hypothesis 6: Commitment will decrease the number of break-offs.

Hypothesis 7: Commitment will increase survey response time.

2.4.Analytical Methods

To examine the effect of commitment on the outcome measures discussed in this chapter, regression models appropriate to the form of the dependent/outcome variable were constructed with commitment as an independent variable, as shown in Figure 2.3.

Binary indicators were also entered for the feedback and CRC (0=control/reference; 1=treatment) conditions to control for possible effects of these treatments in the presence of commitment. Results for the effect of feedback and CRCs will be discussed in chapter 4.

Figure 2.3. Regression Equation for Estimating the Effect of Commitment

$$\gamma = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \varepsilon$$

γ = the dependent variable/outcome measure
β_0 = intercept
β_1 = Commitment
β_2 = Feedback
β_3 = CRCs

Respondent education level (college or graduate degree versus no college) is entered subsequently into all models and examined for possible interactions with commitment. This is because there is reason to believe that the effect of commitment may vary by respondent education level and past studies have found some differential effects for commitment by education. For example, commitment could be more effective for low education respondents if it increases their motivation. Higher education respondents may have a higher level of motivation, regardless of the treatment, due to increased "need for cognition", which Krosnick (1991) has identified as a possible determinant of respondent motivation. A finding from Oksenberg et al. (1977) (in the analysis presented in Cannell et al. (1981)) that commitment increased reporting to open-ended questions among low education respondents only supports this idea. On the other hand, higher education respondents may have more cognitive resources to draw on than low education respondents for certain response tasks. Conrad et al.'s (in press) finding that higher education respondents answered numeracy questions more accurately when they had made a commitment (and were prompted for answering too quickly), whereas lower education respondents did not improve with commitment, supports this theory. However, other studies have found no differential effects of commitment by education level (Miller & Cannell, 1982; Oksenberg et al., 1977a; Vinokur et al., 1977).

Several measures are used to assess the accuracy and direction of error of reported visits by comparing the reported values to the values recorded in the medical record as follows:

- A **binary metric** for a match between the reported number of visits and the number of visits in the records versus no match
- The **simple (signed) difference** between the reported number of visits and the number of visits in the records
- The **absolute difference** between the reported number of visits and the number of visits in the records
- A **count metric** of the count of reported visits relative to reported visits for the reference group
- A **tri-variate measure** – 0 for a match between the reported number of visits and the number of visits in the records, 1 if the difference between reported visits and the records is greater than zero (i.e. overreporting), and 2 if the difference is less than zero (i.e. underreporting)

Each of these measures has strengths and limitations. For example, logistic regression, using a binary metric for match versus no match, may be the most straightforward measure of accuracy – the reported value either matches that in the records or it does not. But this approach is restrictive in that it treats all mismatches equally whether the reported value is off by one visit or off by many and does not provide information on the distance between the reported value and the value in the medical records. It also contains no information about the direction of the error (overreporting versus underreporting).

The simple difference between the reported number of visits and the visits in the records captures directionality of error but if underreports and overreports are relatively symmetrical, then averaging across them effectively cancels out the errors, potentially underestimating the overall amount of error. The absolute value of the difference lacks the directionality of the error provided by the simple difference but captures the magnitude of error. However, it also tends to place greater importance on cases with high visit frequency, where, due to the difficulty of the task, the amount of error is likely to be

greater. The absolute value of the difference also confounds under and overreporting – treating an under-reported visit the same way as an over-reported visit.

A ratio of the count of reported visits relative to the reported visits in the reference group used in a Poisson regression model predicts the count of reported visits while controlling for the number of visits recorded in the records. This measure quantifies the extent of and direction of reporting or what could be seen as the tendency to report visits – a higher count or a lower count – taking into account the actual frequency of visits in the records. The ratio metric is similar to the simple difference measure in that it indicates the overall direction of error, but it also takes into account the actual frequency of visits by controlling for the visits in the records.

The tri-variate measure using multinomial regression makes it possible to examine the components of error – overreporting and underreporting – as opposed to the gross odds of error as estimated with the simple logistic model on a binary outcome measure. But it still does not tell us much about the magnitude of the error. That is because, like the simple binary measure, the trivariate measure treats all mismatches greater than zero and mismatches less than zero equally whether the reported value is off by one visit or off by many.

While each of these measures has limitations, taken together, each of these measures offers insight into the effect of commitment on the number of reported visits and how the reported value compares to the number in the child's medical record – the likelihood of a match, the magnitude of the difference between the reported number of visits and the number in the records, the ratio or tendency to report visits, and the likelihood of an overreport or an underreport. Results from a combination of these measures for the four types of visits asked about in the survey – visits to the pediatrician's office, a specialist, the ER, and the hospital – are discussed below.

For models assessing the accuracy of reported visits, in addition to education, the number of visits in the records, and whether the respondent reported checking records are entered into the models as independent variables along with interaction terms with commitment. The main effect of the number of visits in the records controls for the difficulty of the task, the idea being that the more visits the respondent's child has had in the last 12-months, the more difficult it is to recall the exact number of visits accurately.

In the context of the binary metric for a match versus no match with the records, a positive and significant interaction with commitment would suggest that, as the number of visits increased, and the response task presumably became more difficult, committed respondents were more likely to report a matching number of visits, in other words, maintained a higher level of accuracy. Whether the respondent reported checking a calendar or other relevant documents (checking records) is also entered as a main effect; presumably records are more accurate than recall or other estimation processes, although records may contain error as well. Interaction terms for checking records and commitment are also entered to see if the effect of checking records differed between commitment and control group respondents. For example, checking records could have more of an effect for commitment group respondents if they checked multiple types of records or checked their records more thoroughly.

Three measures are used to examine the accuracy and quality of the reported dates:

- A **binary metric** for a match between the reported date and date in the records versus no match (using logistic regression)
- The **absolute difference** between the reported date and the date in the records (using linear regression)
- A binary metric for whether or not the respondent reported a full date (including a day, month, and year) (using logistic regression) is also examined.

For models assessing the accuracy of reported dates, in addition to education, the number of days since the date of the last visit in the records, and whether the respondent reported checking records are entered into the models along with interaction terms with commitment. The main effect of the number of days since the date of the last visit in the records controls for the difficulty of the task, the idea being that as the number of days since the last visit increases, the more difficult it is to recall the date accurately. In the context of the binary metric for a match versus no match with the records, if commitment increases respondent motivation and effort to provide accurate answers, a positive and significant interaction would suggest that, as the number of days since the last visit increased, and the response task presumably became more difficult, committed respondents were more likely to report a matching date.

2.5. Results

2.5.1. Response to the Request for Commitment

As mentioned above, respondents receiving the commitment treatment were presented with a series of checkboxes corresponding to desired response behaviors. Respondents could select one or more of the checkboxes or none of the above. Respondents who selected none of the above were still able to proceed with the survey.

As shown in Table 2.4, 63% of respondents checked all of the checkboxes. Twenty-seven percent of respondents selected four out of five checkboxes. Of the respondents who selected four checkboxes, 295 out of 302 checked all of the checkboxes except for the one for "looking up information in records or on a calendar, if needed". The remaining respondents selected some combination of two or three checkboxes, only one checkbox, or "none of the above". Among those who selected three checkboxes, 56% agreed to read all of the questions carefully, trying to be as precise as possible with answers, and answering honestly, and did not agree to look up information or provide as much information as possible. Another 19% of the respondents selecting three checkboxes agreed to read all of the questions carefully, answer honestly, and provide as much information as possible, excluding looking up information and trying to be as precise as possible. All seven of the respondents selecting two checkboxes agreed to either read all of the questions carefully or answer honestly. None of these respondents agreed to look up information. Among the respondents selecting only one checkbox, 46% agreed to read all of the questions carefully and 27% agreed to answer honestly. A very small number of respondents, that is .5%, checked the box for none of the above.

Table 2.4. Commitment Checkboxes Selected

Number of commitment checkboxes selected	**n**	**%**
5 checkboxes	700	63%
4 checkboxes	302	27%
3 checkboxes	43	4%
2 checkboxes	7	.6%
1 checkbox	56	5%
None	6	.5%
Total	1,114	100%

The results shown above demonstrate considerable heterogeneity in how respondents responded to the commitment treatment. While this chapter examines how respondents in the commitment treatment groups differ collectively from the control group, the following chapter explores the effect of commitment level, i.e., how the number of behaviors and the particular behaviors the respondents committed to perform affected response accuracy and other indirect indicators of data quality.

2.5.2. Item Nonresponse

This section examines the effect of commitment on item nonresponse overall and specifically to questions about medical visits, which could be considered particularly burdensome compared to the other questions in the survey. Respondents were asked up to forty-two questions, but not all respondents were asked the same number of questions. This is because treatment group respondents were asked some additional debriefing items not posed to control group respondents. Further, not all respondents were asked all of the most recent medical visit date questions – as explained below. Respondents in the treatment group were also asked to respond to the commitment statement. The proportion of item nonresponse overall was examined in two ways: 1) the number of questions not answered out of the total number of questions asked for each respondent, and 2) the total number of questions asked of all respondents (excluding the treatment-related debriefing items and commitment statement posed to the treatment group only). Results were nearly identical so only one set of results, those based on the total number of questions asked (including treatment-related items) is displayed.

For four types of visits, respondents were asked to report the number of visits in the last 12 months and the date of the most recent visit: visits to the pediatrician's office, a specialist, the ER, and to the hospital. All respondents were asked each of the four medical visit questions but only those respondents who reported at least one of a given type of visit were asked to report the date of the most recent visit of that type. Respondents were therefore asked at least four and up to eight of the medical visit questions. As with overall item nonresponse, differences in the number of questions asked were accounted for when calculating the proportion of questions not answered for each respondent.

Comparing the mean proportion of item nonresponse overall and to questions about the number of medical visits and the date of the most recent visit, shown in Table 2.5, we see that the proportion of item nonresponse is greater for the commitment group than the control group for item nonresponse overall and for medical visit and date questions, which fails to support Hypothesis 1. Regression results shown in Table 2.6 show no significant effect of commitment on the proportion of item nonresponse overall. However, results for the medical visit and date questions indicate a significantly higher proportion of item nonresponse for the commitment group.

Table 2.5. Proportion of Item Nonresponse Overall and for Medical Visit and Date Questions

	Overall Item Nonresponse		Item Nonresponse to Medical Visit and Date Questions	
	p	(se)	p	(se)
Control	0.16	0.008	0.07	0.009
Commitment	0.17	0.006	0.08	0.007

Table 2.6: The Main Effect of Commitment on the Proportion of Item Nonresponse Overall and for Medical Visit and Date Questions

VARIABLES	Percent Overall Item Nonresponse	Percent Item Nonresponse to Visit and Date Questions
Commitment	0.02	0.03*
	(0.014)	(0.016)
Feedback	-0.03*	-0.03^
	(0.014)	(0.015)
CRCs	-0.02	-0.03^
	(0.014)	(0.016)
Constant	0.16***	0.07***
	(0.010)	(0.011)
Observations	1,486	1,486
R-squared	0.003	0.004

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

2.5.3. Response Accuracy: Medical Visit Questions

We saw in the previous section that commitment group respondents were less likely to respond to the medical visit and date questions. In this section, we explore the extent to which commitment group respondents were more or less accurate in their responses to the medical visit questions than the control group, when they did respond.

Distributions of the differences between reported and visits in the records for visits to the pediatrician by treatment group, shown in Figure 2.3, indicate a large amount of clustering at and around zero in both directions, with slightly more positive reports suggesting overreporting in all groups. Similarly, Figure 2.4, with the distribution of the absolute differences by treatment group, shows a large number of absolute differences at zero, one and two. Similar figures are provided in the appendix for visits to a specialist, the ER, and hospital. The difference between reported visits and visits in the records and the ratio metric are similar in that they indicate an overall direction of reporting. However, the ratio metric has the added advantage of taking into account differences in the total number of visits between respondents. The large number of absolute differences at and around zero for visits to the pediatrician and the other visit types could result in biased estimates in ordinary least squares (OLS) regression, which assumes a relatively normal distribution for the dependent variable. Therefore, the analysis presented below focuses on the binary metric (match versus no match) using logistic regression, the tri-variate metric (0=match, 1=overreport, 2=underreport) using multinomial logistic regression, and the ratio metric (reported visits to visits in records) using Poisson regression.

Also of note are some potentially extreme observations in the distributions of the differences and absolute differences between reported visits and visits in the records shown in Figures 2.4 and 2.5 for visits to the pediatrician. A few potentially extreme observations can also be observed for the other visits types – similar graphs for these visit types are provided in the appendix. For the regression analyses presented below, models were fitted including all observations and then refitted excluding potentially extreme observations. In cases where extreme observations changed the direction or level of significance of key parameters, results are shown with these observations excluded and this is noted in the text.

Figure 2.3. Visits to the Pediatrician – Difference between Reported Visits and Records

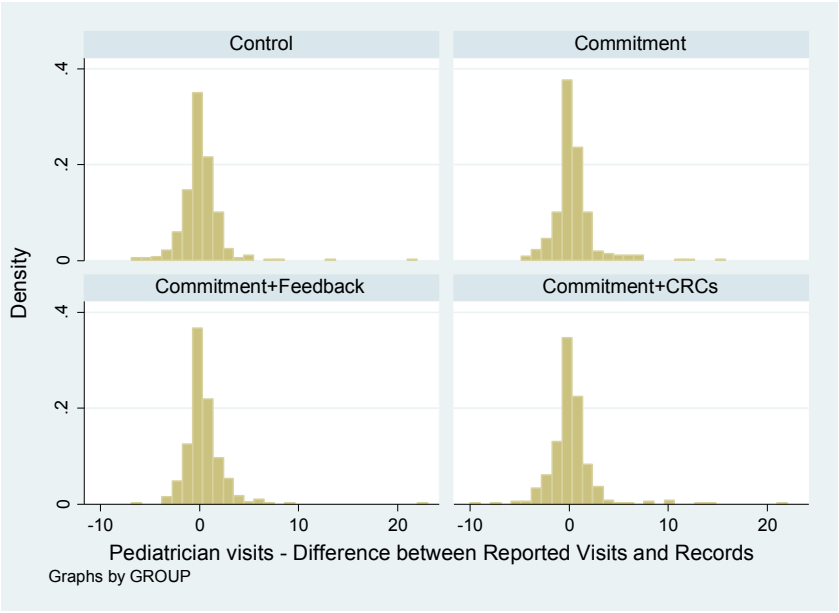
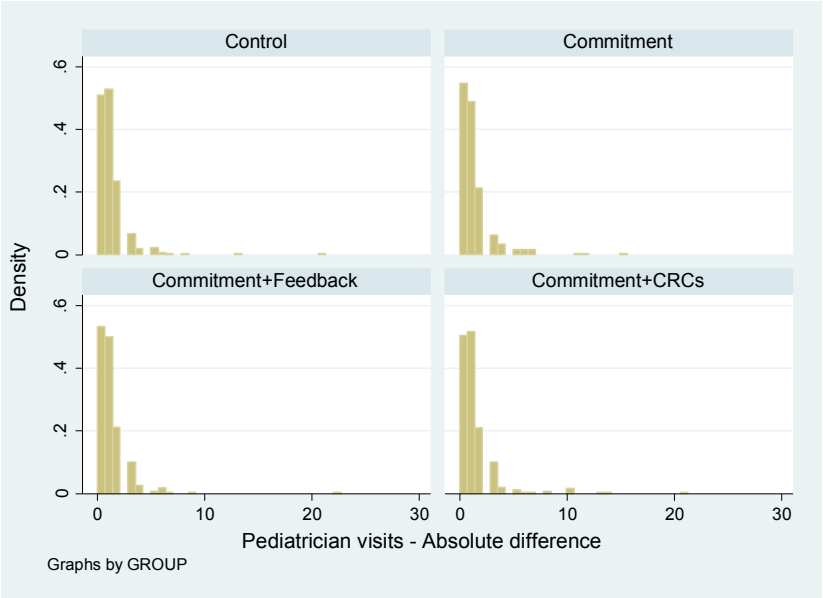


Figure 2.4. Visits to the Pediatrician – Absolute Difference Between Reported and Visits in the Records



If commitment helped improve response accuracy to the visit questions, we would expect higher percentages of a match and lower absolute differences between the reported number of visits and the visits in the records for the commitment group compared to the control group. As shown in the descriptive results in Table 2.7, this does not seem to be the case: both measures are quite small indicating that responses were quite accurate

overall. For percentage of a match, the results are in the expected direction for visits to the pediatrician but not for visits to a specialist or the hospital. There is no difference for visits to the ER. For the absolute difference, the results are in the expected direction for visits to a specialist and to the hospital but not for visits to the pediatrician or to the ER. Responses are generally quite accurate, in terms of the percentage of a match and absolute difference, for visits to the ER and hospital, which could limit the potential effect of the commitment treatment (i.e. a ceiling effect).

The small differences between reported visits and those in the medical record were not significant in the logistic regression results for the main effect of commitment, shown in Table 2.8, failing to support Hypothesis 2. However, results for the ratio metric estimated using Poisson regression, also shown in Table 2.8, suggest differences in the direction of reporting for the different visits types. Positive significant coefficients in the Poisson models for commitment for visits to the pediatrician, specialist, and hospital indicate that the treatment group respondents reported a higher count of visits on average than control group respondents for these types of visits, which could indicate either fewer underreported visits, overreported visits, or some combination thereof.

Results for the tri-variate measure – (0=match, 1=overreport, 2=underreport) – using multinomial regression, are shown in Table 2.9. If commitment helped improve accuracy, we would expect to see significantly less overreporting or underreporting. As we would expect for visits to the pediatrician, we see that the commitment group respondents were significantly less likely to underreport visits to the pediatrician. This suggests that the higher ratio of reported visits seen in Table 2.8 may reflect, at least in part, fewer underreported visits. However, results for hospital and ER visits are not as we would expect. The positive and significant coefficient for commitment for overreporting of hospital visits suggests that the higher ratio of reported hospital visits seen in the ratio metric might reflect some overreporting of these visits. We also see that commitment group respondents were marginally more likely to overreport visits to the ER.

Table 2.7. Percentage of Matches and Mean Absolute Differences between Reported Visits and Records for Commitment Treatment versus Control and by Level of Commitment

	Visits to the Pediatrician	Visits to a Specialist

	Match		Absolute difference		Match		Absolute difference	
	%	(se)	Mean	(se)	%	(se)	Mean	(se)
Control	0.36	0.03	1.17	0.09	0.44	0.03	1.42	0.28
Commitment	0.37	0.01	1.22	0.06	0.43	0.02	1.41	0.13

	Visits to the ER				Visits to the Hospital			
	Match		Absolute difference		Match		Absolute difference	
	%	(se)	Mean	(se)	%	(se)	Mean	(se)
Control	0.87	0.02	0.21	0.04	0.94	0.01	0.07	0.02
Commitment	0.87	0.01	0.16	0.02	0.93	0.01	0.09	0.01

Table 2.8. The Effect of Commitment on Reported Visits to the Pediatrician, Specialist, ER, and Hospital

	Visits to the Pediatrician		Visits to a Specialist	
VARIABLES	Logistic	Poisson	Logistic	Poisson
Commitment	0.12	0.10*	0.13	0.11*
	(0.157)	(0.045)	(0.154)	(0.050)
Feedback	-0.05	-0.01	-0.15	0.02
	(0.154)	(0.044)	(0.152)	(0.049)
CRCs	-0.14	-0.06	-0.32*	0.14**
	(0.158)	(0.045)	(0.157)	(0.050)
Visits in records		0.18***		0.16***
		(0.006)		(0.003)
Constant	-0.57***	-11.49***	-0.24*	-11.77***
	(0.111)	(0.038)	(0.107)	(0.040)
Observations	1,422	1,420	1,397	1,395

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Visits to the ER		Visits to the Hospital	
VARIABLES	Logistic	Poisson	Logistic	Poisson
Commitment	-0.13	0.03	-0.12	0.46*
	(0.222)	(0.112)	(0.315)	(0.232)
Feedback	0.35	-0.29*	0.34	-0.55*
	(0.229)	(0.122)	(0.327)	(0.237)
CRCs	0.09	-0.47***	-0.29	0.03
	(0.222)	(0.126)	(0.292)	(0.206)
Visits in records		0.66***		0.47***
		(0.019)		(0.029)
Constant	1.90***	-13.22***	2.76***	-14.47***
	(0.158)	(0.081)	(0.225)	(0.181)
Observations	1,400	1,398	1,399	1,397

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Table 2.9. The Effect of Commitment on Overreporting and Underreporting

VARIABLES	Visits to the Pediatrician	Visits to a Specialist	Visits to the ER	Visits to the Hospital
logit 2: Overreporting versus zero difference				
Commitment	0.08	0.13	0.41^	0.47*
	(0.168)	(0.182)	(0.214)	(0.239)
Feedback	0.00	-0.07	-0.41^	-0.22
	(0.162)	(0.178)	(0.209)	(0.221)
CRCs	-0.06	0.20	-0.29	-0.12
	(0.169)	(0.180)	(0.210)	(0.223)
Constant	0.15	-0.51***	-1.95***	-2.31***
	(0.121)	(0.131)	(0.161)	(0.183)
logit 3: Underreporting versus zero difference				
Commitment	-0.41*	-0.21	0.02	-12.03
	(0.206)	(0.178)	(0.381)	(426.456)
Feedback	0.11	0.30^	0.14	12.66
	(0.209)	(0.174)	(0.361)	(426.456)
CRCs	0.40^	0.33^	0.15	13.48
	(0.208)	(0.182)	(0.370)	(426.455)
Constant	-0.34*	-0.26*	-3.02***	-5.81***
	(0.137)	(0.122)	(0.264)	(1.002)
Observations	1,486	1,486	1,486	1,486

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

The following covariates were added to the logistic, Poisson, and multinomial models as main effects and as interactions with commitment: the number of visits in the records for the visit type, whether or not the respondent reported checking a calendar or other relevant documents, and the respondent's education level (college degree or higher). Interactions terms for checking records and commitment were dropped from the models because none was significant. Results for the logistic and Poisson models are shown in Table 2.10 and in Table 2.11 for the multinomial models.

As expected, there is a significant main effect for the number of visits in the records for most measures. The number of visits in the records is a negative predictor of a match between reported visits and visits in the records for each of the visit types. It is a positive predictor in the Poisson models for each of the visit types, suggesting that respondents with higher numbers of visits in the records reported a significantly higher ratio of visits, or overreported, on average. It is also a positive predictor of overreporting for visits for three out of four of the visit types.

If commitment helped improve accuracy as the number of visits increased, we would expect to see a positive and significant interaction between commitment and the number of visits in the records in the logistic regression models. This is the case for visits to a specialist, where there is also a significant interaction between commitment and education in the logistic model, suggesting that respondents in the commitment group were more likely to report a matching number of visits as the number of visits increased – particularly among respondents with a college degree or higher education. The regression lines shown in Figure 2.5, where the slopes for commitment (grey and yellow) are less steep than those for the control group as the number of visits increases, illustrates this. This result is also reflected in the multinomial results in Table 2.11, where we see corresponding negative and significant interactions between commitment and the number of visits and between commitment and education for both overreporting and underreporting – so less error in both directions – particularly among those with higher education. This result offers partial support for Hypothesis 2; Commitment improved accuracy for a particular subset of respondents. However, while commitment group respondents were more accurate as the number of visits increases, we also see in Figure 2.5 that they were less likely to report a matching number of visits at zero visits. This is also reflected in the multinomial results in 2.11, where the coefficient for the main effect of commitment is significant indicating more overreporting at zero visits.

The interaction between commitment and the number of visits in the records is not significant in the logistic regression models for the other visits types, nor is it significant in the Poisson models for any of the visit types. This suggests that there was no significant difference in the direction of error between the treatment and control groups as the number of visits increased. Parameter estimates in the Poisson models for the interaction between commitment and the number of visits in the records were sensitive to a few extreme observations for each of the visit types. The extreme observations were therefore excluded from the models shown in Table 2.10.

Checking records significantly increased the likelihood of a match for visits to the pediatrician and the ER but marginally decreased the odds of a match for hospital visits.

Table 2.10. The Effect of Commitment and Covariates – Visits to the Pediatrician, Specialist, ER, and Hospital

	Visits to the Pediatrician		Visits to a Specialist	
VARIABLES	Logistic	Poisson	Logistic	Poisson
Commitment	-0.55^	-0.02	-1.06**	-0.07
	(0.336)	(0.045)	(0.340)	(0.052)
Feedback	-0.09	-0.07	-0.09	0.13*
	(0.164)	(0.046)	(0.163)	(0.052)
CRCs	-0.11	0.11**	-0.27	0.25***
	(0.168)	(0.035)	(0.167)	(0.041)
Checked	0.25*	0.13	0.16	0.10
	(0.124)	(0.095)	(0.124)	(0.091)
Visits in records	-0.33***	0.18***	-0.78***	0.16***
	(0.079)	(0.016)	(0.115)	(0.005)
Visits in records * Commitment	0.08	0.01	0.40**	-0.01
	(0.090)	(0.018)	(0.124)	(0.006)
College or higher	-0.09	0.03	-0.44	-0.02
	(0.252)	(0.071)	(0.274)	(0.078)
College or higher * Commitment	0.56^	-0.11	0.71*	0.01
	(0.293)	(0.081)	(0.310)	(0.090)
Constant	0.18	-11.56***	1.34***	-11.92***
	(0.280)	(0.082)	(0.296)	(0.077)
Observations	1,333	1,332	1,357	1,353

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Visits to the ER		Visits to the Hospital	
VARIABLES	Logistic	Poisson	Logistic	Poisson
Commitment	-0.29	-0.27*	-0.26	-0.28
	(0.381)	(0.124)	(0.474)	(0.251)
Feedback	0.33	-0.09	0.34	0.26
	(0.252)	(0.125)	(0.338)	(0.218)
CRCs	0.05	0.30**	-0.32	0.27
	(0.246)	(0.094)	(0.299)	(0.181)
Checked	0.49**	-0.22	-0.43^	0.10
	(0.182)	(0.171)	(0.243)	(0.388)
Visits in records	-1.17***	0.75***	-1.13*	1.42***
	(0.215)	(0.038)	(0.536)	(0.158)
Visits in records * Commitment	-0.20	0.06	0.30	0.02
	(0.261)	(0.049)	(0.626)	(0.201)
College or higher	0.40	-0.64***	0.34	-0.14
	(0.360)	(0.156)	(0.484)	(0.371)
College or higher * Commitment	0.18	0.33^	0.23	0.11
	(0.416)	(0.188)	(0.548)	(0.425)
Constant	2.08***	-13.17***	2.89***	-14.60***
	(0.315)	(0.134)	(0.405)	(0.329)
Observations	1,364	1,363	1,366	1,364

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Figure 2.5. Regression Lines for the Likelihood of a Match for Visits to a Specialist by Treatment and Education Level and Visits in the Records

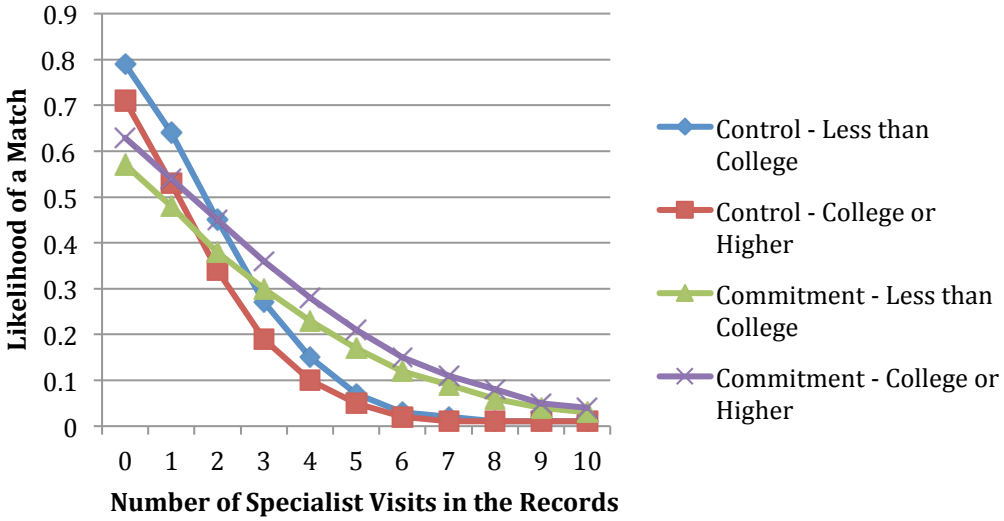


Table 2.11. The Effect of Commitment and Covariates on Overreporting and Underreporting

VARIABLES	Visits to the Pediatrician	Visits to a Specialist	Visits to the ER	Visits to the Hospital
logit 2: Overreporting versus zero difference				
Commitment	0.53	1.22**	0.20	0.17
	(0.362)	(0.398)	(0.394)	(0.459)
Feedback	0.04	-0.19	-0.65*	-0.45
	(0.172)	(0.202)	(0.295)	(0.341)
CRCs	-0.04	0.24	-0.15	0.19
	(0.178)	(0.198)	(0.271)	(0.303)
Checked	-0.14	0.19	-0.09	0.55*
	(0.131)	(0.155)	(0.209)	(0.248)
Visits in records	0.08	0.78***	1.09***	1.35**
	(0.095)	(0.115)	(0.221)	(0.496)
Commitment * Visits in records	-0.02	-0.44***	0.03	-0.67
	(0.108)	(0.126)	(0.277)	(0.584)
College or higher	-0.00	0.36	-0.93*	-0.35
	(0.268)	(0.332)	(0.395)	(0.460)
Commitment * College or higher	-0.52^	-0.63^	0.19	-0.18
	(0.311)	(0.375)	(0.459)	(0.530)
Constant	-0.02	-2.36***	-2.22***	-2.85***
	(0.305)	(0.346)	(0.322)	(0.389)
logit 3: Underreporting versus zero difference				
Commitment	0.45	0.83*	1.06	-2.98
	(0.496)	(0.368)	(0.929)	(2,516.984)
Feedback	0.16	0.29	0.25	23.79

	(0.237)	(0.188)	(0.415)	(1,250.519)
CRCs	0.35	0.27	0.11	24.19
	(0.235)	(0.195)	(0.429)	(1,250.519)
Checked	-0.57**	-0.44**	-1.40***	-0.85
	(0.173)	(0.140)	(0.317)	(1.236)
Visits in records	0.69***	0.77***	1.42***	4.86
	(0.100)	(0.114)	(0.270)	(293.943)
Commitment* Visits in records	-0.11	-0.34**	0.45	-0.40
	(0.115)	(0.124)	(0.335)	(293.946)
College or higher	0.20	0.41	1.36^	3.00
	(0.341)	(0.299)	(0.826)	(2,016.543)
Commitment * College or higher	-0.60	-0.69*	-1.52^	-4.15
	(0.397)	(0.341)	(0.896)	(2,016.544)
Constant	-2.19***	-1.67***	-4.47***	-26.47
	(0.411)	(0.314)	(0.832)	(2,184.355)
Observations	1,370	1,370	1,370	1,370

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

2.5.4. Response Accuracy: Date Questions

Respondents who reported at least one visit for their child in the last 12 months to the pediatrician's office, a specialist, the ER and the hospital were asked to report the date of the most recent visit. Drop-downs appeared for the month, day and year for respondents to use to enter the date.

Table 2.12 provides the percentage of matches and the mean absolute differences between reported dates and dates in the records for commitment treatment versus control. The percentages of matches for visits to the ER and hospital are not reported as there were too few matches (4 and 1 respectively). Examining the distribution of absolute differences between the reported date of the last visit to the pediatrician and the date in the records shown in Figure 2.6, we see that the distribution is highly skewed to the right, violating the assumption of normality. This is also the case for the visits to a specialist, the ER and hospital – figures for which are shown in Appendix 3. Therefore, the analyses presented below focus on the binary metrics (match versus no match and full date (dd/mm/yyyy) versus mm/yyyy or dd/yyyy or dd/mm) using logistic regression.

As shown in the results estimating the main effect of commitment in Table 2.13, the difference in the percentage of matches between commitment and the control group for visits to the pediatrician is positive and statistically significant indicating that

commitment group respondents reported dates more accurately than control respondents. Differences for the remaining measures are not statistically significant.

Table 2.12. Percentage of a Match and Mean Absolute Differences for Date of Last Visit by Commitment

	Date of last visit to the Pediatrician					
	Match		Absolute difference		Full Date	
	%	(se)	%	(se)	%	(se)
Control	0.38	0.03	37.02	4.44	0.84	0.02
Commitment	0.47	0.02	36.67	2.54	0.84	0.01

	Date of last visit to a Specialist					
	Match		Absolute difference		Full Date	
	%	(se)	%	(se)	%	(se)
Control	0.41	0.04	48.64	6.20	0.80	0.03
Commitment	0.50	0.02	39.01	3.32	0.83	0.01

	Date of last visit to the ER				Date of last visit to the Hospital			
	Absolute difference		Full Date		Absolute difference		Full Date	
	%	(se)	%	(se)	%	(se)	%	(se)
Control	32.30	11.51	0.82	0.04	42.57	36.79	0.87	0.07
Commitment	32.18	6.24	0.79	0.03	32.73	21.87	0.80	0.04

Figure 2.6. Date of Last Visit to the Pediatrician - Absolute Difference Between the Reported Date and Date in the Records

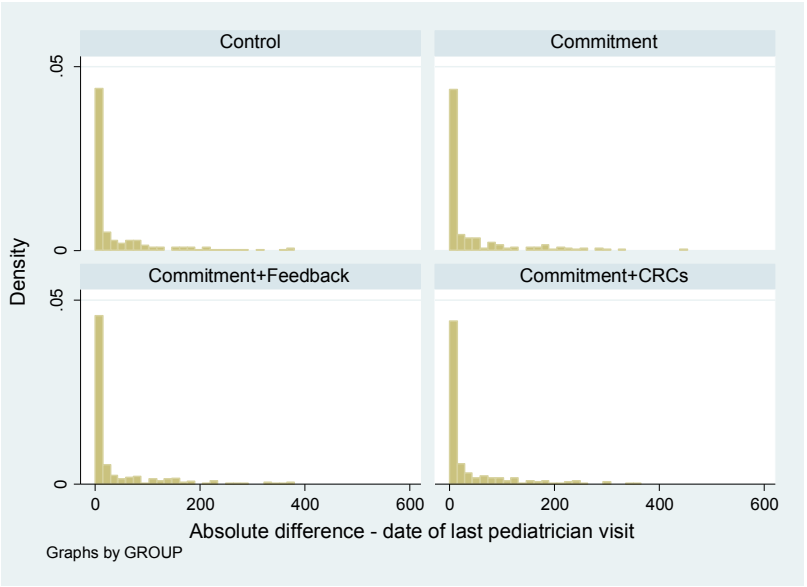


Table 2.13. The Main Effect of Commitment – Date of the Most Recent Visit to the Pediatrician, Specialist, ER, and Hospital

	Date of last visit to the Pediatrician		Date of last visit to a Specialist		Date of last visit to the ER	Date of last visit to the Hospital
VARIABLES	Logistic	Full date	Logistic	Full date	Full date	Full date
Commitment	0.46*	-0.21	0.34	-0.09	-0.56	-0.60
	(0.186)	(0.219)	(0.212)	(0.231)	(0.363)	(0.772)
Feedback	-0.10	0.45*	0.13	0.75**	1.12**	0.60
	(0.175)	(0.218)	(0.200)	(0.254)	(0.410)	(0.772)
CRCs	-0.16	0.23	-0.21	0.18	0.23	-0.11
	(0.183)	(0.218)	(0.210)	(0.234)	(0.358)	(0.631)
Constant	-0.48***	1.65***	-0.43**	1.40***	1.52***	1.90**
	(0.130)	(0.158)	(0.148)	(0.163)	(0.276)	(0.619)
Observations	1,039	1,236	795	963	346	104

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

The following covariates were examined as main effects and as interactions with commitment for the date of the last visit for both the likelihood of a match and that a full date was reported: the number of days since that last visit in the records, whether or not the respondent reported checking a calendar or other relevant documents, and the respondent's education level (college degree or higher). Results for the date of last visit to the pediatrician and a specialist are shown in Table 2.14. There were no significant

effects for any covariates or interactions for the likelihood of reporting a full date for the last visit to the ER or to the hospital.

The length of time since the last visit in the records is a significant negative predictor of a match and that a full date was reported (marginally so for last visit to a specialist), while checking records is a positive predictor. The interaction between checking records and commitment level was not significant and was subsequently dropped from the models.

If commitment helped improve accuracy as the number of days since the last visit in the records increased, we would expect to see positive and significant interactions between commitment and the number of days in the records in the logistic regression models. We see a significant interaction between the length of time since the last visit in the records and commitment for the likelihood of a match for the date of the last visit to a specialist. Comparing the regression lines for the likelihood of a match for the last visit to a specialist, shown in Figure 2.7, the slope of the line is less steep for commitment group respondents suggesting that commitment helped these respondents maintained a higher level of accuracy even as the amount of time since the date of the last visit increased, and presumably the recall task became more difficult. This result offers partial support for Hypothesis 2; Commitment improved accuracy for a particular subset of respondents.

Table 2.14. The Effect of Commitment and Covariates – Date of the Most Recent Visits to the Pediatrician and Specialist

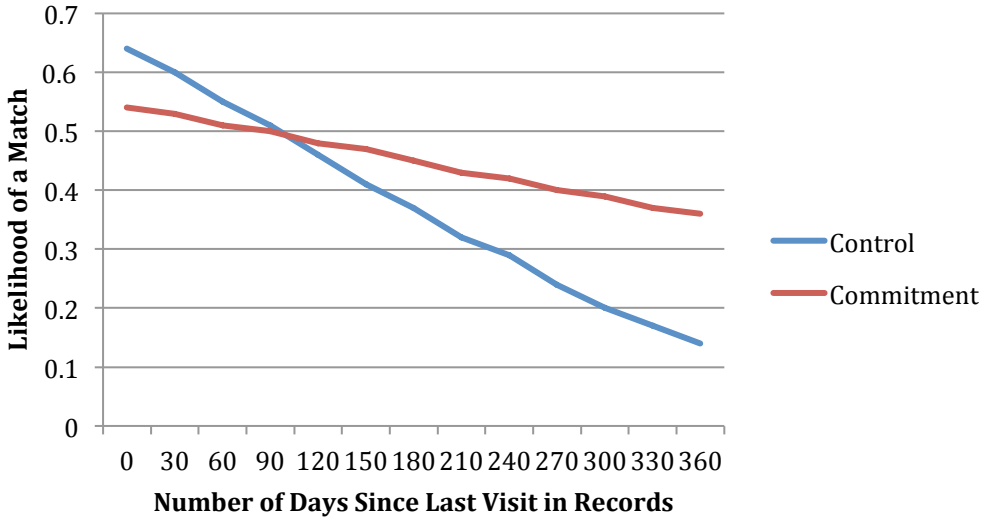
	Date of Last Visit Pediatrician's Office		Date of Last Visit to a Specialist	
VARIABLES	Match	Full date	Match	Full date
Commitment	0.09	-0.28	-0.72	-0.96^
	(0.400)	(0.445)	(0.450)	(0.522)
Feedback	-0.18	0.40^	0.09	0.74*
	(0.202)	(0.237)	(0.227)	(0.291)
CRCs	-0.24	0.30	-0.17	0.30
	(0.210)	(0.238)	(0.237)	(0.272)
Checked records	2.27***	1.65***	2.39***	1.99***
	(0.210)	(0.177)	(0.252)	(0.206)
Recent visit in records	-0.01***	-0.00*	-0.01***	-0.00^
	(0.002)	(0.002)	(0.002)	(0.002)
Commitment * Recent visit in records	0.00	-0.00	0.01*	0.00
	(0.002)	(0.002)	(0.002)	(0.002)
College or higher	0.50	0.18	0.05	-0.63
	(0.351)	(0.358)	(0.384)	(0.394)

Commitment * College or higher	-0.11	-0.36	0.29	0.53
	(0.397)	(0.417)	(0.437)	(0.463)
Constant	-1.49***	1.38***	-1.19**	1.44***
	(0.356)	(0.362)	(0.395)	(0.424)
Observations	988	1,176	747	896

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Figure 2.7. Regression Lines for the Likelihood of a Match for Reported Date of Last Visit to a Specialist by Commitment Level Group and Time Since the Last Visit to a Specialist



2.5.5. Number of Mentions and Words to Open-ended Questions

Respondents were asked the following open-ended questions:

1. What types of foods, vitamins and/or supplements do you give [CHILD] to stay healthy?
2. What have you cut down on giving [CHILD] because you think that it is bad for his/her health?
3. What would you like to do to maintain or improve [CHILD]'s health?

Three one-line text fields appeared for each question. Respondents in the Commitment+Feedback group received feedback if they left the second or third line blank.

The mean number of words and completed response fields, referred to here as mentions, in responses to the open-ended questions, presented in Table 2.15, is higher for

respondents in the commitment group than in the control group but the differences are not statistically significant, as shown in Table 2.16, failing to support Hypothesis 3. There was no significant main effect for college or higher education level or for the interaction education and commitment.

Table 2.15. Mean Number of Mentions and Word Count to Open-ended Questions by Commitment Level

	Word count		Mentions	
	Mean	(se)	Mean	(se)
Control	12.59	0.08	4.36	0.12
Commitment	14.73	0.47	4.93	0.08

Table 2.16. The Main Effect of Commitment on the Number of Mentions and Word Count to Open-ended Questions

VARIABLES	Word count pooled	Mentions pooled
Commitment	1.04	0.12
	(0.964)	(0.158)
Feedback	3.32***	1.16***
	(0.936)	(0.154)
Constant	12.52***	4.40***
	(0.784)	(0.129)
Observations	1,486	1,486
R-squared	0.012	0.045

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

2.5.6. Straightlining and Acquiescence

Answering the same way to five or more out of six (e.g. "near straightlining") items in a battery (Yes/No) is defined as straightlining. This dependent variable includes "near straightlining" because complete straightlining (e.g. the same answer to all six out of six items) was rare – just two cases – mostly likely because the battery of items included a reverse-coded item. For acquiescence, the dependent variable is the number of "agree" or "strongly agree" answers in a battery of questions.

As shown in Table 2.17, the percentage of respondents who straightlined and acquiesced is lower for the commitment group than for the control group. As shown in Table 2.18, this difference is not statistically significant for straightlining but it is significant for acquiesce, lending mixed support for Hypothesis 4. There was no

significant main effect of having a college degree or higher or for the interaction between education and commitment.

Table 2.17. Percentage Straightlining and Mean Number of "Agree" or "Strongly Agree" Answers for Commitment Level

	Straightlining		Acquiescence	
	%	(se)	%	(se)
Control	0.49	0.03	4.33	0.07
Commitment	0.47	0.01	4.17	0.05

Table 2.18. The Main Effect of Commitment on Straightlining and Acquiescence

VARIABLES	Straightlining	Acquiescence
Commitment	-0.06	-0.19*
	(0.128)	(0.096)
Feedback	-0.09	0.07
	(0.125)	(0.093)
Constant	-0.05	4.33***
	(0.104)	(0.078)
Observations	1,486	1,486
R-squared		0.003

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

2.5.7. Socially Desirable Reporting

Responses to the behavioral frequency questions are used as indicators of socially desirable reporting. Three questions asked about socially desirable behaviors – the number of days in the last week (0 – 7) that the respondent's child got enough sleep, exercise, and had a family meal. For these questions, reporting a higher number of days the child got enough sleep, for example, could be considered socially desirable than a lower number. Several other items ask about socially undesirable behaviors – the number of hours the respondent's child watched TV, and number of hours spent on the computer in the last week, how often the respondent raises their voice, spanks their child, smokes in the presence of their child or allows others to smoke in the presence of their child. For these items, lower reports, for instance reporting never raising one's voice, could be interpreted as socially desirable.

We begin with the effect of commitment on reducing reports of the socially desirable behaviors. Based on the mean number of days reported in Table 2.19, we see

that commitment group respondents overall compared with the control group reported slightly less for each of the measures. However, as the regression results for the main effect of commitment in Table 20 indicate, these differences are not statistically significant.

Table 2.19. Mean Number of Days Child Got Enough Sleep, Got Exercise, and Had a Family Meal by Commitment Level

	Number of days child got enough sleep		Number of days child got exercise		Number of days had family meal	
	Mean	(se)	Mean	(se)	Mean	(se)
Control	5.94	0.08	4.35	0.11	5.22	0.10
Commitment	5.89	0.04	4.43	0.06	5.07	0.06

Table 2.20. The Effect of Commitment Level on Number of Days Child Got Enough Sleep, Got Exercise, and Had a Family Meal

VARIABLES	Enough sleep	Exercise	Family meal
Commitment	0.04	0.13	-0.18
	(0.096)	(0.130)	(0.129)
Feedback	-0.25**	-0.13	0.11
	(0.095)	(0.128)	(0.127)
Constant	5.94***	4.35***	5.22***
	(0.078)	(0.105)	(0.105)
Observations	1,364	1,363	1,378
R-squared	0.005	0.001	0.002

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Turning next to socially undesirable behaviors, looking at the means reported in Table 2.21, we see that commitment group respondents reported more TV and computer time and smoking in child's presence, which is as expected, but the same or less for raising voice, spanking and allowing others to smoke in the child's presence. As the regression results for the main effect of commitment in Table 2.22 indicate, none of these differences are statistically significant. These results fail to support Hypothesis 5, that commitment will reduce socially desirable reporting and increase socially undesirable reporting. There was no significant main effect for college or higher education level or for the interaction education and commitment.

Table 2.21. Means for Reports of Socially Undesirable Behaviors by Commitment Level

	> 2-3 hours of TV per day		> 2-3 hours of computer time per day		Raises voice at child > never	
	Mean	(se)	Mean	(se)	Mean	(se)
Control	0.19	0.02	0.12	0.02	0.09	0.02
Commitment	0.23	0.01	0.15	0.01	0.09	0.01

	Spanks child > never		Smokes in child's presence > never		Allows others to smoke in child's presence > never	
	Mean	(se)	Mean	(se)	Mean	(se)
Control	0.14	0.02	0.03	0.01	0.05	0.01
Commitment	0.12	0.01	0.04	0.01	0.05	0.01

Table 2.22. The Main Effect of Commitment Level on Reports of Socially Undesirable Behaviors

VARIABLES	> 2-3 hours of TV per day	> 2-3 hours of computer time per day	Raises voice at child > never	Spanks child > never	Smokes in child's presence > never	Allows others to smoke in child's presence > never
Commitment	0.22	0.19	0.10	-0.10	0.14	0.22
	(0.166)	(0.199)	(0.236)	(0.191)	(0.355)	(0.307)
Feedback	0.13	0.17	-0.08	-0.42^	0.05	-0.40
	(0.153)	(0.180)	(0.231)	(0.216)	(0.331)	(0.321)
Constant	-1.47***	-1.98***	-2.37***	-1.78***	-3.35***	-3.05***
	(0.138)	(0.165)	(0.194)	(0.153)	(0.294)	(0.256)
Observations	1,377	1,367	1,352	1,367	1,391	1,386

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

2.5.8. Break-offs and Response Time

This section examines the effect of commitment on the number of break-offs, which is starting but not completing the survey, and overall response time. Based on the descriptive results shown in Table 3.23, we see that overall, commitment group respondents were more likely to break-off than control group respondents. As shown in Table 3.24, the difference is significant indicating that commitment group respondents were significantly more likely to break off than control group respondents, failing to support Hypothesis 6.

However, looking at median response time, we see that commitment group respondents took longer to complete the survey overall. The regression results in Table

3.24 show a highly statistically significant increase in response time for commitment, which supports Hypothesis 7.

Table 2.23. Percentages of Break-offs and Median Response Time

	Break-off		Median Response Time
	n	%	
Control	19	5.2	10.28
Commitment	95	8.5	13.33

Table 2.24. The Effect of Commitment on Break-offs and Median Response Time

VARIABLES	Break-off	Median Response Time
Commitment	0.72*	2.65***
	(0.294)	(0.630)
Feedback	-0.19	1.07^
	(0.251)	(0.617)
CRCs	-0.41	0.23
	(0.272)	(0.633)
Constant	-2.91***	10.28***
	(0.236)	(0.442)
Observations	1,486	1,486

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

2.6. Discussion

Asking respondents to commit to adhering to certain desired response behaviors produced mixed results. On the one hand, commitment had no effect on response accuracy or data quality for some questions, particularly those concerning the number of medical visits. For several questions, no effect of commitment was observed. For example, there was no significant difference in the likelihood of reporting a matching number of medical visits with the number in the records for visits to the pediatrician, specialist, ER or hospital compared to the control group. Commitment group respondents were also no more likely to report a matching date with the date in the records for the most recent visit to a specialist, the ER or hospital or to report a full date for any of the visit types. There was also no significant difference between commitment and the control group for overall item nonresponse, straightlining, or for reducing socially desirable reports or for increasing socially undesirable reports. Perhaps, the respondents for this

study, the parents of child patients who were sufficiently motivated to log on to complete the survey in the first place, was a fairly conscientious or committed population to begin with, offering little opportunity for improvement.

On the other hand, commitment did improve the quality of reports for other questions. For example, commitment group respondents were significantly more likely to report a matching date of the last visit to the pediatrician. And commitment was particularly effective in improving response quality when the response task was most difficult: Commitment group respondents maintained a higher level of accuracy as the number of visits increased and as the time since the date of the most recent visit to a specialist increased. These response tasks are challenging because the more events one tries to recall the more likely it is that at least some events will be forgotten (Means & Loftus, 1991); moreover, recalling specific dates is widely recognized as being particularly arduous (Tourangeau, Rips, & Rasinski, 2000). It seems likely that asking respondents to commit to being as precise as possible and to looking up information in records as needed, in particular, may have made a difference for these difficult response tasks. In addition to increased accuracy when the response task is difficult, commitment group respondents also took significantly more time to complete the interview, and were significantly less likely to acquiesce, although commitment did not reduce straightlining or socially desirable responding.

Other results suggest that there may be a downside to commitment. While commitment group respondents were more likely to report a matching number of visits to a specialist as the number of visits increased, they were also more likely to overreport visits when there were zero visits in the records. It is possible that commitment group respondents may have been more concerned about missing visits and, thus, tended to err on the side of overreporting when they in fact, had no visits to report. Commitment might motivate respondents to report *something* because they feel that to report no events violates or may give the appearance of violating the agreement they entered into to take the task seriously and invest effort. Further, commitment group respondents had a significantly higher proportion of skipped items for medical visit and date questions. Since the commitment statement emphasized the importance of accuracy, it is possible that more committed respondents opted to not answer the medical visit and date questions

rather than give answers they felt might be inaccurate. Item nonresponse, in this context, might actually be preferred, if the respondent lacks confidence in the accuracy of her answer and, thus, might be more likely to be inaccurate or incomplete. Commitment group respondents also broke off at a significantly higher rate than the control group. In the same vein as with item nonresponse to the medical visit questions, this may be because, more commitment group respondents opted not to complete the interview if they felt that they were unable to provide complete and accurate answers, as requested in the commitment statement.

A key difference between the current study and previous studies examining the effect of commitment was offering a list of response behaviors to which respondents could either commit to or not. The thinking behind this was that it would underscore what is meant by commitment and strengthen the effect of the treatment. The assumption was that respondents would overwhelmingly agree to all of the expected behaviors. Previous studies in which very few refused the commitment including Cannell et al.'s studies (Miller & Cannell, 1982; Oksenberg et al., 1977b, 1977a), which were interviewer-administered (face-to-face and telephone) and Conrad, Couper, Tourangeau, and Zhang's (in press) and Vannette's (2016) web-based studies, seemed to support this assumption. However, contrary to this expectation, there was considerable heterogeneity in how respondents responded to the commitment treatment. The extent to which respondents were willing to commit to the requested behaviors, which this chapter did not take into account, may be associated with different response behaviors in the survey. Study 1, discussed in the previous chapter, found that respondents who were invited to commit but chose not to, i.e., "not committed" respondents, provided poorer quality data, suggesting that those who did not commit to all of the requested behaviors in this study may respond less conscientiously than those who did commit to all of the requested behaviors. The next chapter, therefore, explores the effect of different levels of commitment on response accuracy and data quality.

References

Conrad, F. G., Couper, M. P., Tourangeau, R., & Zhang, C. (in press). Reducing speeding in web surveys by providing immediate feedback. *Cited with Permission.*

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Means, B., & Loftus, E. F. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology*, *5*(4), 297–318.

Miller, P. V., & Cannell, C. F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, *46*(2), 250–269.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977a). Effects of commitment to being a good respondent on interview performance. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 74–108). Department of Health, Education and Welfare.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977b). The effects of instructions, commitment and feedback on reporting in personal interviews. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 133–199). Department of Health, Education and Welfare.

Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.

Vannette, D. L. (2016). The Effects of Asking Attention Filter Questions: Evidence from a 14 Country Survey Experiment. Presented at the Annual Conference of the American Association for Public Opinion Research, Austin, TX.

Vinokur, A., Oksenberg, L., & Cannell, C. F. (1977). Effects of feedback and reinforcement on the report of health information. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 109–132). Department of Health, Education and Welfare.

Appendix 2.1. Parent Study Questionnaire

Introductory items

1.) How long has anyone in your family received care at the University of Michigan?

Less than one year
between 1 and 2 years
between 2 and 3 years
between 3 and 4 years
More than 4 years
DON'T KNOW

The following questions will ask you about the health of [CHILD] and health events and services for [CHILD].

2.) In general, how would you describe [CHILD]'s health? Would you say [his/her] health is excellent, very good, good, fair, or poor?
EXCELLENT
VERY GOOD
GOOD
FAIR
POOR
DON'T KNOW

Health care utilization

3.) During the past 12 months, how many times did [CHILD] see a primary care doctor or nurse practitioner at [his/her] University of Michigan pediatrician's office?

_____ (Please enter a number such as 0,1, 2, etc.)

4.) What was the date of [CHILD]'s most recent visit to [his/her] University of Michigan pediatrician's office?
__ __ Month __ __ Day __ __Year

5.) Specialists are doctors like surgeons, heart doctors, allergy doctors, skin doctors, psychologists, psychiatrists, or eye professionals who specialize in one area of health care.
During the past 12 months, how many times did [CHILD] see a specialist (doctor or nurse practitioner) at the University of Michigan?

_____ (Please enter a number such as 0,1, 2, etc.)

6.) What was the date of the most recent time [CHILD] visited a specialist (doctor or nurse practitioner) at the University of Michigan?
__ __ Month __ __ Day __ __Year

7.) During the last 12 months, how many times has [CHILD] been seen in the Emergency Room (ER) at the University of Michigan (please include visits that resulted in a hospital admission)?

_____ (Please enter a number such as 0,1, 2, etc.)
8.) What was the date of [CHILD]'s most recent Emergency Room (ER) visit at the University of Michigan?
__ __ Month __ __ Day __ __Year

9.) During the last 12 months, how many times has [CHILD] been admitted to CS Mott Children's Hospital?

_____ (Please enter a number such as 0,1, 2, etc.)

10.) What was the date of the beginning of [CHILD]'s most recent stay at CS Mott Children's Hospital?
__ __ Month __ __ Day __ __Year

Experience with Care
Thinking about your child's most recent visit to the pediatrician's office at the University of Michigan, please tell us how much you agree or disagree with each of the following statements:

11.) The provider explained things in a way that was easy to understand.

12.) The provider listened carefully to me.

13.) The provider knew important information about [CHILD]'s medical history.

14.) The provider showed respect for what I had to say.

15.) The provider spent too little time with [CHILD].

16.) I would recommend this provider to my friends and family.

Health Practices
17.) What types of foods, vitamins and/or supplements do you give [CHILD] to stay healthy?

18.) What have you cut down on giving [CHILD] because you think that it is bad for his/her health?

19.) What would you like to do to maintain or improve [CHILD]'s health?

20.) During the past week, on how many nights did [CHILD] get enough sleep for a child [his/her] age?

Zero
One
Two
Three
Four
Five
Six
Seven
DON'T KNOW

21.) During the past week, on how many days did [CHILD] exercise, play a sport, or participate in physical activity outside of school?

Zero
One
Two
Three
Four
Five
Six
Seven
NOT APPLICABLE
DON'T KNOW

22.) In the past week, how much time did [CHILD] spend in front of a TV watching TV programs, movies, or playing video games?

None
Less than 1 hour
1 to 2 hours
2 to 3 hours
3 to 4 hours
4 to 5 hours
More than 5 hours

NOT APPLICABLE
DON'T KNOW

23.) In the past week, how much time did [CHILD] spend in front of a computer watching TV programs or movies, surfing the Internet, or playing video games?

None
Less than 1 hour
1 to 2 hours
2 to 3 hours

3 to 4 hours
4 to 5 hours
More than 5 hours

NOT APPLICABLE
DON'T KNOW

24.) During the past week, how many days did all the family members who live in the household eat AT LEAST ONE meal together?

Zero
One
Two
Three
Four
Five
Six
Seven
DON'T KNOW

25.) In the past three months, how often did [CHILD] attend a religious service?

____ NUMBER
NOT APPLICABLE - DOES NOT ATTEND
DON'T KNOW

26.) How often do you smoke cigarettes in the presence of your child?
Never
Rarely
Sometimes
Often
Most of the time
All of the time

27.) How often do others in your household smoke cigarettes in the presence of your child?
Never
Rarely
Sometimes
Often
Most of the time
All of the time

28.) How often do you raise your voice or yell at your child?

Daily

Almost daily
Two or three times a week
About once a week
Less often than once a week
Never

29.) How often do you spank your child?
Daily
Almost daily
Two or three times a week
About once a week
Less often than once a week
Never

Demographics
30.) What is your relationship to [CHILD]?

MOTHER (BIOLOGICAL, STEP, FOSTER, ADOPTIVE)
FATHER (BIOLOGICAL, STEP, FOSTER, ADOPTIVE)
GRANDMOTHER
GRANDFATHER
AUNT
UNCLE
FEMALE GUARDIAN
MALE GUARDIAN
OTHER ----- please specify _____

31.) In what year were you born? ____

32.) What is your marital status?

Married
Divorced
Widowed
Separated
Never married
A member of an unmarried couple

33.) What is the highest grade or level of school that you completed?

8th GRADE OR LESS
9th-12th GRADE NO DIPLOMA
HIGH SCHOOL GRADUATE OR GED COMPLETED
COMPLETED A VOCATIONAL, TRADE, OR BUSINESS SCHOOL PROGRAM
SOME COLLEGE CREDIT BUT NO DEGREE
ASSOCIATE DEGREE (AA, AS)

BACHELOR'S DEGREE (BA, BS, AB)
MASTER'S DEGREE (MA, MS, MSW, MBA)
DOCTORATE (PhD, EdD) or PROFESSIONAL DEGREE (MD, DDS, DVM, JD)
DON'T KNOW

34.) What is your total household income?

Less than $10,000
$10,000 to $19,999
$20,000 to $29,999
$30,000 to $39,999
$40,000 to $49,999
$50,000 to $59,999
$60,000 to $69,999
$70,000 to $79,999
$80,000 to $89,999
$90,000 to $99,999
$100,000 to $149,999
$150,000 or more

Debriefing

35.) How important do you think the results are from this survey?
Not at all important
A little important
Somewhat important
Very important
36.) How confident are you about the accuracy of your responses to the questions about your child's number of visits to the pediatrician's office, specialists, the ER, and the hospital?

Not confident at all
Somewhat confident
Fairly confident
Extremely confident

37.) How confident are you about the accuracy of your responses to the questions about the dates of your child's most recent visits to the pediatrician's office, specialists, the ER, and the hospital?

Not confident at all
Somewhat confident
Fairly confident
Extremely confident

[FOR RESPONDENTS IN THE TREATMENT GROUPS ONLY]

38.) At the beginning of the survey, we asked you to check boxes indicating your commitment to several approaches to completing the survey to insure responses that are as accurate, complete, and honest as possible.

In the space below, please tell us as much as you remember about what we asked you to commit to:

[FOR RESPONDENTS IN THE CRC GROUP ONLY]
39.) We presented the following information along with questions about your child's number of visits to the pediatrician's office, specialists, the ER, and the hospital:

"TO HELP YOU REMEMBER: It may help to think about the reason for the visits, the season of the year, who took [CHILD] to the visits, if you took off work, how you travelled to the visits, who else was with you and [CHILD], and if you waited long."

How helpful did you find this information when coming up with your answers to these questions?

Very helpful
Somewhat helpful
A little helpful
Not helpful at all
I didn't read this information

40.) Did you check a calendar or other relevant documents to answer one or more of the questions?

Yes – Skip to 40a
No – Skip to 40b

> 40a.) Under which circumstances did you check a calendar or relevant documents?
>
> I checked a calendar or relevant documents to make sure that I gave the correct answers
> I looked at a calendar or relevant documents for the correct answer when I was not sure of my answers
>
> 40b.) Which of the following best describes why you did not check a calendar or relevant documents?
>
> I was sure that I knew the correct answer to all the questions.
> I thought my estimates would be close enough.

41.) How much effort would you say you put into this survey?
Not much effort at all
A little effort
A moderate amount of effort
A fair amount of effort
A great deal of effort

42.) How burdensome was this survey to you?
Very burdensome
Somewhat burdensome
Not at all burdensome
A little burdensome

43.) Do you have any comments about our survey?

Appendix 2.2. Distributions of Medical Visits

Figure 2.8. Visits to a Specialist – Visits in the Records

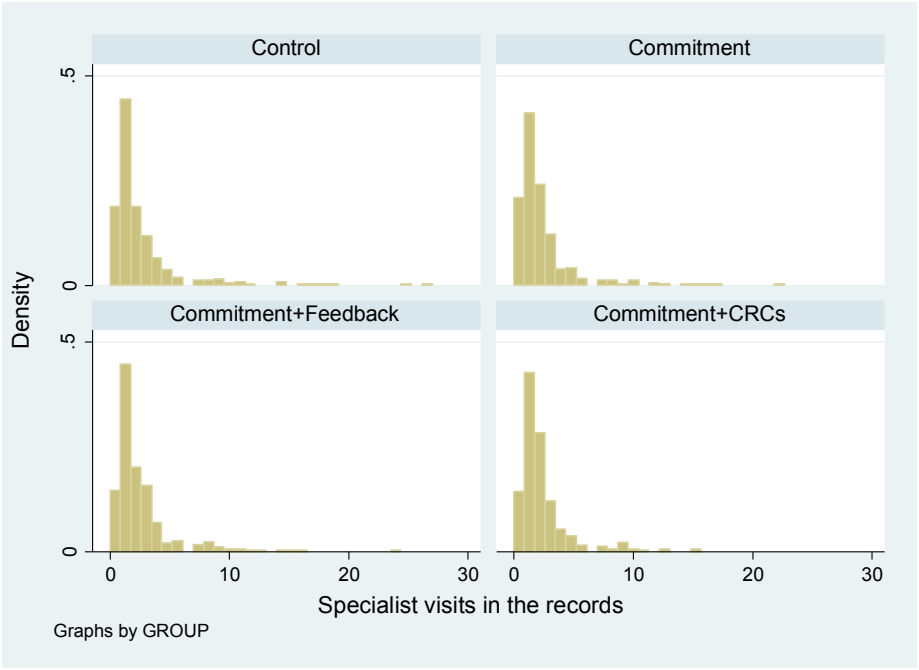


Figure 2.9. Visits to a Specialist – Difference Between Reported Visits and Visits in the Records

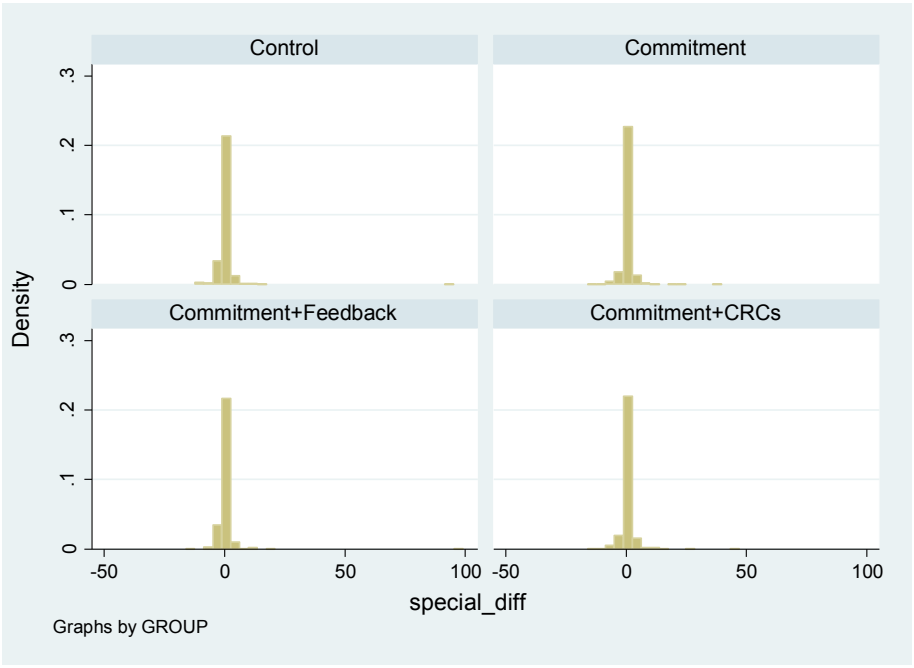


Figure 2.10. Visits to a Specialist – Absolute Difference Between Reported Visits and Visits in the Records

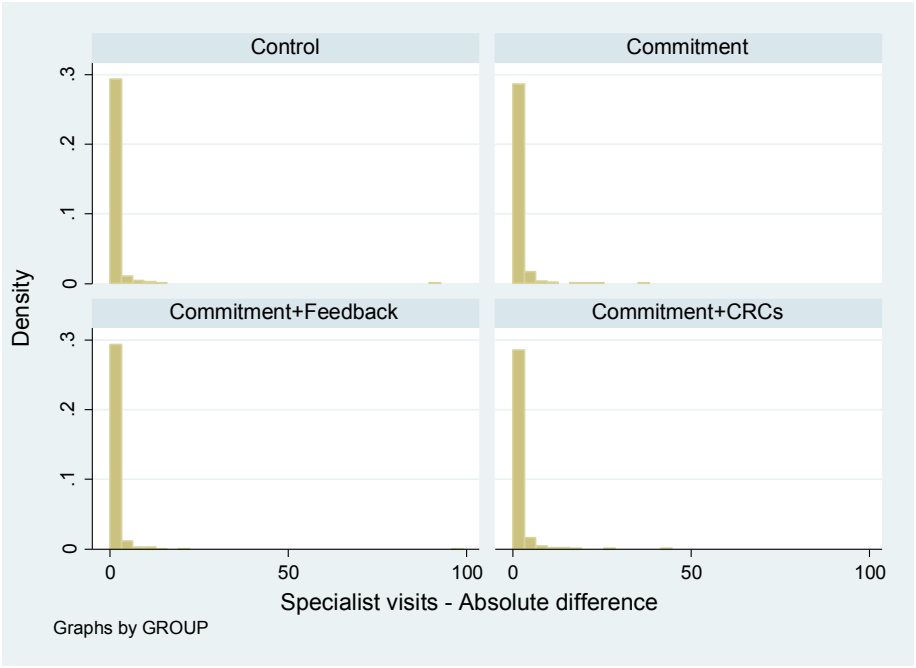


Figure 2.11. Visits to the ER – Visits in the Records

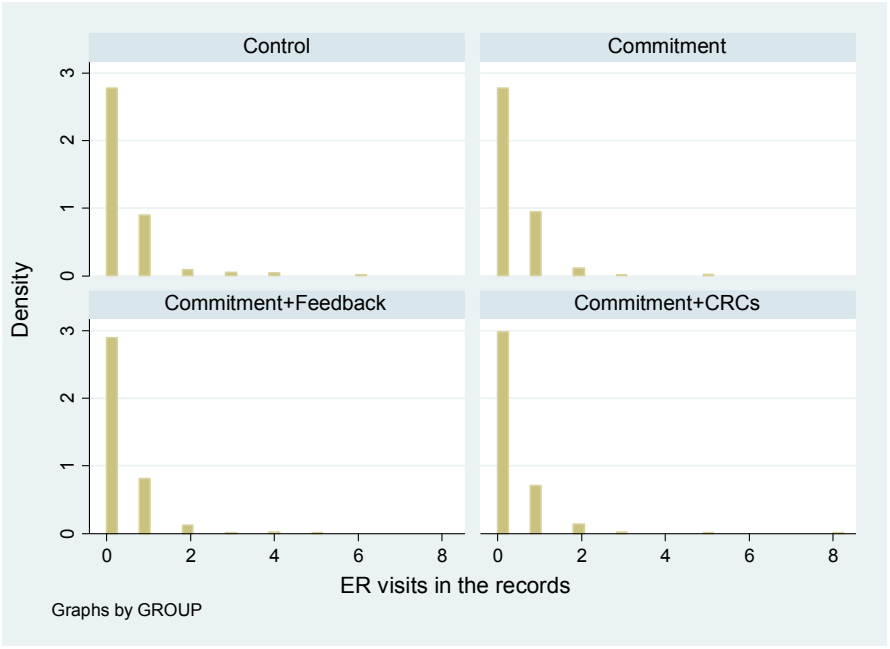


Figure 2.12. Visits to the ER – Difference Between Reported and Visits in the Records

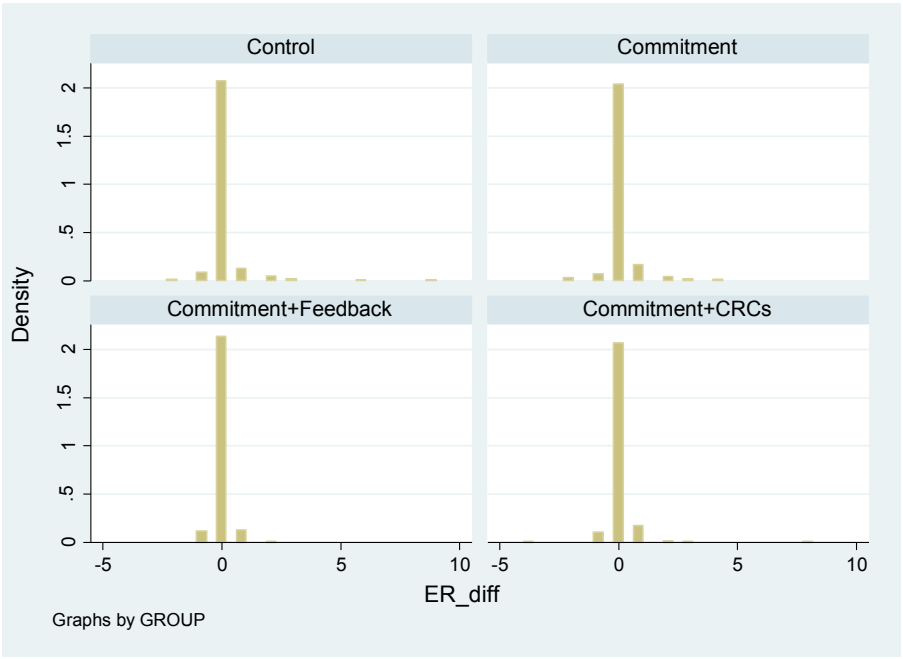


Figure 2.13. Visits to the ER – Absolute Difference Between Reported and Visits in the Records

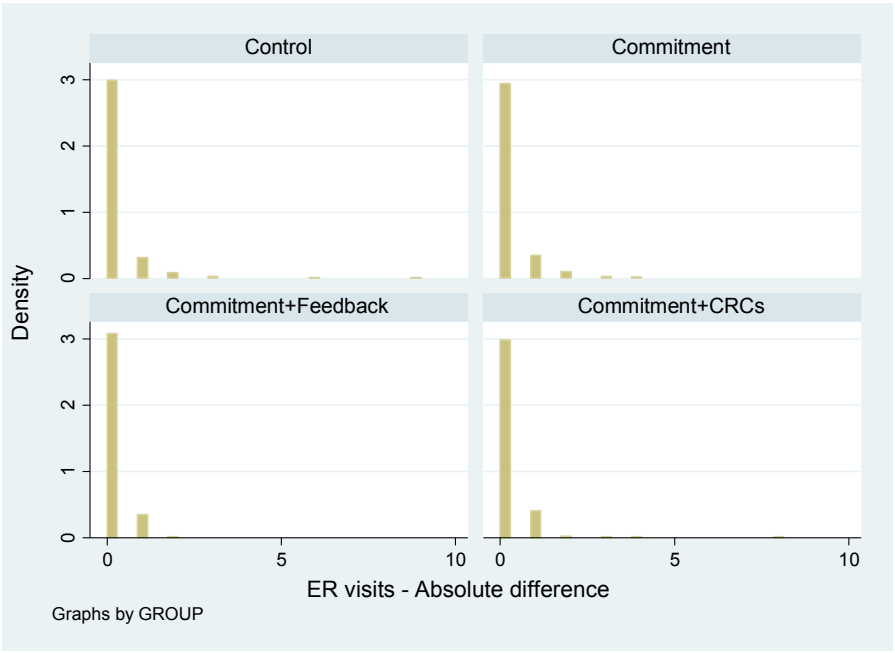


Figure 2.14. Visits to the Hospital – Visits in the Records

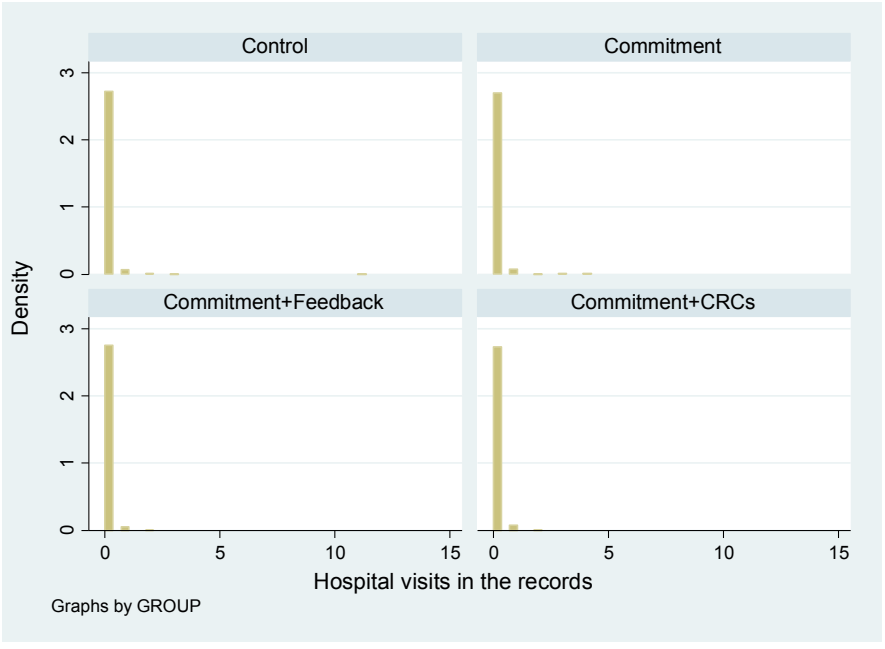


Figure 2.15. Hospital Visits – Absolute Difference Between Reported and Visits in the Records

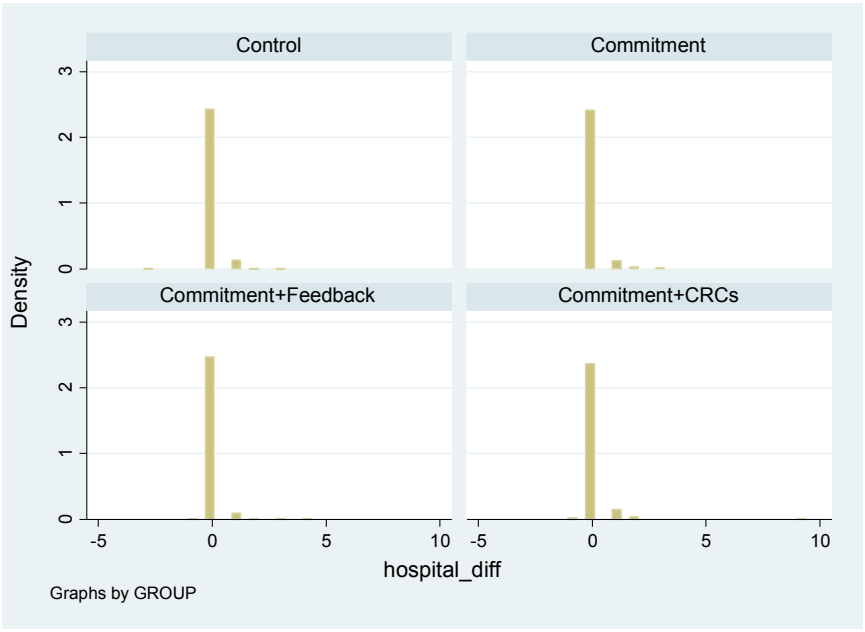
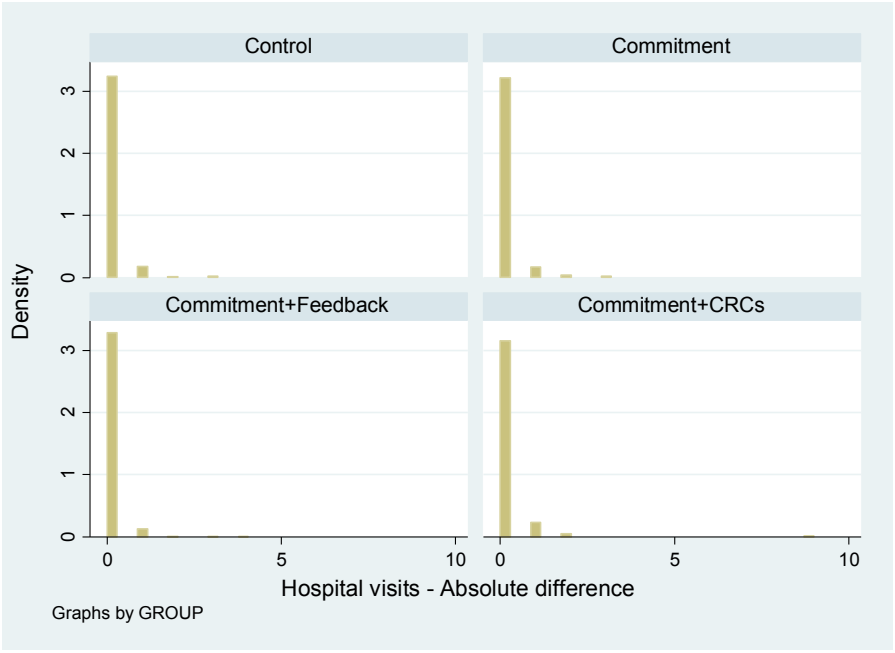


Figure 2.16. Hospital Visits – Absolute Difference Between Reported and Visits in the Records



Appendix 2.3. Distributions for the Date of the Last Visit

Figure 2.17. Date of Last Visit to the Pediatrician – Absolute difference Between the Reported Date and the Date in the Records

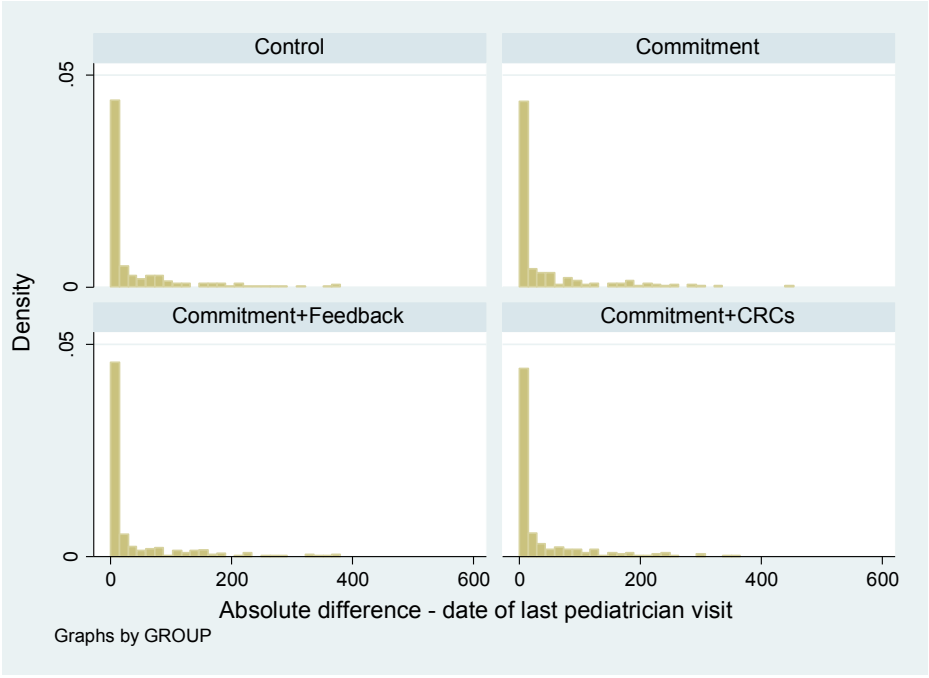


Figure 2.18. Date of Last Visit to a Specialist – Absolute difference Between the Reported Date and the Date in the Records

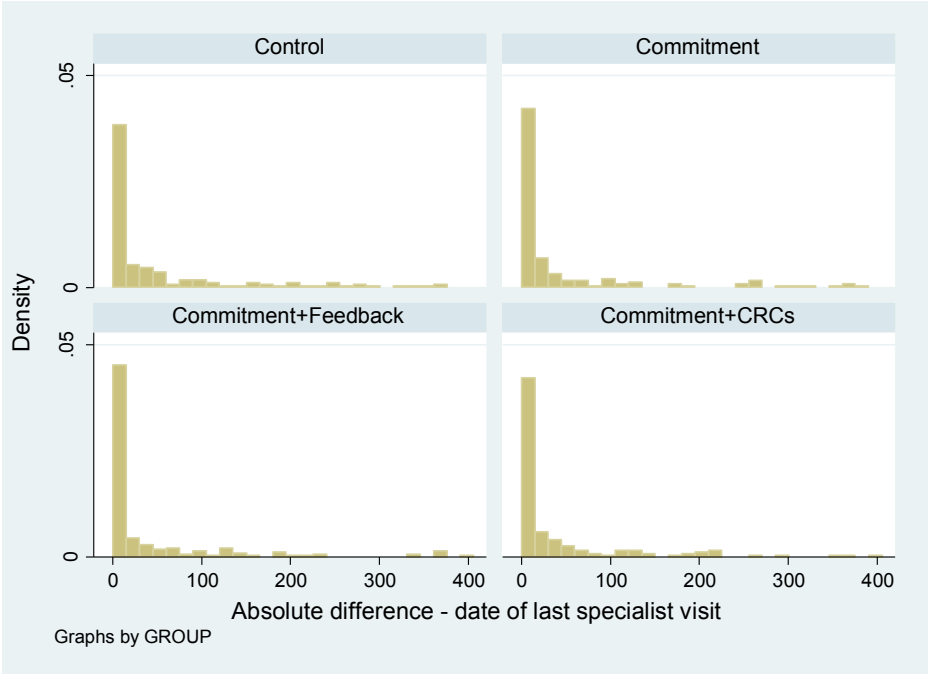
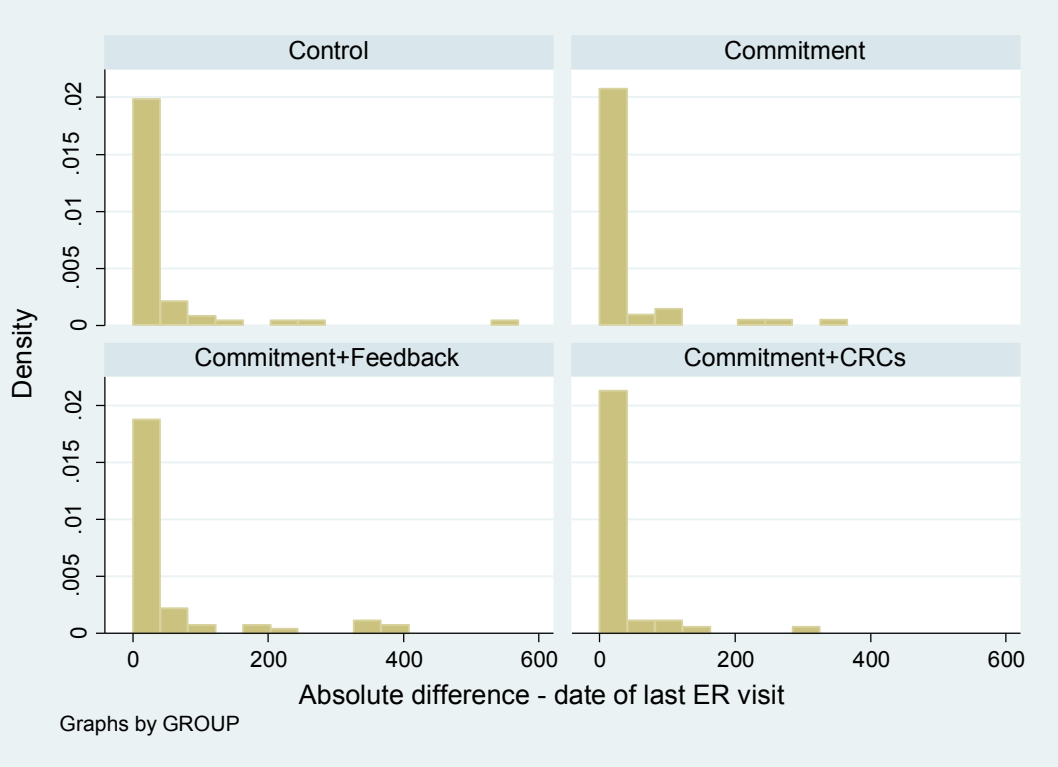


Figure 2.19. Date of Last Visit to the ER – Absolute difference Between the Reported Date and the Date in the Records



Chapter 3: The Effect of Level of Commitment in a Survey of Parents about their Child's Health and Healthcare (Study 2)

3.1.Introduction

Cannell and his associates (Cannell, Miller, & Oksenberg, 1981; Miller & Cannell, 1982; Oksenberg, Vinokur, & Cannell, 1977b, 1977a) at the University of Michigan proposed asking respondents to commit to providing complete and accurate responses in a global, binary way – respondents could choose either to commit or not to commit. The commitment request in the current study was actually comprised of asking respondents to commit to engaging in five specific response behaviors that seem likely promote data quality, such as reading the questions carefully, and trying to be as precise as possible. Respondents could commit to some but not all of these practices so commitment as implemented here was continuous or graded in contrast to the binary commitment requested by Cannell and his associates in their pioneering research. This chapter concerns the effects of *commitment level*, i.e., how many and which of these particular commitment requests respondents pledged to carry out, extending the findings reported in the previous chapter concerning the overall effects of commitment. Those results demonstrated that any commitment versus none sometimes improves data quality but in other circumstances it does not. The present chapter explores how the level of commitment might affect response accuracy and other indirect indicators of data quality. Most of the comparisons are within treatment, i.e., between different levels of commitment; however, some comparisons are also made to the control group.

3.2. Study Design and Methods

As mentioned in the previous chapter, respondents receiving the commitment treatment were presented with a series of checkboxes corresponding to the following expected response behaviors:

- Reading all of the questions carefully
- Trying to be as precise as possible
- Looking up information in records or an a calendar, if needed
- Providing as much information as possible
- Answering honestly

Respondents could select one or more of the checkboxes or none of the above. Respondents who selected none of the above were still able to proceed with the survey. Please see the previous chapter for additional details on the study design and methods.

3.3. Analytical Methods

As discussed in the previous chapter, respondents responded in several different ways to the request for commitment. As shown in Table 3.1, three levels of commitment can be identified, and are used in the analysis reported below.

Table 3.1. Response Behaviors Selected and Corresponding Levels of Commitment

Number of Response Behaviors Selected	Commitment Level	n	%
All five checkboxes	"Fully Committed"	700	63%
4 checkboxes	"Moderately Committed"	302	27%
3 checkboxes	"Least Committed"	43	4%
2 checkboxes		7	.6%
1 checkbox		56	5%
None		6	.5%
Total		1,114	100%

As shown in Table 3.1, 63% of respondents checked all of the checkboxes and could be considered "fully committed". Twenty-seven percent of respondents selected four out of five checkboxes. Of the respondents who selected four checkboxes, 295 out of 302 checked all of the checkboxes except for the one for "looking up information in records or on a calendar, if needed". Because they committed to all of the requested

response behaviors, these respondents could be considered "moderately committed". The remaining respondents selected some combination of two or three checkboxes, only one checkbox, or "none of the above". Among those who selected three checkboxes, 56% agreed to read all of the questions carefully, try to be as precise as possible with answers, and answer honestly, and did not agree to look up information or provide as much information as possible. Another 19% of the respondents selecting three checkboxes agreed to read all of the questions carefully, answer honestly, and provide as much information as possible, excluding looking up information and trying to be as precise as possible. All seven of the respondents selecting two checkboxes agreed to either read all of the questions carefully or answer honestly. None of these respondents agreed to look up information. Among the respondents selecting only one checkbox, 46% agreed only to read all of the questions carefully and 27% agreed only to answer honestly. In contrast to the respondents who agreed to all of the requested response behaviors (the "fully committed") and those who selected all of the requested behaviors with the only, near universal exception of looking up information (the "moderately committed"), the respondents who selected only three, two, one or none of the requested response behaviors could be considered the "least committed".

Table 3.2 shows regression results examining predictors of fully committed, compared to the least committed, including the child's health status (very good or excellent health compared to good, fair, or poor), total number of visits of any type in the medical records (visits to the pediatrician, a specialist, ER, and hospital), child's age, parent's age, the relationship of the respondent to the child, and parent education. Respondents with a college degree or higher education were significantly more likely to fully commit while respondents with an older child and other than the child's mother were less likely to fully commit. Fully committed is also associated with having marginally more total visits in the records.

Table 3.2. Predictors of Fully Committing

VARIABLES	Predictors of Fully Committing
Child health status	0.11
	(0.383)
Total visits in records	0.03^

	(0.020)
Child age	-0.06*
	(0.027)
Parent age	-0.01
	(0.012)
Other than child's mother	-0.38*
	(0.185)
College degree or higher	0.56***
	(0.145)
Constant	1.00*
	(0.435)
Observations	1,015

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

It is important to note that the results of the analysis reported below are observational as opposed to the result of the experimental design. It is not possible to determine whether the response behavior for different commitment levels simply classifies respondents' degree of compliance with the commitment request or whether the act of committing to perform in a particular way changed respondents' behavior accordingly. For example, we do not know if the least committed respondents, who were less accurate than the fully committed respondents would have been similarly accurate with a binary ("yes"-"no") commitment request, or if it is because having been asked to commit and not committing fully affected their behavioral in ways that helped them justify in not engaging in all of the behaviors listed in the commitment request.

The analysis reported below draws on the same outcome measures that were used in the previous chapter examining the main effect of commitment, as follows:

Accuracy measures: Accuracy is assessed by comparing the reported number of visits as well as the dates of the most recent visit to the pediatrician, a specialist, the ER and the hospital to the information in the child's medical records in several different ways, as detailed below.

Indirect measures: The proposed study also examines the effect of commitment on several indirect measures of data quality. These measures include item nonresponse, the number of mentions and length of responses to open-ended questions, straightlining, acquiescence, socially desirable reporting, break-offs and response time. Please see Chapter 1 for further detail and rationale on these outcome measures.

To examine differences between the fully committed, moderately committed and least committed for the outcome measures discussed in this chapter, regression models are fit with the following indicator variables: moderately committed (0 or 1); and fully committed (0 or 1) with the reference group (0) mapping to the least committed group, as shown in an example for linear regression shown in Equation 3.1. Binary indicators were also entered for the feedback and CRC (0=control/reference; 1=treatment) conditions to control for possible effects of these treatments in the presence of commitment. Differences tend to be greatest between the fully committed and least committed for most of the outcome measures, with the fully committed providing more accurate and higher quality responses than the least committed. However, it is clear from the descriptive results for some measures, such as for the date of the most recent visit to the pediatrician and specialist, that the fully committed responded more accurately than the least committed as well as the control group, making the comparison to the control group pertinent. To compare the fully, moderately, and least committed to the control group, models are fit with the control group as the reference group.

Equation 3.1. Regression Equation for Estimating the Effect of Commitment Level Groups

$$\gamma = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \varepsilon$$

γ = the outcome measure
β_0 = intercept
β_1 = Fully committed
β_2 = Moderately committed
β_3 = Feedback
β_4 = CRCs

Respondent education level (college or graduate degree versus no college) is entered subsequently into all models and examined for possible interactions with commitment. As discussed in Chapter 2, this is because there is reason to believe that the effect of commitment may vary by respondent education level and past studies have found some differential effects for commitment by education. Commitment could be more effective for low education respondents if higher education respondents have a higher level of motivation, regardless of the treatment, due to increased "need for cognition", which

Krosnick (1991) has identified as a possible determinant of respondent motivation. One study by Oksenberg et al. (1977) (in the analysis presented in Cannell et al. (1981)) found that commitment increased reporting to open-ended questions among low education respondents. Other the other hand, higher education respondents may have more cognitive resources to draw on than low education respondents for certain response tasks. Conrad et al.'s (in press) finding that higher education respondents answered numeracy questions more accurately when they had made a commitment (and were prompted for answering too quickly), whereas lower education respondents did not improve with commitment, supports this theory. However, other studies have found no differential effects of commitment by education level (Miller & Cannell, 1982; Oksenberg et al., 1977a; Vinokur et al., 1977).

Several measures are used to assess the accuracy and direction of error of reported visits by comparing the reported values to the values recorded in the medical records as follows:

- A **binary metric** for a match between the reported number of visits and the number of visits in the records versus no match
- The **simple (signed) difference** between the reported number of visits and the number of visits in the records
- The **absolute difference** between the reported number of visits and the number of visits in the records
- A **ratio metric** of the count of reported visits relative to reported visits for the reference group
- A **tri-variate measure** – 0 for a match between the reported number of visits and the number of visits in the records, 1 if the difference between reported visits and the records is greater than zero (i.e. overreporting), and 2 if the difference is less than zero (i.e. underreporting)

As in the analysis assessing the main effect of commitment in Chapter 2, for models assessing the accuracy of reported visits, in addition to education, the number of visits in the records, and whether the respondent reported checking records are entered into the models along with interaction terms with commitment. The main effect of the number of visits in the records controls for the difficulty of the task, the idea being that

the more visits the respondent's child has had in the last 12-months, the more difficult it is to recall the exact number of visits accurately. In the context of the binary metric for a match versus no match with the records, a positive and significant interaction with commitment would suggest that, as the number of visits increased, and the response task presumably became more difficult, committed respondents were more likely to report a matching number of visits, in other words, maintained a higher level of accuracy. Whether the respondent reported checking a calendar or other relevant documents (checking records) is also entered as a main effect; presumably records are more accurate than recall or other estimation processes, although records may contain error as well. Interaction terms for checking records and commitment are also entered to see if the effect of checking records differed between commitment and control group respondents. For example, checking records could have more of an effect for commitment group respondents if they checked multiple types of records or checked their records more thoroughly.

Three measures are used to examine the accuracy and quality of the reported dates:

- A **binary metric** for a match between the reported date and date in the records versus no match (using logistic regression)
- The **absolute difference** between the reported date and the date in the records (using linear regression)
- A binary metric for whether or not the respondent reported a full date (including a day, month, and year) (using logistic regression) is also examined.

As in the analysis assessing the main effect of commitment presented in Chapter 2, for models assessing the accuracy of reported dates, in addition to education, the number of days since the date of the last visit in the records, and whether the respondent reported checking records are entered into the models along with interaction terms with commitment. The main effect of the number of days since the date of the last visit in the records controls for the difficulty of the task, the idea being that as the number of days since the last visit increases, the more difficult it is to recall the date accurately. In the context of the binary metric for a match versus no match with the records, if commitment increases respondent motivation and effort to provide accurate answers, a positive and

significant interaction would suggest that, as the number of days since the last visit increased, and the response task presumably became more difficult, committed respondents were more likely to report a matching date.

All analyses were conducted using Stata (*Stata Statistical Software*, 2015).

3.4. Results

3.4.1. Item Nonresponse

This section examines the effect of commitment level group on item nonresponse overall and specifically, to questions about medical visits and dates. As noted in the previous chapter, respondents were asked up to forty-two questions but not all respondents were asked the same number of questions. Therefore, the proportion of item nonresponse overall was examined in two ways: 1) based on the number of questions not answered out of the total number of questions asked, and 2) based on the total number of questions asked of all respondents (excluding the treatment-related debriefing items and commitment statement). Results were nearly identical so only one set of results, those based on the total number of questions asked (including treatment-related items), is displayed.

For four types of visits, respondents were asked to report the number of visits in the last 12 months and the date of the most recent visit: visits to the pediatrician's office, a specialist, the ER, and to the hospital. All respondents were asked each of the four medical visit questions but only those respondents who reported at least one of a given type of visit was asked to report the date of the most recent visit of that type. Respondents were therefore asked at least four and up to eight of the medical visit questions. As with overall item nonresponse, differences in the number of questions asked were accounted for when calculating the proportion of questions not answered for each respondent.

As shown in Table 3.3, the mean proportion of item nonresponse overall and to the medical visit and date questions was highest among the least committed respondents. For item nonresponse overall, the mean proportion is the same between fully committed and moderately committed respondents. But for the medical visits and date questions, the mean proportion of item nonresponse for moderately committed respondents was higher

than that of the fully committed respondents, i.e., item nonresponse increased monotonically from the most to least committed respondents.

Table 3.3. Proportion of Item Nonresponse Overall and for Medical Visit and Date Questions

	Overall		Medical Visit and Date Questions	
	p	(se)	p	(se)
Control	0.16	0.008	0.068	0.009
Commitment	0.17	0.006	0.084	0.007
Least committed	0.20	0.007	0.121	0.009
Moderately committed	0.16	0.010	0.096	0.012
Fully committed	0.16	0.022	0.066	0.025

As shown in the regression results shown in Table 3.4, comparing the fully and moderately committed to the least committed, fully committed respondents skipped significantly fewer items overall ($p < .05$) and of the medical visit and date questions ($p < .01$) specifically. Moderately committed respondents skipped marginally fewer questions overall than the least committed ($p < .1$). The R-squared values for these models are low, suggesting that commitment levels may not explain much of the variation in item nonresponse. Nonetheless, these regression results further demonstrate the trend observed in the descriptive results. A possible main effect of education level of college degree or higher and interaction with the commitment level groups was examined but was not significant.

Table 3.4. The Effect of Commitment Level Groups on Item Nonresponse Overall and to Medical Visit and Date Questions

	a. Reference: Control		b. Reference: Least Committed	
VARIABLES	Overall	Medical Visit and Date Questions	Overall	Medical Visit and Date Questions
Fully committed	0.02	0.02	-0.04*	-0.06**
	(0.014)	(0.016)	(0.018)	(0.021)
Moderately committed	0.02	0.05**	-0.04^	-0.03
	(0.016)	(0.018)	(0.020)	(0.023)
Least committed	0.06**	0.08**		
	(0.021)	(0.024)		
Feedback	-0.04**	-0.04**	-0.04**	-0.04^
	(0.013)	(0.015)	(0.013)	(0.015)
CRCs	-0.02	-0.03*	-0.02	-0.03*

	(0.014)	(0.015)	(0.014)	(0.015)
Constant	0.16***	0.07***	0.22***	0.15***
	(0.009)	(0.011)	(0.019)	(0.022)
Observations	1,481	1,481	1,481	1,481
R-squared	0.009	0.012	0.009	0.012

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

3.4.2. Response Accuracy: Medical Visit Questions

In the previous section, we saw that fully committed respondents were significantly more likely to respond to the medical visit and date questions than the least committed. In this section, we explore the extent to which commitment level groups were more or less accurate in their responses to the medical visit questions, when they did respond.

Looking at the frequency of the number of visits in the records for visits to the pediatrician, specialist, ER and hospital by commitment level and the control group in Table 3.5, we see that the number of visits in the records is reasonably balanced across the groups. This suggests that there is somewhat of a baseline in the difficulty of the response task for the each of the medical visit questions and that the respondent's level of commitment appears to be independent of the number of their child's medical visits.

Table 3.5. Visits in the Records by Commitment Level and Control Group

	Fully Committed		Moderately Committed		Least Committed		Control			
	n	%	n	%	n	%	n	%		
	700	42.3	302	20.4	112	7.6	367	24.7		
Pediatrician visits									p-value of χ^2	df
1	232	33.9	112	38.4	35	32.7	123	34.2		
2 – 4	347	50.7	152	52.1	59	55.1	186	51.7	0.302	6
5 or more	106	15.5	28	9.6	13	12.15	51	14.2		
Specialist visits										
0	99	14.1	45	14.9	17	15.2	60	16.4		
1	262	37.4	107	35.4	47	42.0	142	38.7	0.558	9
2 – 3	231	33.0	103	34.1	30	26.8	98	26.7		
4 or more	108	15.4	47	15.6	18	16.1	67	18.3		
ER visits										
0	513	73.3	225	74.5	90	80.4	263	71.7		
1	158	22.6	62	20.5	16	14.3	85	23.2	0.500	6
2 or more	29	4.1	15	5.0	6	5.4	19	5.2		
Hospital visits										

0	677	96.7	294	97.4	106	94.6	354	96.5	0.597	3
1 or more	23	3.3	8	2.7	6	5.4	13	3.5		

If a higher level of commitment helped improve response accuracy to the visit questions, we would expect better performance for both match and difference measures for the fully and moderately than least committed group: higher percentages of responses that match the recorded response and lower absolute differences between the reported number of visits and the visits in the records. As shown in the descriptive results in Table 3.6, this is largely the case. For the percentage of answers matching the records and for the absolute difference, results are in the expected direction for visits to the pediatrician, a specialist, and to the ER. In some cases, such as visits to the specialist and visits to the hospital, the moderately committed had a slightly higher percentage of a match or smaller absolute difference than the most committed. For visits the hospital, results are in the expected direction for the moderately committed but not for the fully committed respondents, who had a slightly smaller percentage of a match and larger absolute difference than both the moderately committed and least committed. However, responses here were quite accurate overall.

Table 3.6. Percentage of Matches and Mean Absolute Differences between Reported Visits and Records for Commitment Treatment versus Control and by Level of Commitment

	Visits to the Pediatrician				Visits to a Specialist			
	Match		Absolute difference		Match		Absolute difference	
	%	(se)	Mean	(se)	%	(se)	Mean	(se)
Control	0.36	0.03	1.17	0.09	0.44	0.03	1.42	0.28
Commitment	0.37	0.01	1.22	0.06	0.43	0.02	1.41	0.13
Least committed	0.29	0.04	1.26	0.15	0.31	0.05	1.44	0.20
Moderately committed	0.34	0.03	1.21	0.11	0.46	0.03	1.38	0.36
Fully committed	0.40	0.02	1.22	0.07	0.44	0.02	1.42	0.13

	Visits to the ER				Visits to the Hospital			
	Match		Absolute difference		Match		Absolute difference	
	%	(se)	Mean	(se)	%	(se)	Mean	(se)
Control	0.87	0.02	0.21	0.04	0.94	0.01	0.07	0.02
Commitment	0.87	0.01	0.16	0.02	0.93	0.01	0.09	0.01

Least committed	0.86	0.03	0.20	0.06	0.93	0.03	0.07	0.03
Moderately committed	0.86	0.02	0.15	0.03	0.97	0.01	0.03	0.01
Fully committed	0.88	0.01	0.16	0.02	0.92	0.01	0.12	0.02

The regression analyses reported below focuses on a binary metric (match versus no match) using logistic regression, a tri-variate metric (0=match, 1=overreport, 2=underreport) using multinomial logistic regression, and a ratio metric (count of reported visits relative to reported visits for the reference group) using Poisson regression. The binary metric gives us the most direct measure of accuracy while the tri-variate measure indicates whether more respondents overestimated in the commitment than control conditions and whether more underestimated in the commitment than control conditions. The ratio metric gives us an aggregate sense of the relative rate of reporting – a higher versus lower count of reported visits, on average, relative to the reference group. The ratio metric may detect broad differences in response behavior that may be associated with more or less accuracy but it needs to be interpreted in the context of the other measures. For example, a significant positive difference in the ratio metric could indicate 1) overreporting compared to the reference group or, alternatively, 2) it could mean more accurate reporting of visits and that the reference group is underreporting its true visits.[4] Evidence of overreporting in the tri-variate metric and a lower rate of matches (e.g. less accuracy) would support the first explanation while evidence of less underreporting and a higher rate of matches would support the second explanation.

As expected, based on the descriptive results, regression results comparing fully and moderately committed to the least committed respondents, shown in Table 3.7, show that fully committed respondents were significantly more likely to report the number of visits matching the number in the records for both visits to the pediatrician and to a specialist ($p < .05$). This result is also significant for moderately committed respondents for visits to a specialist. Results for the Poisson models indicate that fully and moderately committed respondents reported a significantly higher rate of visits to the pediatrician and

[4] A third explanation could be that the group of interest has more actual visits (i.e. in the records) than the reference group. However, chi-square tests have ruled out any significant differences between the control and commitment level groups in the number of visits in the records, as shown in Table 3.5.

to a specialist relative to the least committed respondents. In the results for the tri-variate measure – (0=match, 1=overreport, 2=underreport) – using multinomial regression, shown in Table 3.8, we see that the fully committed were less likely to underreport, for visits to the pediatrician and to a specialist ($p < .001$). Moderately committed respondents were also less likely to underreport ($p < .01$) for visits to a specialist. These results support the idea that fully and moderately committed respondents were more accurate (more likely to report a match) for visits to the pediatrician and to a specialist because they tended to report more visits, as seen in the results for the ratio measure.

Looking at the results for ER visits, there were no significant differences between the fully and moderately committed and the least committed in the likelihood of a match between reported visits and the medical records. The marginally significant result for the ratio measure for fully committed respondents without a corresponding increase in matches, which could indicate some overreporting. However, there is no evidence of overreporting in the multinomial results in Table 3.8b. For visits to the hospital, there were too few cases to estimate parameters in the Poisson model and for moderately committed in the logistic model. The result for the fully committed is not significant.

While there is no evidence that the fully and moderately committed were any more or less accurate than the least committed for visits to the ER or hospital, there is evidence that fully committed respondents were less accurate for the relatively infrequent visits to the ER and hospital when compared to the control group. As shown in Table 3.8a, fully committed respondents were significantly more likely to overreport visits to the ER and hospital compared to the control group. As a possible downside to being fully committed, commitment might motivate respondents to report *something* because they feel that to report no events violates or may give the appearance of violating the agreement they entered into to take the task seriously and invest effort.

Table 3.7. The Effect of Commitment Level Groups – Visits to the Pediatrician, Specialist, ER, and Hospital – Reference: Least Committed

	Visits to the Pediatrician		Visits to a Specialist	
VARIABLES	Logistic	Poisson	Logistic	Poisson
Fully committed	0.48*	0.19**	0.57*	0.39***
	(0.231)	(0.066)	(0.230)	(0.082)
Moderately committed	0.23	0.17*	0.62*	0.42***
	(0.250)	(0.071)	(0.247)	(0.087)

Feedback	-0.04	-0.00	-0.14	0.03
	(0.154)	(0.044)	(0.153)	(0.049)
CRCs	-0.13	-0.06	-0.30^	0.15**
	(0.159)	(0.045)	(0.157)	(0.050)
Visits in records		0.18***		0.16***
		(0.006)		(0.003)
Constant	-0.83***	-11.56***	-0.66**	-12.04***
	(0.239)	(0.071)	(0.238)	(0.086)
Observations	1,422	1,420	1,397	1,395

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Visits to the ER		Visits to the Hospital
VARIABLES	Logistic	Poisson	Logistic
Fully committed	0.13	0.35^	-1.36
	(0.312)	(0.206)	(1.129)
Moderately committed	0.02	0.29	
	(0.336)	(0.221)	
Feedback	0.35	-0.28*	0.04
	(0.229)	(0.122)	(1.236)
CRCs	0.09	-0.47***	1.73
	(0.223)	(0.127)	(1.118)
Visits in records		0.66***	
		(0.019)	
Constant	1.68***	-13.50***	-0.28
	(0.321)	(0.209)	(1.117)
Observations	1,400	1,398	37

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Table 3.8. The Effect of Commitment Level Groups on Overreporting and Underreporting

	a. Reference: Control			
VARIABLES	Visits to the Pediatrician	Visits to a Specialist	Visits to the ER	Visits to the Hospital
	logit 2: Overreporting versus zero difference			
Fully committed	-0.03	0.23	0.46*	0.56*
	(0.176)	(0.196)	(0.230)	(0.250)
Moderately committed	0.23	-0.11	0.41	0.03
	(0.201)	(0.227)	(0.264)	(0.305)
Least committed	0.42	0.53^	0.69*	0.67^
	(0.281)	(0.318)	(0.341)	(0.356)
Feedback	-0.02	-0.18	-0.49*	-0.27
	(0.163)	(0.186)	(0.218)	(0.230)
CRCs	-0.06	0.15	-0.28	-0.10
	(0.170)	(0.186)	(0.216)	(0.228)

Visits in records	0.07	0.41***	0.88***	0.86***
	(0.044)	(0.044)	(0.116)	(0.223)
Constant	0.01	-1.24***	-2.33***	-2.37***
	(0.151)	(0.154)	(0.177)	(0.185)
logit 3: Underreporting versus zero difference				
Fully committed	-0.63**	-0.31	0.13	-3.89
	(0.239)	(0.200)	(0.447)	(1,519.660)
Moderately committed	-0.07	-0.14	0.52	-20.25
	(0.268)	(0.219)	(0.482)	(1,975.866)
Least committed	0.36	0.69*	0.66	-3.04
	(0.349)	(0.296)	(0.611)	(1,519.661)
Feedback	0.10	0.22	0.06	23.93
	(0.230)	(0.185)	(0.395)	(1,146.697)
CRCs	0.37	0.28	0.11	24.89
	(0.229)	(0.192)	(0.408)	(1,146.696)
Visits in records	0.62***	0.46***	1.54***	4.65***
	(0.049)	(0.043)	(0.148)	(1.182)
Constant	-2.11***	-1.16***	-4.20***	-27.02
	(0.204)	(0.149)	(0.346)	(997.239)
Observations	1,479	1,479	1,479	1,479

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	b. Reference: Least committed			
VARIABLES	Visits to the Pediatrician	Visits to a Specialist	Visits to the ER	Visits to the Hospital
logit 2: Overreporting versus zero difference				
Fully committed	-0.45^	-0.29	-0.24	-0.11
	(0.245)	(0.280)	(0.294)	(0.294)
Moderately committed	-0.19	-0.64*	-0.28	-0.64^
	(0.266)	(0.305)	(0.323)	(0.345)
Least committed				
Feedback	-0.02	-0.18	-0.49*	-0.27
	(0.163)	(0.186)	(0.218)	(0.230)
CRCs	-0.06	0.15	-0.28	-0.10
	(0.170)	(0.186)	(0.216)	(0.228)
Visits in records	0.07	0.41***	0.88***	0.86***
	(0.044)	(0.044)	(0.116)	(0.223)
Constant	0.43	-0.71*	-1.63***	-1.70***
	(0.267)	(0.293)	(0.301)	(0.306)
logit 3: Underreporting versus zero difference				
Fully committed	-0.99***	-1.00***	-0.53	-0.85
	(0.302)	(0.257)	(0.514)	(1.364)
Moderately committed	-0.43	-0.84**	-0.14	-17.21
	(0.326)	(0.276)	(0.544)	(1,262.805)

Least committed				
Feedback	0.10	0.22	0.06	23.93
	(0.230)	(0.185)	(0.395)	(1,146.697)
CRCs	0.37	0.28	0.11	24.89
	(0.229)	(0.192)	(0.408)	(1,146.696)
Visits in records	0.62***	0.46***	1.54***	4.65***
	(0.049)	(0.043)	(0.148)	(1.182)
Constant	-1.74***	-0.46^	-3.55***	-30.06
	(0.339)	(0.272)	(0.540)	(1,146.700)
Observations	1,479	1,479	1,479	1,479

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

In the previous chapter, we saw that commitment group respondents overall, compared to the control group, maintained a higher level of accuracy (likelihood of a match) for visits to a specialist as the number of visits in the records increased (a proxy measure for difficulty). To examine potential reporting differences by commitment level group, interaction terms for commitment level and the number of visits in the records are added to the logistic, Poisson and multinomial models, in addition to the other covariates: education level (college degree or higher) and whether or not the respondent reported checking a calendar or other relevant documents. The reference group for these analyses is the control group.

Results for the logistic and Poisson models for visits to the pediatrician, specialist, and ER are shown in Table 3.9 and multinomial in Table 3.10. Results are not presented for visits to the hospital because there were no significant effects for any of the covariates and the small number of observations made the parameter estimates for the interactions unstable.

No significant interactions are observed for visits to the pediatrician and the interaction terms for visits to the ER were dropped as the small number of observations made the parameter estimates unstable. However, in predicting the likelihood of a match for visits to a specialist, we see significant interactions between the number of visits in the records and (1) each of the commitment level groups (fully, moderately, and least) and (2) college degree or higher for the fully and least committed. Please note that the model is restricted to those with fewer than eight visits in the records because of the small

number of cases with eight or more visits. Figure 3.1 shows the regression lines for the likelihood of a match for visits to a specialist by commitment level group and education level by number of visits. Only one line is displayed for the moderately committed as there was no significant effect of education level for these respondents. Compared to the control group, the slope of the lines for the fully and moderately committed respondents (at both education levels) is less steep and nearly flat for the least committed compared to the control group suggesting that respondents in the commitment treatment, at all levels of commitment, were less inaccurate (more likely to report a match) as the number of visits increased. As expected, it is harder to accurately report a larger than smaller amount of visits for respondents at all commitment levels. However, education appears to moderate this effect for the fully committed, and the least committed in particular, where those with a college degree or higher tended to be more accurate than those without a college degree across number of visits. In the case of the least committed, it is important to keep in mind that these respondents started out with a low level of accuracy even when the number of visits is small, and that the number of cases with a high number of visits is relatively small. Overall, this suggests a college degree or higher and at least modest commitment helps for several visits but if the number of visits is large enough, even these advantages may not help much.

For visits to a specialist, we also see, in the Poisson model, a significant negative interaction for fully committed and a significant positive interaction for moderately committed with the number of visits in the records. This suggests that, on average, fully committed respondents reported fewer visits and that moderately committed reported more visits relative to the control group as the number of visits increased. Given the corresponding positive and significant results for these respondents in the logistic model and significantly less overreporting for the fully committed and significantly less underreporting for the moderately committed in the multinomial model, this suggests reporting fewer visits helped improve accuracy for the fully committed and reporting more increased accuracy for the moderately committed as the number of visits increased.

Table 3.9. The Effect of Commitment Level Groups and Covariates – Visits to the Pediatrician, Specialist, and ER – Reference: Control

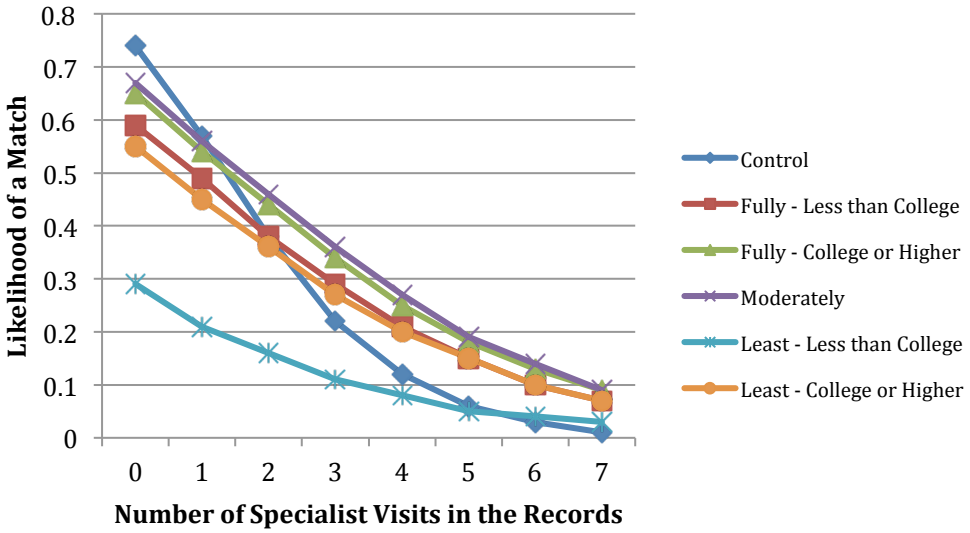
* Note: Hospital visits are not reported because none of the covariates or interaction terms was significant

	Visits to the Pediatrician		Visits to a Specialist		Visits to the ER	
VARIABLES	Logistic	Poisson	Logistic	Poisson	Logistic	Poisson
Fully committed	-0.38	0.12	-1.00**	0.12	-0.36	0.16
	(0.364)	(0.105)	(0.366)	(0.098)	(0.267)	(0.122)
Moderately committed	-0.63	0.23*	-0.72^	0.00	-0.22	-0.01
	(0.426)	(0.118)	(0.410)	(0.116)	(0.294)	(0.143)
Least committed	-0.75	0.17	-2.29***	0.03	-0.41	-0.33
	(0.638)	(0.175)	(0.644)	(0.169)	(0.396)	(0.232)
Feedback	-0.08	-0.03	-0.01	-0.09^	0.33	-0.25*
	(0.165)	(0.046)	(0.167)	(0.053)	(0.251)	(0.124)
CRCs	-0.10	-0.06	-0.22	0.12*	0.07	-0.08
	(0.169)	(0.047)	(0.169)	(0.053)	(0.245)	(0.125)
Checked	0.14	0.08*	0.17	0.22***	0.53**	0.19^
	(0.134)	(0.038)	(0.136)	(0.044)	(0.197)	(0.102)
Visits in records	-0.32***	0.18***	-0.78***	0.16***	-1.32***	0.79***
	(0.079)	(0.016)	(0.116)	(0.005)	(0.124)	(0.024)
Fully committed * Visits in records	0.08	0.00	0.35**	-0.02*		
	(0.094)	(0.019)	(0.135)	(0.007)		
Moderately committed * Visits in records	0.08	0.01	0.36*	0.04**		
	(0.125)	(0.025)	(0.154)	(0.012)		
Least committed * Visits in records	-0.17	-0.03	0.39^	-0.01		
	(0.223)	(0.048)	(0.235)	(0.022)		
College or higher	-0.07	0.04	-0.44	-0.02	0.53**	-0.41***
	(0.251)	(0.071)	(0.275)	(0.078)	(0.182)	(0.090)
Fully committed * College or higher	0.48	-0.04	0.67*	0.06		
	(0.317)	(0.087)	(0.334)	(0.097)		
Moderately committed * College or higher	0.44	-0.24	0.48	-0.04		
	(0.379)	(0.103)	(0.383)	(0.118)		
Least committed * College or higher	1.01^	-0.24	1.54*	-0.26		
	(0.583)	(0.154)	(0.610)	(0.188)		
Constant	0.21	-11.45***	1.33***	-11.90***	2.08***	-13.31***
	(0.280)	(0.072)	(0.298)	(0.077)	(0.229)	(0.114)
Observations	1,333	1,333	1,287	1,353	1,362	1,363

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Figure 3.1. Regression Lines for the Likelihood of a Match for Visits to a Specialist by Commitment Level Group and Education Level by Number of Visits in the Records



While there is evidence that reporting accuracy improved for the fully and moderately committed as the number of visits increased, there is also evidence that these respondents were less accurate when there were zero visits in the medical records for visits to a specialist. Compared to the control group, fully committed respondents were significantly less likely to report a matching number of visits to a specialist when there were zero visits in the records and moderately committed respondents were marginally less likely. This is apparent in Figure 3.3, where the regression lines for the fully and moderately committed are at a lower percentage of a match at zero visits in the records. Similar to the results discussed above for the fully committed for ER and hospital visits, this could well reflect a possible cost of commitment: respondents may feel that to report zero visits gives the impression of not trying hard, especially since they know they have been sampled because their child has been a patient in the university's medical system, so they report visits that did not actually occur, at least within the study's reference period. Results in the multinomial models are consistent with these results; fully committed respondents were significantly more likely to overreport, and moderately respondents were marginally so.

Also of note is the main effect for checking records in the Poisson models for visits to the pediatrician and specialist that suggest that respondents who reported

checking records reported more visits on average compared to the control group. This may not have improved accuracy of reported visits in all cases in terms of exact match since there is no corresponding main effect for checking records in the logistic model. However, negative and significant coefficients for checking records for underreporting in the multinomial regression model shown in Table 3.10 indicates that checking records improved accuracy for visits to the pediatrician, specialist, and the ER. For visits to the ER, checking records is also a significant predictor of a match (as shown in Table 3.9) in addition to significantly less underreporting.

Table 3.10. The Effect of Commitment Level Groups and Covariates on Overreporting and Underreporting

VARIABLES	Visits to the Pediatrician	Visits to a Specialist	Visits to the ER
logit 2: Overreporting versus zero difference			
Fully committed	0.33	1.09*	0.29
	(0.390)	(0.442)	(0.291)
Moderately committed	0.71	0.71	0.17
	(0.462)	(0.508)	(0.331)
Least committed	0.72	2.12**	0.32
	(0.687)	(0.780)	(0.462)
Feedback	0.03	-0.20	-0.66*
	(0.173)	(0.208)	(0.298)
CRCs	-0.04	0.20	-0.14
	(0.179)	(0.203)	(0.273)
Checked	-0.04	0.03	-0.12
	(0.143)	(0.168)	(0.225)
Visits in records	0.07	0.62***	1.22***
	(0.095)	(0.129)	(0.139)
Fully committed * Visits in records	-0.00	-0.36*	
	(0.112)	(0.154)	
Moderately committed * Visits in records	-0.09	-0.31^	
	(0.150)	(0.181)	
Least committed * Visits in records	0.19	-0.47	
	(0.255)	(0.289)	
College or higher	-0.01	0.40	-0.81***
	(0.268)	(0.336)	(0.204)
Fully committed * College or higher	-0.47	-0.57	
	(0.336)	(0.406)	
Moderately committed * College or higher	-0.39	-0.65	
	(0.400)	(0.482)	
Least committed * College or higher	-1.02^	-1.54*	
	(0.608)	(0.753)	
Constant	-0.05	-1.98***	-2.28***
	(0.305)	(0.362)	(0.254)

logit 3: Underreporting versus zero difference			
Fully committed	0.35	0.79^	0.57
	(0.539)	(0.430)	(0.471)
Moderately committed	0.11	0.64	0.24
	(0.641)	(0.474)	(0.503)
Least committed	0.50	2.34***	0.50
	(0.890)	(0.688)	(0.641)
Feedback	0.12	0.17	0.24
	(0.241)	(0.196)	(0.414)
CRCs	0.34	0.22	0.17
	(0.238)	(0.201)	(0.422)
Checked	-0.41*	-0.35*	-1.54***
	(0.186)	(0.157)	(0.347)
Visits in records	0.68***	0.87***	1.89***
	(0.100)	(0.124)	(0.171)
Fully committed * Visits in records	-0.15	-0.30*	
	(0.120)	(0.147)	
Moderately committed * Visits in records	0.04	-0.37*	
	(0.165)	(0.165)	
Least committed * Visits in records	0.24	-0.38	
	(0.277)	(0.242)	
College or higher	0.17	0.39	0.11
	(0.340)	(0.312)	(0.319)
Fully committed * College or higher	-0.55	-0.70^	
	(0.433)	(0.386)	
Moderately committed * College or higher	-0.41	-0.31	
	(0.517)	(0.436)	
Least committed * College or higher	-0.72	-1.47*	
	(0.734)	(0.652)	
Constant	-2.22***	-1.91***	-3.96***
	(0.411)	(0.342)	(0.424)
Observations	1,370	1,294	1,369

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

3.4.3. Response Accuracy: Date Questions

Table 3.11 provides the percentage of matches and the mean absolute differences between reported dates and dates in the records for commitment treatment versus control and by level of commitment for each visit type. The percentages of matches for visits to the ER and hospital are not reported as there were too few matches (4 and 1 respectively). As is evident in the table, the fully committed respondents, compared to the least committed, had a higher percentage of matches, lower absolute differences, and higher

percentages of reporting a full date for the date of the last visit to the pediatrician, specialist, ER and hospital.

Table 3.11. Percentage of a Match, Mean Absolute Differences and Percentage Providing Full Date for Date of Last Visit by Commitment and Commitment Level

	Date of last visit to the Pediatrician					
	Match		Absolute difference		Full Date	
	%	(se)	%	(se)	%	(se)
Control	0.38	0.03	37.02	4.44	0.84	0.02
Commitment	0.47	0.02	36.67	2.54	0.84	0.01
Least committed	0.31	0.06	58.12	11.78	0.78	0.04
Moderately committed	0.39	0.04	43.95	6.19	0.71	0.03
Fully committed	0.51	0.02	32.19	2.76	0.90	0.01

	Date of last visit to a Specialist					
	Match		Absolute difference		Full Date	
	%	(se)	%	(se)	%	(se)
Control	0.41	0.04	48.64	6.20	0.80	0.03
Commitment	0.50	0.02	39.01	3.32	0.83	0.01
Least committed	0.18	0.06	57.02	14.07	0.79	0.06
Moderately committed	0.37	0.05	47.62	8.79	0.64	0.04
Fully committed	0.53	0.02	35.04	3.63	0.91	0.01

	Date of last visit to the ER				Date of last visit to the Hospital			
	Absolute difference		Full Date		Absolute difference		Full Date	
	%	(se)	%	(se)	%	(se)	%	(se)
Control	32.30	11.51	0.82	0.04	42.57	36.79	0.87	0.07
Commitment	32.18	6.24	0.79	0.03	32.73	21.87	0.80	0.04
Least committed	61.42	58.83	0.61	0.14	123.18	121.03	0.67	0.21
Moderately committed	31.59	9.32	0.57	0.06	0.29	0.11	0.57	0.14
Fully committed	30.97	7.23	0.88	0.02	23.90	22.74	0.87	0.04

Given the higher percentage of a reporting a full and matching date among the fully committed in the descriptive results compared to the moderately and least committed as well as the control group, we begin my comparing the commitment level groups to the control group. The regression results in Table 3.12 show that fully committed respondents were significantly more likely to report a matching date for both visits to the pediatrician and a specialist. Fully committed respondents were also significantly more likely to report a full date for the last visit to a specialist. Moderately committed respondents, on the other hand, were no more likely to report a matching date for any visits and were significantly less likely to report a full date for the date of the last visit to the pediatrician, specialist, ER, and marginally so to the hospital. The least committed were significantly less likely to report a matching date for the last visit to a specialist and were significantly less likely to report a full date for the last visit to the pediatrician and ER.

Table 3.12. The Effect of Commitment Level Groups – Date of the Most Recent Visits to the Pediatrician, Specialist, ER and Hospital – Reference: Control

	Date of Last Visit Pediatrician's Office		Date of Last Visit to a Specialist		Date of Last Visit to the ER	Date of Last Visit to the Hospital
VARIABLES	Logistic	Full date	Logistic	Full date	Full date	Full date
Fully committed	0.61**	0.27	0.56*	0.66*	0.22	-0.01
	(0.192)	(0.239)	(0.220)	(0.267)	(0.415)	(0.835)
Moderately committed	0.13	-0.99***	-0.12	-1.10***	-1.75***	-1.62^
	(0.235)	(0.250)	(0.273)	(0.264)	(0.433)	(0.895)
Least committed	-0.21	-0.68*	-1.08*	-0.46	-1.89**	-1.43
	(0.320)	(0.337)	(0.444)	(0.404)	(0.724)	(1.158)
Feedback	-0.07	0.49*	0.16	0.75**	1.45**	0.58
	(0.177)	(0.224)	(0.205)	(0.268)	(0.453)	(0.814)
CRCs	-0.18	0.19	-0.24	0.10	-0.02	-0.28
	(0.185)	(0.225)	(0.215)	(0.249)	(0.395)	(0.672)
Constant	-0.48***	1.65***	-0.43**	1.40***	1.52***	1.90**
	(0.130)	(0.158)	(0.148)	(0.163)	(0.276)	(0.619)
Observations	1,039	1,236	795	963	346	104

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Comparing the fully committed to the least committed, we see in the regression results in Table 3.13, that the fully committed respondents were significantly more likely to report a matching date and to report a full date for both visits to the pediatrician and a

specialist. Fully committed respondents were also significantly more likely to report a full date for the last visit to the ER than the least committed.

Table 3.13. The Effect of Commitment Level Groups – Date of the Most Recent Visits to the Pediatrician, Specialist, ER and Hospital – Reference: Least committed

	Date of Last Visit Pediatrician's Office		Date of Last Visit to a Specialist		Date of Last Visit to the ER	Date of Last Visit to the Hospital
VARIABLES	Logistic	Full date	Logistic	Full date	Full date	Full date
Fully committed	0.83**	0.94**	1.64***	1.12**	2.11**	1.41
	(0.277)	(0.293)	(0.404)	(0.367)	(0.677)	(0.983)
Moderately committed	0.35	-0.31	0.96*	-0.64^	0.14	-0.20
	(0.311)	(0.301)	(0.438)	(0.366)	(0.671)	(1.052)
Feedback	-0.07	0.49*	0.16	0.75**	1.45**	0.58
	(0.177)	(0.224)	(0.205)	(0.268)	(0.453)	(0.814)
CRCs	-0.18	0.19	-0.24	0.10	-0.02	-0.28
	(0.185)	(0.225)	(0.215)	(0.249)	(0.395)	(0.672)
Constant	-0.70*	0.98**	-1.51***	0.94*	-0.37	0.47
	(0.292)	(0.298)	(0.418)	(0.369)	(0.670)	(0.978)
Observations	1,039	1,236	795	963	346	104

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

The following covariates were examined as main effects and as interactions with commitment level group for the date of the last visit for both the likelihood of a match and that a full date was reported: the length of time since that last visit in the records (in days), whether or not the respondent reported checking a calendar or other relevant documents, and the respondent's education level (college degree or higher). The interaction between checking records and commitment level was not significant and was subsequently dropped from the models. Results for the date of last visit to the pediatrician and a specialist are shown in Table 3.14. No results are reported for the last visit to the ER or hospital because none of the covariates or interactions was significant.

If commitment level helped improve accuracy as the number of days since the last visit in the records increased, we would expect to see positive and significant interactions between the commitment level groups and the number of number of days in the records in the logistic regression models. We see a significant interaction between the date of the last visit in the records and fully committed for the likelihood of a match for the fully committed for the last visit to the pediatrician and for the fully committed and the least

committed for the likelihood of a match for the last visit to a specialist. However, the result for the least committed is based on very few cases.

The regression lines for the likelihood of a match for the last visit to the pediatrician by commitment level group by number of days since the visit are presented in Figure 3.2. The main thing to note is that the slope of the line for the fully committed respondents is less steep suggesting that being fully committed helped improve accuracy for these respondents even as the task became more difficult. This is also the case for the date of the last visit to a specialist, as shown in Figure 3.3.

Table 3.14. The Effect of Commitment Level Groups and Covariates – Date of the Most Recent Visits to the Pediatrician and Specialist

	Date of Last Visit Pediatrician's Office		Date of Last Visit to a Specialist	
VARIABLES	Match	Full date	Match	Full date
Fully committed	0.06	0.05	-0.42	-0.24
	(0.420)	(0.501)	(0.469)	(0.612)
Moderately committed	0.20	-0.51	-1.15^	-1.39*
	(0.520)	(0.524)	(0.614)	(0.588)
Least committed	0.36	-0.83	-2.85*	-1.96*
	(0.727)	(0.703)	(1.212)	(0.991)
Feedback	-0.14	0.42^	0.11	0.73*
	(0.204)	(0.243)	(0.230)	(0.307)
Checked records	2.26***	1.44***	2.31***	1.66***
	(0.216)	(0.187)	(0.258)	(0.217)
CRCs	-0.23	0.28	-0.18	0.22
	(0.212)	(0.243)	(0.243)	(0.289)
Recent visit in records	-0.01***	-0.00*	-0.01***	-0.00^
	(0.002)	(0.002)	(0.002)	(0.002)
Fully committed * Recent visit in records	0.00*	-0.00	0.00*	0.00
	(0.002)	(0.002)	(0.002)	(0.002)
Moderately committed * Recent visit in records	-0.00	0.00	0.00	0.00
	(0.003)	(0.002)	(0.003)	(0.002)
Least committed * Recent visit in records	-0.01	0.00	0.01*	0.01^
	(0.005)	(0.003)	(0.006)	(0.007)
College or higher	0.50	0.21	0.06	-0.57
	(0.351)	(0.356)	(0.382)	(0.389)
Fully committed * College or higher	-0.26	-0.10	0.08	0.54
	(0.414)	(0.468)	(0.453)	(0.554)
Moderately committed * College or higher	0.38	-0.87^	0.85	0.24
	(0.569)	(0.501)	(0.629)	(0.534)
Least committed * College or higher	0.05	-0.20	0.53	0.89
	(0.799)	(0.686)	(1.048)	(0.931)
Constant	-1.49***	1.42***	-1.14**	1.48***
	(0.357)	(0.360)	(0.393)	(0.418)

Observations	988	1,176	747	896

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Figure 3.2. Regression lines for the Likelihood of a Match for Date of Last Visit to the Pediatrician by Commitment Level Group and Time Since the Last Visit to the Pediatrician

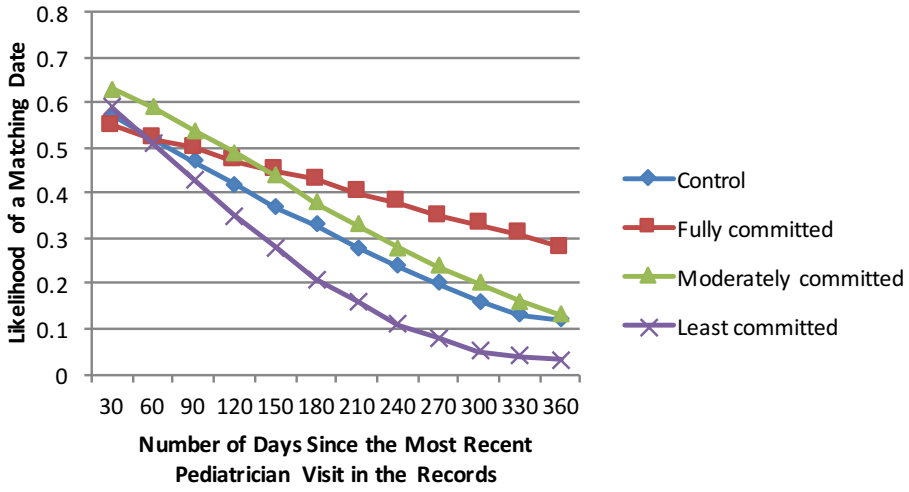
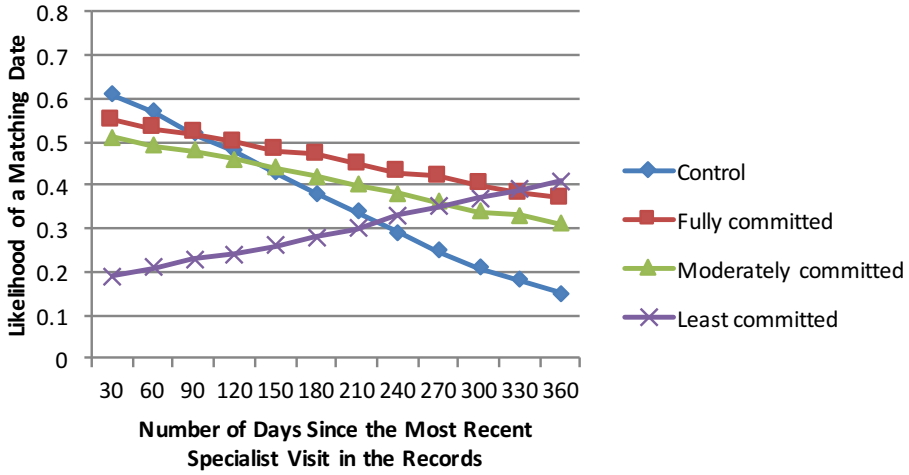


Figure 3.3. Regression Lines for the Likelihood of a Match for Reported Date of Last Visit to a Specialist by Commitment Level Group and Time Since the Last Visit to a Specialist



3.4.4. Number of Mentions and Words to Open-ended Questions

Respondents were asked the following open-ended questions:

4. What types of foods, vitamins and/or supplements do you give [CHILD] to stay healthy?
5. What have you cut down on giving [CHILD] because you think that it is bad for his/her health?
6. What would you like to do to maintain or improve [CHILD]'s health?

Three one-line text fields appeared for each question. Respondents in the Commitment+Feedback group received feedback if they left the second or third line blank. The regression results shown below control for the additional effect of feedback.

Comparing the mean number of words and mentions among the commitment level groups in Table 3.15, we see that the least committed used fewer words and fewer mentions than the moderately and fully committed. "Mentions" refers to the number of response fields filled in by the respondent; three response fields were provided for each open-ended question.

Table 3.15. Mean Number of Mentions and Word Count to Open-ended Questions by Commitment Level

	Word count		Mentions	
	mean	(se)	Mean	(se)
Control	12.59	0.08	4.36	0.12
Commitment	14.73	0.47	4.94	0.10
Least committed	12.42	1.22	4.64	0.29
Moderately committed	13.21	0.69	4.71	0.14
Fully committed	15.92	0.66	5.12	0.10

Based on the regression results in Table 3.16a, we see that fully committed respondents used significantly more words ($p < .05$) and marginally more mentions ($p < .1$) than the control group. The comparison is stronger between fully committed respondents and their least committed counterparts, as shown in Table 3.16b. There was a marginally significant main effect of college or higher education for word count but no significant interaction effect with the commitment level groups.

Table 3.16. The Effect of Commitment Level Groups on the Word Count and Number of Mentions to Open-ended Questions

	a. Reference: Control		b. Reference: Least committed	
VARIABLES	Word count	Mentions	Word count	Mentions
Fully committed	2.16*	0.29^	3.74*	0.52*
	(1.020)	(0.167)	(1.524)	(0.249)
Moderately committed	-0.54	-0.11	1.05	0.12
	(1.208)	(0.197)	(1.656)	(0.271)
Least committed	-1.59	-0.23		
	(1.657)	(0.271)		
Feedback	3.50***	1.22***	3.50***	1.22***
	(0.937)	(0.153)	(0.937)	(0.153)
Constant	12.52***	4.40***	10.94***	4.17***
	(0.781)	(0.128)	(1.461)	(0.239)
Observations	1,481	1,481	1,481	1,481
R-squared	0.021	0.055	0.021	0.055

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

3.4.5. Straightlining and Acquiescence

Answering the same way to five out of six (e.g. "near straightlining") items in a battery (Yes/No) is the dependent variable for straightlining. The dependent variable includes "near straightlining" because complete straightlining (e.g. the same answer to all six out of six items) was rare – just two cases – mostly likely because the battery of items included a reverse-coded item. For acquiescence, the dependent variable is the number of "agree" or "strongly agree" answers in a battery of questions.

As shown in Table 3.17, we can see by looking at the percentage of straightlining for the commitment level groups that, contrary to expectation, it is the least committed respondents who were less likely to straightline while higher percentages of both the moderately committed and the fully committed straightlined. Similarly, for acquiescence, the mean number of "agree" or "strongly agree" answers is also lower for the least committed group and this difference is statistically significant ($p < .05$). Here again it is the least committed who were less acquiescent with their answers than the moderately and fully committed. This may have had more to do with true differences between the fully committed and least committed on the topic of the questions rather than with the effect of commitment level: fully committed respondents may have actually been more satisfied with their child's pediatrician than were less committed respondents; their

willingness to fully commit to the survey may have been due, in part, to their satisfaction with the care they receive at the university's health system.

Consistent with these results, Table 3.18b shows that fully committed respondents were marginally more likely to straightline and acquiesce compared to the least committed. As shown in Table 3.18a, the amount of straightlining and acquiescence for the fully and moderately committed is not significantly different from the control group. There was no significant main effect for college or higher education level or in interaction with commitment level or commitment level group.

Table 3.17. Percentage Straightlining and Mean Number of "Agree" or "Strongly Agree" Answers for Commitment Level

	Straightlining		Acquiescence	
	%	(se)	%	(se)
Control	0.49	0.03	4.33	0.07
Commitment	0.47	0.01	4.17	0.05
Least committed	0.38	0.05	3.95	0.16
Moderately committed	0.47	0.03	4.18	0.08
Fully committed	0.48	0.02	4.22	0.06

Table 3.18. The Effect of Commitment Level Groups on Straightlining and Acquiescence

	a. Reference: Control		b. Reference: Least committed	
VARIABLES	Straightlining	Acquiescence	Straightlining	Acquiescence
Fully committed	-0.01	-0.15	0.38^	0.28^
	(0.136)	(0.101)	(0.209)	(0.151)
Moderately committed	-0.03	-0.19	0.36	0.24
	(0.162)	(0.119)	(0.226)	(0.164)
Least committed	-0.40^	-0.43**		
	(0.226)	(0.164)		
Feedback	-0.07	0.11	-0.07	0.11
	(0.126)	(0.093)	(0.126)	(0.093)
Constant	-0.05	4.33***	-0.45*	3.90***
	(0.104)	(0.077)	(0.200)	(0.145)
Observations	1,481	1,481	1,481	1,481
R-squared		0.005		0.005

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

3.4.6. Socially desirable reporting

Responses to the behavioral frequency questions are used as indicators of socially desirable reporting. Three questions asked about socially desirable behaviors – the number of days in the last week (0 – 7) that the respondent's child got enough sleep, exercise, and had a family meal. For these questions, reporting a higher number of days the child got enough sleep, for example, could be considered socially desirable. Several other items ask about socially undesirable behaviors – the number of hours the respondent's child watched TV, and number of hours the child spent on the computer in the last week, how often the respondent raises their voice, spanks their child, smokes in the presence of their child or allows others to smoke in the presence of their child. For these items, lower reports, for instance reporting never raising one's voice, could be interpreted as socially desirable.

We begin with the effect of commitment level on reporting socially desirable behaviors. Based on the mean number of days reported in Table 3.19, we see that commitment group respondents overall compared with the control group reported slightly less for each of the measures. However, examining the mean number of days reported by commitment level group, we see that it is the least committed respondents who reported fewer days compared to the moderately and fully committed. Regression results presented in Table 3.20 show that the difference between fully committed and least committed is highly significant for exercise ($p < .001$) but not significant for the other two measures. It is possible that the kind of respondent who will fully commit to conscientious responding in a survey is also the kind of parent who is particularly conscientious about taking steps to improve their child's health, such as insuring that they get enough exercise.

Table 3.19. Mean Number of Days Child Got Enough Sleep, Got Exercise, and Had a Family Meal by Commitment Level

	Number of days child got enough sleep		Number of days child got exercise		Number of days had family meal	
	Mean	(se)	Mean	(se)	Mean	(se)
Control	5.94	0.08	4.35	0.11	5.22	0.10
Commitment	5.89	0.04	4.43	0.06	5.07	0.06
Least committed	5.72	0.17	3.87	0.19	4.91	0.22

Moderately committed	5.93	0.08	4.24	0.12	4.94	0.12
Fully committed	5.90	0.06	4.60	0.07	5.16	0.07

Table 3.20. The Effect of Commitment Level Groups on the Days Child Got Enough Sleep, Got Exercise, and Had a Family Meal – Reference: Least committed

VARIABLES	Enough sleep	Exercise	Family meal
Fully committed	0.16	0.73***	0.26
	(0.158)	(0.211)	(0.212)
Moderately committed	0.18	0.37	0.04
	(0.171)	(0.229)	(0.229)
Feedback	-0.25**	-0.12	0.11
	(0.095)	(0.127)	(0.127)
Constant	5.83***	3.92***	4.86***
	(0.153)	(0.204)	(0.204)
Observations	1,364	1,363	1,378
R-squared	0.006	0.013	0.004

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Turning next to socially undesirable behaviors, we see in the descriptive results in Table 3.21, that the fully committed respondents were less likely to report more TV and computer time, raising their voice, and spanking but slightly more likely to report smoking or allowing others to smoke in the child's presence, compared to the least committed. These results are reflected in the regression shown in Table 3.22, which show that fully committed respondents reported significantly less TV and computer time ($p < .001$), raising one's voice ($p < .01$), and marginally less spanking. So, as with the socially desirable behaviors, it is the least committed respondents who were more likely to report in the expected direction, as far as increased disclosure of socially undesirable behaviors.

Table 3.21. Means for Reports of Socially Undesirable Behaviors by Commitment Level Group

	> 2-3 hours of TV per day		> 2-3 hours of computer time per day		Raises voice at child > never	
	Mean	(se)	Mean	(se)	Mean	(se)
Control	0.19	0.02	0.12	0.02	0.09	0.02
Commitment	0.23	0.01	0.15	0.01	0.09	0.01
Least committed	0.37	0.05	0.31	0.05	0.20	0.04
Moderately committed	0.22	0.02	0.17	0.02	0.07	0.02
Fully committed	0.21	0.02	0.12	0.01	0.08	0.01

	Spanks child > never		Smokes in child's presence > never		Allows others to smoke in child's presence > never	
	Mean	(se)	Mean	(se)	Mean	(se)
Control	0.14	0.02	0.03	0.01	0.05	0.01
Commitment	0.12	0.01	0.04	0.01	0.05	0.01
Least committed	0.18	0.04	0.01	0.01	0.04	0.02
Moderately committed	0.10	0.02	0.07	0.01	0.05	0.01
Fully committed	0.12	0.01	0.03	0.01	0.05	0.01

Table 3.22. The Effect of Commitment Level Groups on Reports of Socially Undesirable Behaviors – Reference: Least committed

VARIABLES	> 2-3 hours of TV per day	> 2-3 hours of computer time per day	Raises voice at child > never	Spanks child > never	Smokes in child's presence > never	Allows others to smoke in child's presence > never
Fully committed	-0.79***	-1.17***	-0.97**	-0.52^	1.20	0.19
	(0.229)	(0.254)	(0.299)	(0.295)	(1.029)	(0.542)
Moderately committed	-0.72**	-0.74**	-1.18***	-0.75*	1.96^	0.25
	(0.252)	(0.274)	(0.355)	(0.337)	(1.033)	(0.576)
Feedback	0.11	0.15	-0.11	-0.44*	0.08	-0.40
	(0.154)	(0.182)	(0.233)	(0.217)	(0.333)	(0.321)
Constant	-0.56**	-0.88***	-1.37***	-1.36***	-4.63***	-3.02***
	(0.218)	(0.235)	(0.278)	(0.279)	(1.015)	(0.522)
Observations	1,377	1,367	1,352	1,367	1,391	1,386

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

3.4.7. Break-offs and Response Time

This section examines the effect of commitment level group on the number of break-offs, which is starting but not completing the survey, and overall response time. Based on the descriptive results shown in Table 3.23, we see that overall, commitment group respondents were more likely to break-off than control group respondents. While a higher percentage of the least committed respondents broke off, the moderately and fully committed respondents also broke off at a rate higher than that of the control group. As shown in Table 3.24, the difference is significant for the fully and least committed, and marginally significant for the moderately committed.

However, looking at median response time, we see that commitment group respondents took longer to complete the survey overall, with fully committed respondents taking the most time, followed by the moderately committed and then the least committed. The regression results in Table 3.24 show that the increase in response time is significant for all commitment levels but especially for the fully committed group.

Table 3.23. Percentages of Break-offs and Median Response Time by Commitment Level Group

	Break-off		Median Response Time
	n	%	
Control	19	5.2	10.28
Commitment	95	8.5	13.33
Least committed	14	12.5	12.63
Moderately committed	23	7.6	12.73
Fully committed	53	7.6	13.88

Table 3.24. The Effect of Commitment Level Group on Break-offs and Median Response Time by Commitment Level Group – Reference: Control

VARIABLES	Break-off	Median Response Time
Fully committed	0.66*	3.38***
	(0.308)	(0.651)
Moderately committed	0.66^	1.92**
	(0.345)	(0.737)
Least committed	1.25**	2.28*
	(0.402)	(0.986)
Feedback	-0.34	1.02^
	(0.260)	(0.610)
CRCs	-0.47^	-0.18
	(0.276)	(0.625)
Constant	-2.91***	10.28***
	(0.236)	(0.436)
Observations	1,481	1,481

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

3.5. Discussion

The results reported in this chapter show that by allowing respondents to indicate which desired response behaviors they will and will not commit to – commitment level – can have a profound effect on data quality. A number of significant effects were

observed, within treatment and compared to the control group, based on how respondents responded to the request to commit to certain desired response behaviors, which included – Reading all of the questions carefully; trying to be as precise as possible; looking up information in records or on a calendar, if needed; providing as much information as possible, and; answering honestly. Respondents were considered fully committed if they agreed to adhere to all of the requested behaviors, moderately committed if they agreed to four out of five (overwhelmingly, respondents in this category excluded looking up information in records or on a calendar), or least committed if they agreed to three or fewer or none of the behaviors.

Compared to the least committed, fully committed respondents were significantly more accurate in the reported number of visits to the pediatrician and to a specialist in terms of the most direct – but most stringent – measure of accuracy, exact matches with the records. Moderately committed respondents were also significantly more likely than the least committed to report an exact match for visits to a specialist. Additional analyses comparing reported values to the values in the medical records as well as the results of key covariates shows that the response behavior of fully and moderately committed respondents differed from that of the least committed in a number of ways. Some of these differences in response behavior appeared to improve accuracy while others had little or no effect on accuracy while still others, in some cases, may have harmed accuracy.

Fully and moderately committed respondents reported more visits to the pediatrician and a specialist, on average, relative to the least committed. Fully committed respondents were less likely to underreport visits to the pediatrician and to a specialist, while the moderately committed were also less likely to underreport visits to a specialist. These results support the idea that fully and moderately committed respondents were more accurate (more likely to report a match) for visits to the pediatrician and to a specialist because they tended to report more visits.

It is possible that the fully and moderately committed respondents used different recall strategies or that they used them more successfully. For example, if using a recall-and-extrapolate strategy, recalling a few events to estimate a rate and then projecting over the reference period (Tourangeau, Rips, & Rasinski, 2000), they may have tended to adjust their estimate upwards, which in the case of visits to the pediatrician and a

specialist may have helped them to not underreport visits, particularly with high numbers of visits. Or they may have used a recall-and-count strategy, recalling each event and counting the events to get the total. This may be less likely since this strategy typically leads to underreporting (Conrad, Brown, & Cashman, 1998). However, fully and moderately committed respondents may have put in more effort, checked outside records, thereby more successfully retrieving visits and underreporting less.

For visits to a specialist – commitment at each level improved accuracy compared to the control group as the number of visits increased. This suggests that as the task became harder, commitment had a more pronounced effect – even marginally so for the least committed, keeping in mind that these respondents started out with a low level of accuracy even when the number of visits is small, and that the number of cases with a high number of visits is relatively small. It is possible that this effect was observed for visits to a specialist and not for visits to the pediatrician because higher numbers of visits occur more frequently for visits to a specialist which may have made it more likely to detect this effect. It is also possible that the request for commitment was particularly motivating for parents of children requiring multiple specialist visits and who engage with the University of Michigan Health System routinely.

Looking at the date of the most recent visits, compared to the least committed, fully committed respondents were more likely to report a date matching the date in the records for visits to the pediatrician and a specialist and to report a full date (day, month, and year) for the most recent visit to the pediatrician, a specialist and the ER. Interaction effects between fully committed and the number of days since the date of the last visit in the records suggest that the effect of commitment for fully committed respondents, increased as the task became more difficult.

Fully committed respondents provided higher quality data based on several other measures. They had significantly less item nonresponse overall and to the medical visit and date questions specifically, and provided significantly more mentions and longer responses to open-ended questions. While the fully committed were marginally more likely to straightline and acquiesce, contrary to expectation, this may well have had more to do with true differences between the fully committed and least committed on the topic of the questions than with the effect of commitment level: they may have actually been

more satisfied with their child's pediatrician than were less committed respondents because respondents who were satisfied with the care their child receives were more likely to fully commit to a survey request from a University of Michigan Health System pediatrician, one of the Principle Investigators for the survey. Results for socially desirable reporting were also not in the expected direction but this also may have had to do with true differences in parenting style between the fully and least committed, than lack of disclosure: it is possible that the kind of respondent who will fully commit to conscientious responding is also the kind of parent who is particularly conscientious about taking steps to improve their child's health. Further, while all respondents in the commitment group were more likely to break-off than in the control group, those who did not break-off took significantly more time to complete the interview. This was particularly pronounced among the fully committed, who took the longest amount of time, on average, to respond.

While reporting accuracy improved for the fully and moderately committed as the number of visits increased, there is also evidence that these respondents were less accurate when there were zero visits in the medical records for visits to a specialist. Compared to the control group, fully committed respondents were significantly less likely to report a matching number of visits to a specialist when there were zero visits in the records and moderately committed respondents were marginally less likely. There is also evidence that the fully committed, also compared to the control group, overreported the relatively infrequent ER and hospital visits. This may represent a possible downside to commitment; respondents may feel compelled to report visits that did not actually occur, at least during the study's reference period, so as not to appear that they were not taking the task seriously or erring on the side of reporting as opposed to potentially underreporting.

While the fully committed performed significantly better on a range of measures compared to the least committed, they did not perform significantly better than the control group on many measures. It is therefore perhaps not surprising that, as shown in the previous chapter, there were few significant differences when comparing commitment overall to the control group. The fully committed would have needed to perform significantly better to make up for the response behavior of the least committed, which

was significantly weaker than that of the control group on nearly every measure, and presumably overwhelmed the gains made by the fully, and in some cases moderately, committed respondents.

As noted in the previous chapter, a key difference between the current study and previous studies examining the effect of commitment was offering a list of response behaviors to which respondents could either commit to or not. The assumption was that respondents would overwhelmingly agree to all of the expected behaviors and that it would strengthen the treatment. In the original studies by Cannell and his associates (Miller & Cannell, 1982; Oksenberg et al., 1977b, 1977a) very few sample members refused the commitment request in interviewer-administered (face-to-face and telephone) data collection. This was also the case in Conrad et al.'s (in press) and Vannette's (2016) web-based studies. But unlike the binary commitment request in the earlier studies, the commitment request in the current study may have implied to respondents that they could be selective about which behaviors they would commit to. Did offering a list of response behaviors which respondents could either commit to or not give license to those who committed to the least number of behaviors not to work very hard? Or would these respondents have behaved similarly (just as badly) if the request had been binary? It does seem plausible that indicating a low level of commitment may, in the minds of some respondents, excuse them from having to follow through on behaviors they did not commit to. However, a future study including a simple "commit/do not commit" treatment group would be needed to answer this definitively.

The sizeable number of respondents who did not fully commit may be related to the web mode and the lack of social presence of an interviewer to not only increase the number of respondents who would commit fully but also motivate them to adhere to the behaviors to which they had committed. Or it may have to do with attitudes toward survey participation and effort fifty or more years after the original studies.[5]

On the whole, data quality is not improved on most measures without excluding responses from the least committed respondents. In a few cases, commitment overall, may harm quality and this can be localized to those who take their pact most seriously

[5] The recent web studies involved paid volunteer samples (Conrad, Couper, Tourangeau, & Zhang, in press; Vannette, 2016) so may not be comparable to the current, probability web sample.

(i.e. being fully committed) – e.g. overreporting when there are zero visits to report for visits to a specialist, overreporting relatively infrequent ER and hospital visits, and some increased break-offs. Further, it may give license to low commitment respondents to not try very hard – further research is needed to know the extent to which this happens as a result of asking respondents to commit to a list of response behaviors. Nonetheless, the downsides to commitment may be a worthwhile price to pay for the potential gains in data quality. Commitment may motivate a substantial proportion of respondents (63% in this study) to be as accurate and thoughtful as possible, potentially improving data quality on some measures. In particular, this study offers evidence that commitment has the greatest effect when the response task is the most difficult. There is little to no cost operationally in requesting a commitment, assuming sample members who do not commit are still included in data collection. To the extent that sample members do commit or do not commit fully, if the current approach, is used in the future it would make most sense to take commitment into account in the analyses.

References

Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology*, *12*, 389–437.

Conrad, F. G., Brown, N. R., & Cashman, E. R. (1998). Strategies for estimating behavioural frequency in survey interviews. *Memory*, *6*(4), 339–366.

Conrad, F. G., Couper, M. P., Tourangeau, R., & Zhang, C. (in press). Reducing speeding in web surveys by providing immediate feedback. *Cited with Permission.*

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Miller, P. V., & Cannell, C. F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, *46*(2), 250–269.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977a). Effects of commitment to being a good respondent on interview performance. In *Experiments in interviewing techniques: Field*

experiments in health reporting, 1971-1977 (pp. 74–108). Department of Health, Education and Welfare.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977b). The effects of instructions, commitment and feedback on reporting in personal interviews. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 133–199). Department of Health, Education and Welfare.

Stata Statistical Software. (2015). (Version Release 14). College Station, TX: StataCorp LP.

Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.

Vannette, D. L. (2016). The Effects of Asking Attention Filter Questions: Evidence from a 14 Country Survey Experiment. Presented at the Annual Conference of the American Association for Public Opinion Research, Austin, TX.

Chapter 4: The Effect of Feedback in the Context of Commitment in a Survey of Parents about their Child's Health and Healthcare (Study 2)

4.1. Introduction

Charles Cannell and his associates at the University of Michigan pioneered the application of commitment and feedback in interviewer-administered surveys (Cannell, Marquis, & Laurent, 1977; Cannell, Oksenberg, & Converse, 1977; Miller & Cannell, 1982; Oksenberg, Vinokur, & Cannell, 1977). Seeing complementary roles for commitment and feedback, they believed that informing respondents on the survey's goals and the expected response process via commitment (and instructions), followed by feedback on the adequacy of the respondent's performance on the response process, could reinforce good respondent performance (e.g. taking time, expending effort, providing the expected level of detail) and improve data quality (for detailed discussion, see Cannell et al. (1981)). Recent research has explored the effect of feedback or "interventions" triggered by undesirable respondent behavior in web surveys such as responding too quickly (Conrad, Couper, Tourangeau, & Zhang, in press) and skipping questions (DeRouvray and Couper 2002) as well as increasing answers to open-ended questions (Holland & Christian, 2009). Findings from studies examining the effect of interactive feedback messages in web surveys suggest that interventions can be reasonably successful in affecting the targeted behavior (e.g., less speeding and reduced item nonresponse).

This chapter details the design, methods, and additional effects of feedback in the context of commitment in a survey of parents about their child's health and healthcare. The feedback treatment was nested within the principal treatment of commitment, in which respondents were asked to commit to certain desired response behaviors. The experimental design is discussed in more detail below. The current study examines the effect of feedback on reducing several targeted behaviors including responding too quickly or "speeding", reporting incomplete dates,

and skipping response fields in response to open-ended questions. It further assesses the effect of providing feedback in response to these undesired behaviors on improving measures of data quality including response accuracy to medical visit and date questions, reducing straightlining and acquiescence, and increasing responses to open-ended questions, and whether feedback improves data quality beyond the commitment treatment alone.

4.2. Background

Feedback plays an essential role in everyday conversation. According to conversation researchers, feedback consists of continuers or assessments provided by the listener to indicate comprehension or interest. Also referred to as "back-channels", feedback can take different forms such as a non-lexical gestures (e.g. head nodding), phrases (e.g. "uh hmm", "I see."), or more substantive utterances ("I understand what you mean") (Schegloff, 1991). Feedback can reveal either negative evidence about the success of the communication – mishearing or misunderstanding – or positive evidence – acceptance or understanding. We rely on, and are thus highly attuned to, incremental feedback as evidence and for help in reaching the goal of mutual understanding or "grounding" in conversation (Clark & Brennan, 1991, p. 128). The role and importance of feedback extends to dialogue with and across different media, including computer-interfaces, as discussed further below.

Survey interviews have been described as a "conversation with a purpose" (Kahn and Cannell (1957) cited in Maynard and Schaeffer (2006)) and research indicates that respondents draw on conversational norms (Grice, 1975) when responding to surveys (Schwarz, 2007). It therefore follows that respondents are likely to be attuned to and draw on interviewer feedback for cues during the question and answering process.

Studies by Vinokur, Oksenberg, and Cannell (1977) and Oksenberg et al. (1977) examined the effect of substantive feedback "tailored" to the expected response process for different types of questions. What they referred to as "tailored feedback" depends on the respondent's performance on the response process, and indicates adequate or poor performance to the respondent. Educational theory and research about feedback and effective learning supports predicted benefits for tailored feedback.

In education, feedback is considered information provided by an agent (e.g. teacher, peer, book, parent, and experience) regarding aspects of one's performance or understanding. It occurs after instructions that seek to impart knowledge or skills. According to Winnie and Butler (1994), feedback provides specific information that a learner can use to confirm, add to, overwrite, tune, or restructure different kinds of information in memory such as domain-specific knowledge, beliefs about self and tasks, or cognitive tactics and strategies. In a review of educational research on feedback, Hattie and Timperley (2007) found that feedback information about the processes underlying a task validates for the learner whether the task was carried out correctly and draws the learner's attention to the relationship between a task strategy and the probability of a successful performance thereby improving future use of successful task strategies. They also found that the effectiveness of feedback is enhanced by clear instructions about expectations and goals and should be task-related and repeated (2007).

Applied to a survey interview, education theory and research suggests that providing instructions on goals and the expected response process, followed by feedback on the adequacy of the respondent's performance on the response process, could reinforce good respondent performance (e.g. taking time, expending effort, providing the expected level of detail). Studies examining the effect of feedback in interviewer-administered surveys indicate that respondents are attuned to feedback and that it can help improve the quality of reported data. For example, feedback has been found to improve reporting on measures designed to be demanding in terms of memory and effort. Compared to a control condition, Vinokur, Oksenberg, and Cannell (1977) found that respondents in the feedback condition provided significantly more answers to 17 out of 24 open-ended items. Date precision was in the expected direction but not significant and no significant differences were observed for reporting undesirable information or checking outside sources. Oksenberg et al. (1977) found stronger results. Compared with the control condition, respondents receiving feedback increased the number of items reported to open-ended questions, checking outside sources, date precision of medical event dates, reported doctor visits, and activity curtailment.

The role of feedback transfers readily to the online environment. Feedback from computer interfaces grounds both our interactions with and through computers (Brennan,

1998). When interacting with a computer, computer feedback communicates information about the current state of the computer to the user. For example, feedback may acknowledge that the computer has registered a user's action or input, thereby satisfying the communication expectations that users have when engaging in a dialogue (Pérez-Quiñones & Sibert, 1996). Computer interfaces also provide substantive feedback that can mediate human interactions mediated through a computer. For example, feedback messages may be programmed to appear when there is an error in the information provided or an omission according to the expectations or requirements of the receiving individual or institution.[6] For example, a business or government entity may require that certain pieces of information be entered in order to create an account or to complete an online transaction.

In the context of surveys on the web, researchers have begun to examine the effect of intervention messages to curtail speeding (Conrad et al., in press; Zhang & Conrad, 2014), reduce item nonresponse (DeRouvray and Couper 2002), and increase answers to open-ended questions (Holland and Christian 2009). Findings from these studies suggest that interventions can be reasonably successful in affecting the targeted behavior (e.g., less speeding and reduced item nonresponse). For example, DeRouvray and Couper (2002) found that prompts substantially reduced the average skip rate (from 7.3% to 1.4%). Conrad et al. (in press) and Zhang and Conrad (in press) found that speeding prompts increased the amount of time spent on subsequent questions and also reduced straightlining (providing the same answer to multiple questions) in grid questions for respondents who were responsive to the speeding prompts (see also Kunz and & Fuchs, 2014a, 2014b). Holland and Christian (2009) found modest successful results in an experiment that probed respondents following two open-ended questions. 25% of respondents responded to probes asking if they had anything to add for the first of two open-ended questions, and only 9% did so for the second open-ended question. However, it is important to note that the probes were not contingent upon the nature of the respondent's initial response to the open-ended question. It is possible that respondents perceived these probes as generic and, therefore, more readily ignored them.

[6] See http://uxdesign.smashingmagazine.com/2009/01/19/12-useful-techniques-for-good-user-interface-design-in-web-applications/ for examples of web interface designs for feedback messages to users on the quality of the information they have entered.

Previous studies by Charles Cannell and colleagues included positive feedback ("Thanks for taking your time"; "Thanks. This is the exact kind of information we need") in addition to negative feedback ("You answered that quickly.") (Oksenberg et al., 1977; Vinokur et al., 1977). They were not able compare the relative effectiveness of positive and negative feedback. To avoid confounding the effects of positive and negative feedback, the feedback treatment examined in the current study focuses on negative feedback, as detailed below.

4.3. Study Design and Methods

As discussed in the previous chapters, the study employed a nested design with Feedback and CRCs nested within the principal treatment of Commitment, as follows:

Group 1: Commitment

Group 2: Commitment+Feedback

Group 3: Commitment+CRCs

Group 4: Control

Respondents assigned to the Commitment+Feedback group received question-specific instructions and feedback in response to their behavior in responding to different types of questions. Question-specific instructions were provided to clarify the objective of the question and advise the respondent on how they could go about producing complete and accurate answers for a set of items. Examples of instructions for the sets of questions and feedback phrases tailored to the respondent's response behavior are provided in Table 4.1 below. Screenshots with an example of each type of feedback are provided in Appendix 1.

Drawing on Conrad et al. (in press), speeding is defined as below a psychologically-based threshold as opposed to a relative one (i.e., responses that are faster than others), as defined by Greszki, Meyer, & Schoen (2015) and Malhotra (2008), for example. As Conrad et al. (in press) note, a true "speeding threshold" is likely to vary across people, based on cognitive ability, whether the respondent has ready answer, but it is technically difficult to distinguish between a legitimate fast response from a response that is fast due to speeding. Therefore, the same response time threshold is used for all respondents based on the estimated average reading speed of 300 milliseconds multiplied

by the number of words in the question (e.g. 3,000 msec. or 3 seconds for a 10-word question). While generic, Conrad et al. (in press) have found this threshold to reliably discriminate between more and less conscientious responses and respondents. The timing threshold was changed following a pilot study for one question, which was at the beginning of a set of questions using a five-point Likert-scale about satisfaction with care. The timing threshold was adjusted to account for the extra text that appeared with the first statement instructing respondents to indicate their level of agreement with the following statements.

Table 4.1. Instructions and Feedback

Question	Instructions	Response Behavior → Feedback
Medical visits e.g. During the past 12 months, how many times did [CHILD] see a primary care doctor? Date of most recent visit e.g. What was the date of [CHILD]'s most recent visit to a primary doctor?	For the next set of questions, we'd like you to be as exact as you can about the number of visits and dates of the most recent visits. To be the most accurate, you may need to take time to think carefully before you answer.	Quick response (e.g. faster than the established threshold) → "You seem to have responded very quickly. Please be sure you have given the question sufficient thought to provide an accurate answer. Year (yy) or month and year only (mm/yy) → "Could you be more precise about the date?"
Open-ended – health practices [three empty text fields will appear for each question] e.g. What type of foods, vitamins and/or supplements do you give [CHILD] to stay healthy?	For the next items, we'd like you to provide as much information as you can even things that may not seem important to you.	Any blank response fields out of the three response fields presented → "Are you certain you haven't missed anything?"
Satisfaction with care – e.g. The University of Michigan Health System delivers on its promises. [Five-point	For the next few questions, you may need to take time to think carefully before you answer to be the	Quick response (e.g. faster than the established threshold) → "You seem to have responded very quickly. Please be sure you have given the question sufficient thought

response scale - strongly agree to strongly disagree] Behavioral frequency – e.g. How often do you smoke in the presence of your child?	most accurate.	to provide an accurate answer.”

Respondents in conditions including CRCs were presented with the following text along with the medical visit questions:

> *“TO HELP YOU REMEMBER: It may help to think about the reason for the visits, the season of the year, who took [CHILD] to the visits, if you took off work, how you travelled to the visits, who else was with you and [CHILD], and if you waited long.”*

While CRCs were a factor in the study design, they had no effect on data quality and so are not discussed further.

4.4. Outcome Measures and Hypotheses

The outcome measures for the analysis reported below include the response behaviors targeted by the feedback interventions: (1) responding faster than the established threshold (discussed above), (2) incomplete dates that is a mm/yyyy or dd/mm versus dd/mm/yyyy, and (3) any blank response fields out of the three response fields presented for open-ended questions. In addition to the targeted response behaviors, the analysis examines the effect of feedback for the targeted response behaviors on increasing response accuracy and data quality based on the same indicators used in the previous chapters on the effect of commitment. For the purposes of this analysis, the likelihood of a reported number of visits that matches the records and the likelihood of a reported date that matches the records are used as indicators of accuracy. Other data quality indicators include straightlining, acquiescence, and the number of words and mentions to open-ended questions.

To summarize, the hypotheses for the study are as follows:

Hypothesis 1: Feedback about speeding will decrease speeding on subsequent questions.

Hypothesis 2: Feedback about incomplete dates will increase the reporting of complete dates in subsequent questions.

Hypothesis 3: Feedback about incomplete responses to open questions will reduce blank response fields for subsequent open-ended questions.

Hypothesis 4: Feedback about speeding will increase the accuracy of reported medical visits and feedback about incomplete dates will increase the accuracy of the date of the most recent visit.

Hypothesis 5: Feedback for speeding will reduce straightlining and acquiescence.

Hypothesis 6: Feedback will increase the number of words and mentions to open-ended questions.

4.5. Analytical Methods

To examine the effect of feedback on reducing speeding in subsequent medical visit questions, satisfaction, and behavioral frequency questions for which respondents could receive speeding-related feedback, logistic regression models are constructed with speeding or no speeding as the outcome variable and feedback as the independent variable. Binary indicators are also entered for commitment and CRCs (0=control/reference; 1=treatment) to control for possible effects of these treatments, as shown in an example for linear regression in Equation 4.1. Included in the analysis are respondents from any of the treatment groups who sped when responding to any of the previous question(s). Respondents in the Commitment+Feedback group would have received feedback upon responding too quickly, while respondents in the control group and the other treatment groups would not have received feedback. This model estimates the effect of receiving feedback for respondents in the Commitment+Feedback group on the probability of speeding on subsequent questions compared to speeders in the other groups without feedback. Similar regression models are fit to examine the effect of feedback on increasing the likelihood that a complete date was reported in subsequent

date questions and decreasing the likelihood that response fields were left blank in subsequent open-ended questions.

Equation 4.1. Regression Equation for Estimating the Effect of Feedback

$$\gamma = \beta_0 + \beta_1 + \beta_2 + \beta_3 + \varepsilon$$

γ = the dependent/outcome variable
β_0 = intercept
β_1 = Feedback
β_2 = Commitment
β_3 = CRCs

A similar approach is used to examine the effect of feedback on improving response accuracy to the reported number of visits and dates, reducing straightlining and acquiescence, and increasing the number of words and mentions to open-ended questions. Included in the analysis are respondents from any of the treatment groups who exhibited the targeted behavior in response to any of the previous question(s) in the set. For example, respondents were prompted for speeding for the medical visit questions, so respondents from any of the treatment groups who sped when answering the medical visit questions are included when assessing the effect of feedback on increasing reporting accuracy to these questions.

There is reason to believe that some respondents using the Chrome Internet browser for the survey may have been able to block feedback messages after receiving at least one feedback message, which may have weakened the feedback treatment. In some versions of the browser, Chrome adds a checkbox underneath the text in pop-up boxes after the first pop-up is displayed giving users the option to "Prevent this page from displaying future dialogs." Essentially, this feature enables respondents, who notice it and check the box, to block any further potential pop-ups, which would, in effect, eliminate future feedback messages. A screenshot in Appendix 2 shows an example. Testing by the author suggests that this feature may have only been available in older versions of Chrome and revealed no evidence that this type of feature exists in other web browsers including Safari, Firefox, and Internet Explorer, even when third-party pop-up blockers were installed.

Respondents in the Commitment+Feedback group were asked in a debriefing question if they were using Chrome and, if yes, whether they blocked feedback messages. Nineteen out of 131 respondents reported that they blocked feedback messages. Survey paradata, specifically the user agent string, which contains information about the respondents' web browser, confirms that all of the respondents who reported blocking feedback were using Chrome. To detect a possible effect of blocking feedback on the main effect of feedback, an additional regression model was estimated for each of the models estimating the effect of feedback for each of the outcome measures, controlling for those who reported blocking feedback. Results for these additional models are provided in Appendix 3. In several cases, there was an insufficient number of cases or amount of variation to estimate parameters for those who reported blocking feedback. Overall, there were very few instances where blocking feedback had a significant effect or changed the result for the main effect of feedback. These instances are noted in the results section below.

Because some respondents using Chrome may not have reported blocking feedback, even if they did, the same regression models were also fit with terms entered into the model for Chrome browser use and an interaction between Chrome browser use and feedback. A statistically significant coefficient for the interaction between Chrome and feedback would indicate that the effect of feedback on reducing undesirable behaviors is moderated (reduced) by Chrome versus other browsers. Results for these models are provided in Appendix 3. Overall, there were very few instances where Chrome use interacted with feedback significantly or changed the result for the main effect of feedback. These instances are noted in the results section below.

4.6. Results

4.6.1. The Effect of Feedback on Reducing Speeding

Respondents in the Commitment+Feedback group could receive feedback for responding faster than the established threshold for three sets of questions in the survey: questions asking about the number of medical visits of different types of their child had in the last 12 months, satisfaction questions related to the care their child received at their most recent visit to the pediatrician, and a series of behavioral frequency questions

(Hypothesis 1). We first examine the effect of feedback in reducing speeding for medical visit questions. This section also examines the effect of feedback on increasing the accuracy of reported medical visits (Hypothesis 4) and on reducing straightlining and acquiesces to the satisfaction questions (Hypothesis 5).

Medical Visit Questions

The number and percent of speeding instances by treatment group for the medical visit questions can be found in Table 4.2. Respondents in the control group, Commitment only, and Commitment+Feedback groups begin with a similar percent of speeding. Different timing thresholds were calculated for respondents in the Commitment+CRCs group because of the additional text containing the recall cues with the medical visit questions. The substantially higher percentage of speeding for respondents in this group suggests that many respondents receiving the CRC treatment were unlikely to have thoroughly read the text presented to them.

Table 4.2. Number and Percent of Speeding Instances by Treatment Group for Medical Visit Questions

	Control		Commitment only		Commitment + Feedback		Commitment + CRCs	
	n	%	n	%	n	%	n	%
Q1. Visits to the Pediatrician	24	6.5	19	5.3	24	6.0	138	38.2
Q2. Visits to a Specialist	152	41.4	121	33.9	126	31.4	228	63.2
Q3. Visits to the ER	221	66.2	200	56.0	233	58.1	316	87.5
Q4. Visits to the Hospital	75	20.4	66	18.5	58	14.5	325	90.0

Regression results for the likelihood of speeding for those who were speeding for the initial question, visits to the pediatrician, on the subsequent medical visit questions are presented in Table 4.3. While no significant effect for feedback is seen for the first two items, visits to a specialist or to the ER, perhaps after enough instances of feedback for speeding have been delivered, we see that feedback had a marginally significant effect on reducing speeding for the last question, visits to the hospital. This result for the effect of feedback is statistically significant ($p < .05$) when controlling for those who reported

blocking feedback as shown in Appendix 3, Table 1, supporting Hypothesis 1 for this item, at least.

Table 4.3. The Effect of Feedback for Speeding on Previous Items on the Probability of Speeding in Subsequent Questions

VARIABLES	Q2.Visits to a Specialist	Q3.Visits to the ER	Q4.Visits to the Hospital
Feedback	0.44	0.20	-0.37^
	(0.618)	(0.276)	(0.211)
Commitment	-0.80	-0.20	-0.04
	(0.631)	(0.265)	(0.201)
CRCs	1.30**	1.90***	3.64***
	(0.502)	(0.328)	(0.265)
Constant	0.69	1.10***	-0.97***
	(0.433)	(0.183)	(0.138)
Observations	205	687	1,097

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

How did feedback affect accuracy of responses to the medical visit questions? Table 4.4 presents regression results for the likelihood that the reported number of visits matches the records for visits to a specialist, the ER and the hospital. While feedback did not appear to have a significant effect on reducing speeding for visits to the ER, respondents receiving feedback were more likely to report a matching number of visits to the ER compared to speeding respondents in the other groups that did not receive feedback. This result lends partial support to Hypothesis 4.

Table 4.4. The Effect of Feedback on the Likelihood of a Matching Report to Subsequent Visit Questions

VARIABLES	Q1.Visits to a Specialist	Q2.Visits to the ER	Q3.Visits to the Hospital
Feedback	0.03	0.79*	0.54
	(0.722)	(0.381)	(0.466)
Commitment	-0.59	-0.52	-0.01
	(0.691)	(0.341)	(0.411)
CRCs	0.43	0.25	-0.48
	(0.561)	(0.291)	(0.360)
Constant	-0.29	2.03***	2.93***
	(0.441)	(0.251)	(0.285)
Observations	191	665	1,075

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Satisfaction Questions

As shown in the percentages of respondents speeding in response to a series of Likert-scale items about satisfaction with care shown in Table 4.5, close to half of the respondents in each of the treatment groups responded faster than the threshold for the first question (Q1. Understand). The percentage of respondents speeding in the Commitment+Feedback group drops sharply after this initial question but does not appear to go down in the other groups. This is reflected in the regression results shown in Table 4.6, which shows a positive and statistically significant effect of feedback on reducing speeding for all of the subsequent questions in the set, clearly supporting Hypothesis 1.

Table 4.5. Percent of Respondents Responding Below Speeding Threshold by Treatment Group for Satisfaction Questions

	Control		Commitment only		Commitment + Feedback		Commitment + CRCs	
	n	%	n	%	n	%	n	%
Q1. Understand	190	51.8	173	48.5	182	45.4	170	47.1
Q2. Listened	223	60.8	199	55.7	2	0.5	195	54.0
Q3. History	196	53.4	161	45.1	4	1.0	186	51.5
Q4. Respect	283	77.1	263	73.7	22	5.5	261	72.3
Q5. Time	125	34.1	113	31.7	2	0.5	106	29.4
Q6. Recommend	275	74.9	245	68.6	21	5.2	271	75.1

Table 4.6. The Effect of Feedback for Speeding on Previous Items on the Probability of Speeding in Subsequent Questions

VARIABLES	Q2.Listened	Q3.History	Q4.Respect	Q5.Time	Q6.Recommend
Feedback	-6.40***	-4.67***	-4.14***	-4.64***	-3.97***
	(1.011)	(0.590)	(0.285)	(1.006)	(0.289)
Commitment	-0.02	-0.03	-0.04	-0.12	0.09
	(0.216)	(0.160)	(0.216)	(0.143)	(0.176)
Constant	1.23***	0.60***	1.89***	-0.48***	1.42***
	(0.173)	(0.128)	(0.174)	(0.115)	(0.140)
Observations	715	922	991	1,104	1,115

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

While feedback appears to have had a strong effect on reducing speeding in the context of the Likert-scale satisfaction questions, it does not appear to have helped reduce acquiescence or straightlining as shown in Table 4.7. In fact, contrary to expectation,

feedback appears to have had a marginally positive effect on acquiescence. These results fail to support Hypothesis 5. The marginal increase in acquiescence may reflect a possible downside of feedback, which is to increase self-consciousness by making it salient that respondents are being evaluated (Henderlong & Lepper, 2002) thereby increasing socially desirable or acquiescent reporting. This is consistent with a study on interactive feedback by Zhang and Conrad (in press) in which respondents in each of the intervention conditions gave more socially desirable answers compared to the no-intervention condition.

Table 4.7. The Effect of Feedback on Reducing Acquiescence and Straightlining

VARIABLES	Acquiescence	Straightlining
Feedback	0.17^	0.08
	(0.096)	(0.164)
Commitment	-0.04	0.01
	(0.079)	(0.135)
Constant	4.46***	0.06
	(0.064)	(0.109)
Observations	1,159	1,159
R-squared	0.003	

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Behavioral Frequency Questions

Looking at the percentages of respondents speeding in response to a series behavioral frequency questions shown in Table 4.8, we see that respondents in the Commitment+Feedback group began with a smaller percentage of speeders in response to the first question. This group had a slightly higher percentage of speeders for Q2 and Q3, a higher percentage for Q4, and then smaller again for the remaining four questions Q6-Q9. This pattern is reflected in the regression results presented in Table 4.9, where we see a marginally greater chance of speeding after feedback for Q4 and significantly less chance of speeding for Q6-Q9. These results offer mixed support for Hypothesis 1.

Table 4.8. Percent of Respondents Speeding by Treatment Group for Behavioral Frequency Questions

	Control		Commitment only		Commitment + Feedback		Commitment + CRCs	
	n	%	n	%	n	%	n	%
Q1. Enough sleep	33	9.0	36	10.1	11	2.7	41	11.4
Q2. Exercise	8	2.2	12	3.4	34	8.5	15	4.2
Q3. TV time	23	6.3	24	6.7	37	9.2	21	5.8
Q4. Computer time	62	16.9	62	17.4	125	31.2	57	15.8
Q5. Family meal	12	3.3	14	3.9	23	5.7	14	3.9
Q6. Smoking	185	50.3	169	47.3	10	2.5	177	49.0
Q7. Others smoking	257	70.0	240	67.2	143	35.7	242	67.0
Q8. Raise voice	23	6.3	17	4.8	3	0.8	26	7.2
Q9. Spanking	214	58.3	199	55.7	1	0.25	216	59.8

Table 4.9. The Effect of Feedback for Speeding on Previous Items on the Probability of Speeding in Subsequent Questions

VARIABLES	Q2.Exercise	Q3.TV	Q4.Computer	Q5.Family meal	Q6.Smoking
Feedback	0.71	-0.20	0.60^	0.04	-4.20***
	(0.746)	(0.447)	(0.307)	(0.334)	(0.421)
Commitment	1.05	-0.03	-0.20	0.46	0.28
	(0.794)	(0.439)	(0.339)	(0.454)	(0.278)
Constant	-2.74***	-1.06**	-0.23	-2.48***	0.84***
	(0.730)	(0.367)	(0.279)	(0.393)	(0.222)
Observations	121	173	236	436	451

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

VARIABLES	Q7.Others smoking	Q8.Raise voice	Q9.Spanking
Feedback	-1.57***	-2.22**	-6.12***
	(0.215)	(0.728)	(1.007)
Commitment	0.14	0.01	0.04
	(0.239)	(0.269)	(0.156)
Constant	1.70***	-2.45***	0.72***
	(0.189)	(0.217)	(0.125)
Observations	772	1,040	1,041

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

It is possible that feedback could create a feeling of being monitored or a sense of social presence among respondents who receive one or more instances of feedback. This could inhibit socially undesirable responses or encourage respondents to provide more

socially desirable responses. One study found that respondents receiving feedback interventions for speeding reported fewer socially undesirable answers compared to a control condition (Conrad & Zhang, in press). Examining the effect of feedback on the behavioral frequency questions in this study, which are moderately sensitive and could be susceptible to socially desirable reporting, yielded few significant results. However, consistent with Conrad and Zhang (in press), for one item, respondents who had received at least one instance of feedback for a previous item reported significantly more days when all family members in the household had at least one meal together.

4.4.2. The Effect of Feedback on Responses to Date Questions

This section examines the effect of feedback on reducing the reporting of incomplete dates (Hypothesis 2) and on increasing the accuracy of the date of the most recent medical visits (Hypothesis 4). Table 4.10 shows similar percentages of incomplete dates for each of the treatment groups in response to (Q1.Last visit to the pediatrician). The percentage of respondents reporting an incomplete date in the Commitment+Feedback group then goes down for the next two questions – visits to a specialist and ER. The percentage of incomplete dates for the last visit to the hospital is lower in the Commitment+Feedback than the Commitment only and Commitment+CRCs groups but is the same as the control group. However, the number of cases for the last visit to the hospital is quite small. This pattern is reflected in the regression results shown in Table 4.11, which shows a positive and statistically significant effect of feedback on reducing incomplete dates for the last visit to a specialist and the ER, which supports Hypothesis 2. The effect of feedback for the last visit to the hospital is in the expected direction but not significant.

Table 4.10. Percent of Respondents Reporting an Incomplete Date by Treatment Group

	Control		Commitment only		Commitment + Feedback		Commitment + CRCs	
	n	%	n	%	n	%	n	%
Q1.Last visit to the Pediatrician	48	13.1	54	15.1	63	15.7	48	13.3
Q2.Last visit to a Specialist	47	12.8	47	13.2	33	8.2	44	12.2
Q3.Last visit to the ER	16	4.4	25	7.0	12	3.0	18	5.0
Q4.Last visit to the Hospital	3	0.8	6	1.7	3	0.8	7	1.9

Table 4.11. The Effect of Feedback for an Incomplete Date on Previous Items on the Probability of Reporting a Full Date for Subsequent Questions

VARIABLES	Q2.Last Visit to a Specialist	Q3.Last Visit to the ER	Q4.Last Visit to the Hospital
Feedback	2.83***	2.09**	1.79
	(0.546)	(0.667)	(1.118)
Commitment	-0.59	-0.71	-0.29
	(0.464)	(0.655)	(1.323)
Constant	-0.90*	-0.51	-1.10
	(0.358)	(0.516)	(1.155)
Observations	143	68	24

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

While feedback appears to have had a strong effect on reducing the reporting of incomplete dates, it does not appear to have helped increase the accuracy of reported dates in terms of a match with the date in the records, as shown in Table 4.12. This finding fails to support Hypothesis 4. There are too few matching reported dates for the date of the last visit to the ER and hospital to estimate the effect of feedback for these measures.

Table 4.12. The Effect of Feedback on the Likelihood of a Match to Subsequent Date Questions

VARIABLES	Last Visit to a Specialist
Feedback	-0.13
	(0.759)
Commitment	-1.48^
	(0.862)
Constant	0.56
	(0.627)
Observations	48

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

4.4.3. The Effect of Feedback on Responses to Open-ended Questions

This section examines the effect of feedback on reducing skipped response fields (Hypothesis 3) and on increasing the number of words and mentions (Hypothesis 6) in response to open-ended questions. As shown in Table 4.13, a smaller percentage of respondents in the Commitment+Feedback group left response fields blank in response to

the first open-ended question. This suggests that previous feedback for speeding and incomplete dates may be helping here because no feedback about answering open-ended questions fully had yet been given. The percentage goes up in response to the second and third question. However, these percentages are smaller in comparison to the other treatment groups. The regression results shown in Table 4.14 show a significant effect of feedback in reducing the probability of blank response fields to the two subsequent open-ended questions, which supports Hypothesis 3. Chrome browser usage was not significant for either question, but the effect of feedback loses significance when controlling for Chrome usage for Q2 ("Cutdown").

Not surprisingly, the significant effect of feedback in reducing missing response fields resulted in significantly more words and mentions in responses to the open-ended questions in a pooled analysis for word count and mentions, as shown in Table 4.15, lending support for Hypothesis 6.

Table 4.13. Percent of Respondents with Missing Response Fields by Treatment Group for Open-ended Questions

	Control		Commitment only		Commitment + Feedback		Commitment + CRCs	
	n	%	n	%	n	%	n	%
Q1.Foods, vitamins, or supplements	204	55.6	206	57.7	115	28.7	184	51.0
Q2.Cut down	322	87.4	298	83.5	278	69.3	316	87.5
Q3.Maintain child's health	312	85.0	301	84.3	240	59.9	298	82.5

Table 4.14. The Effect of Feedback for Missing Response Fields on the Probability of Blank Response Fields in Subsequent Questions

VARIABLES	Q2.Cut down	Q3.Improve child's health
Feedback	-1.07**	-1.27***
	(0.342)	(0.183)
Commitment	0.13	0.20
	(0.361)	(0.212)
Constant	2.69***	1.97***
	(0.287)	(0.167)
Observations	709	1,266

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Table 4.15. The Effect of Feedback on the Number of Words and Mentions to Open-ended Questions

VARIABLES	Word count	Number of mentions
Feedback	3.85***	1.23***
	(0.968)	(0.158)
Commitment	0.74	0.13
	(0.997)	(0.163)
Constant	12.59***	4.36***
	(0.808)	(0.132)
Observations	1,409	1,409
R-squared	0.015	0.050

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

4.4.4. Interaction between Feedback and Commitment Level

Results in the previous chapter showed a number of significant differences, within treatment and compared to the control group, based on how respondents responded to the request to commit to the desired response behaviors that comprised the commitment treatment in this study. Overall, respondents who agreed to adhere to all of the requested behaviors, the fully committed, and in some cases, the moderately committed, those who agreed to four out of five of the requested behaviors (overwhelmingly, respondents in this category excluded looking up information in records or on a calendar), provided more accurate and better data quality responses based on a number of indicators than those who agreed to three or fewer or none of the behaviors. One might expect that the fully committed, in particular, and the moderately committed, compared to the least committed might be more attuned to and responsive to feedback in an effort to uphold their commitment to providing complete and accurate answers. Interaction terms for fully committed and moderately committed and feedback were added to each of the models above. There was one marginally significant result for fully committed and feedback for the likelihood of reporting a matching number of visits with the records for visits to the ER. The results for the other measures (e.g. speeding, incomplete dates) were mostly in the expected direction but not significant. The lack of significant results is likely due to limited statistical power because of the limited number of cases assigned to the feedback group, which were then further broken down by commitment level group.

4.5. Discussion

Overall, providing feedback produced a number of significant effects in reducing unwanted respondent behaviors including speeding, reporting incomplete dates and leaving response fields blank to open-ended questions. Specifically, feedback significantly reduced speeding for all items in a battery of Likert-scale items, and for four out of eight behavioral frequency questions. While no significant effect for feedback was seen for the first two medical visit items, perhaps after enough prior instances of speeding, feedback had a marginal effect on reducing speeding for the last question, visits to the hospital. This result is significant when controlling for those who blocked feedback. These results are consistent with Conrad et al. (in press) who found that speeding prompts increased the amount of time spent on subsequent questions and Conrad and Zhang (in press) who also found that feedback significantly curtailed speeding.

Feedback also significantly reduced the likelihood that an incomplete date was reported for two date questions and was in the expected direction for the third, most likely due to insufficient statistical power. We also saw a significant effect of feedback on reducing the probability of blank response fields to subsequent open-ended questions. The effect of feedback in the context of open-ended questions in this study appears to be more pronounced than the effect of non-contingent probes examined by Christian and Holland (2009).

While there is good evidence from this study that feedback reduces unwanted respondent behaviors, the results are mixed when it comes to improvements in reporting accuracy and quality. There is some evidence that feedback improved data quality based on some measures. For example, respondents receiving feedback were more likely to report a matching number of visits to the ER compared to speeding respondents in the other groups that did not receive the feedback treatment. Even though the number of speeding respondents was not significantly reduced for visits to the ER specifically, it is possible that those who were speeding on the previous question subsequently slowed down and provided a more accurate response to this question. It is also possible that some respondents who were speeding were still able to provide accurate answers. For example, it may have been easy for those with zero visits to report this answer very quickly.

Further, feedback resulted in significantly more words and mentions to open-ended questions.

On the other hand, no significant effect was seen for feedback on increasing the accuracy of a reported date (the last date to a specialist) and there were too few matching reported dates for the other date questions (date of the last visit to the ER and hospital) to estimate the effect of feedback for these items. This study also found no effect of feedback on reducing satisficing behaviors such as straightlining, unlike Zhang and Conrad (in press) and Kunz and Fuchs (2014a; 2014b), or acquiescence. In fact, feedback was associated with marginally more acquiesce. The latter may reflect a possible downside of feedback, which is to increase self-consciousness by making it salient that respondents are being evaluated (Henderlong & Lepper, 2002) thereby increasing socially desirable reporting. This is consistent with a study on interactive feedback by Zhang and Conrad (in press) in which respondents in each of the intervention conditions gave more socially desirable answers compared to the no-intervention condition.

Why don't we see a stronger effect for feedback in improving accuracy and data quality? It is possible that speed, in particular, is not always associated with inaccurate or poor quality responding. For example, some respondents may be able to read and respond to a question very quickly and accurately, particularly, when the response task is easy. For example, some respondents may have needed less time than the established threshold to comprehend and respond to the question about the number of visits to the ER in the last 12 months, since questions about pediatrician and specialist visits, which follow the same format, immediately proceeded it, and they may have known how many visits their child to the ER in the last 12 months without having to search their memory or having to look up the information. Alternatively, some respondents who are speeding may be unable or unwilling to improve the accuracy of their response by taking more time. These respondents may in fact be speeding because they know that they are not in a position to, or are unwilling to, take the time to provide an accurate answer.

Along this vein, providing feedback for incomplete dates increased the likelihood that a complete date was reported for subsequent questions, but this was not associated with increased accuracy for reported dates. It may be that respondents who reported a partial date did so because they did not know the complete date. The hope was that

prompting them to report a full date would encourage respondents to look up the information or give it more thought, which may have helped the respondent come up with a more complete and accurate date. But it also may have encouraged some respondents to guess or to enter a complete, yet randomly selected, date.

Perhaps this study did not replicate past findings for feedback in reducing straightlining and showed a marginal increase in acquiescence because of the nature of the questions. This may be a case where straightlining or agreement to all items in a battery of questions may reflect the respondent's true beliefs. In this case, respondents may truly feel satisfied with the care their child received from their pediatrician across the range of items included in the battery.

In this study, the feedback treatment was nested within the principal treatment of commitment. It is interesting to note that while feedback had a significant effect on reducing speeding, reporting incomplete dates and leaving open-ended response fields blank, there were no main effects for commitment in reducing these unwanted respondent behaviors. While statistical power was limited, there is also some evidence that feedback had more of an effect among the fully and moderately committed respondents, suggesting that the effect of feedback may be further enhanced for those who committed to all or most of the requested response behaviors. This is consistent with Cannell et al.'s (1981) original notion of commitment (and instructions) and feedback as complementary and reinforcing, particularly among those with a high level of commitment.

References

Brennan, S. E. (1998). The grounding problem in conversations with and through computers. *Social and Cognitive Psychological Approaches to Interpersonal Communication*, 201–225.

Cannell, C. F., Marquis, K. H., & Laurent, A. (1977). *A summary of studies of interviewing methodology*. US Government Printing Office, Washington, DC 20402.

Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology*, *12*, 389–437.

Cannell, C. F., Oksenberg, L., & Converse, J. M. (1977). Striving for response accuracy: experiments in new interviewing techniques. *Journal of Marketing Research*, 306–315.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, *13*(1991), 127–149.

Conrad, F. G., Couper, M. P., Tourangeau, R., & Zhang, C. (in press). Reducing speeding in web surveys by providing immediate feedback. *Cited with Permission.*

Conrad, F. G., & Zhang, C. (in press). Intervening to Reduce Satisficing in Web Surveys: Evidence from two experiments. *Social Science Computer Review*.

Greszki, R., Meyer, M., & Schoen, H. (2015). Exploring the Effects of Removing "Too Fast" Responses and Respondents from Web Surveys. *Public Opinion Quarterly*, *79*(2), 471–503.

Grice, H. P. (1975). Logic and conversation. *1975*, 41–58.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112.

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: a review and synthesis. *Psychological Bulletin*, *128*(5), 774.

Holland, J. L., & Christian, L. M. (2009). The influence of topic interest and interactive probing on responses to open-ended questions in web surveys. *Social Science Computer Review*, *27*(2), 196–212.

Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public Opinion Quarterly*, *72*(5), 914–934.

Maynard, D. W., & Schaeffer, N. C. (2006). Standardization-in-interaction: The survey interview. In P. Drew, G. Raymond, & D. Weinberg (Eds.), *Talk and interaction in social research methods* (pp. 9–27). London, U.K.: Sage Publications, Inc.

Miller, P. V., & Cannell, C. F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, *46*(2), 250–269.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977). The effects of instructions, commitment and feedback on reporting in personal interviews. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 133–199). Department of Health, Education and Welfare.

Pérez-Quiñones, M., & Sibert, J. L. (1996). A collaborative model of feedback in human-computer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 316–323). Retrieved from http://dl.acm.org/citation.cfm?id=238535

Schegloff, E. A. (1991). Conversation analysis and socially shared cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 150–171). Washington D.C.: American Psychological Association.

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology*, *21*(2), 277–287.

Vinokur, A., Oksenberg, L., & Cannell, C. F. (1977). Effects of feedback and reinforcement on the report of health information. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 109–132). Department of Health, Education and Welfare.

Zhang, C., & Conrad, F. G. (2014). Investigation of speeding in web surveys: tendency to speed and its association with straightlining. *Survey Research Methods*, *8*(2), 127–135.

Appendix 4.1. Feedback Examples

Figure 4.1. Speeding Feedback Example

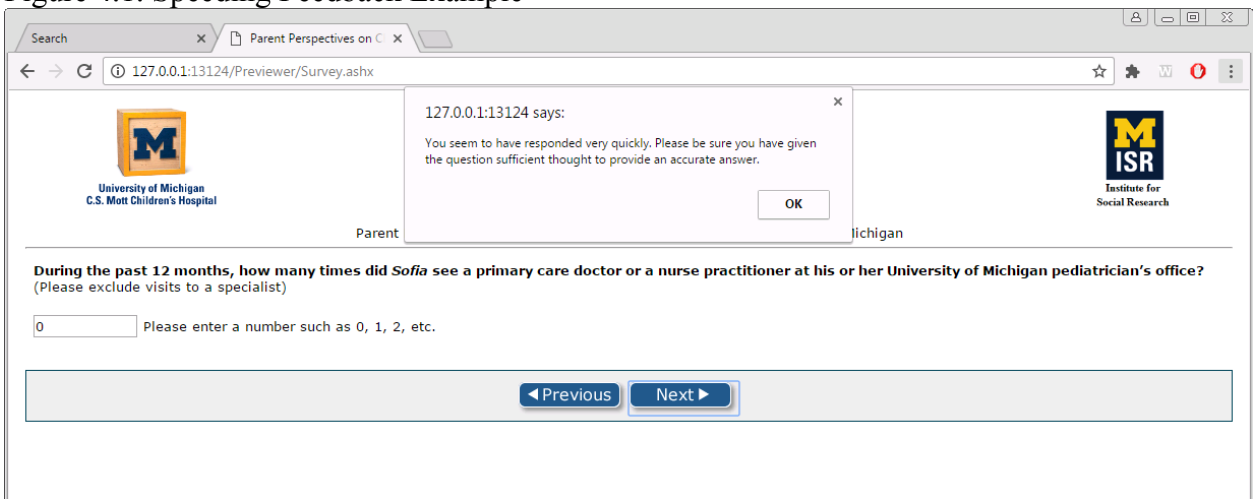


Figure 4.2. Incomplete Date Feedback Example

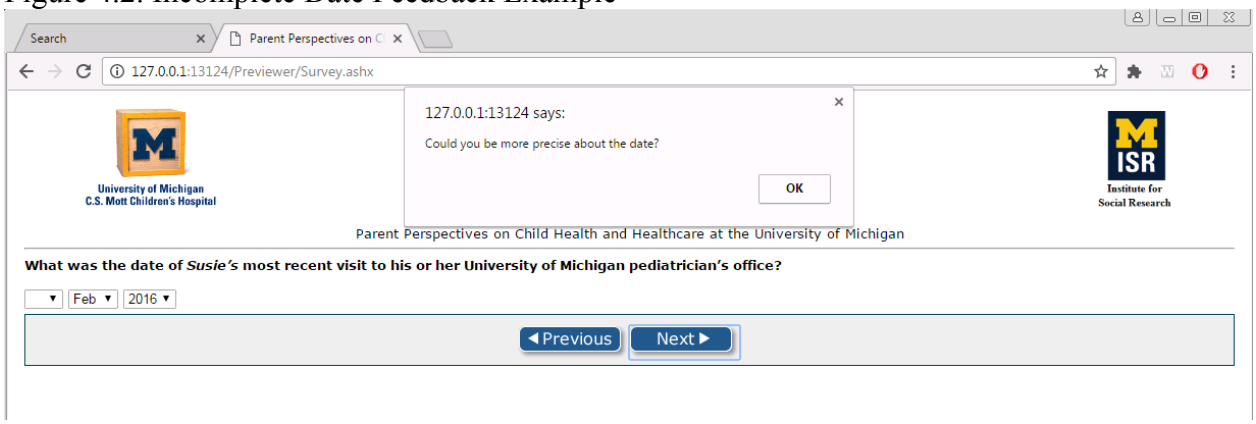
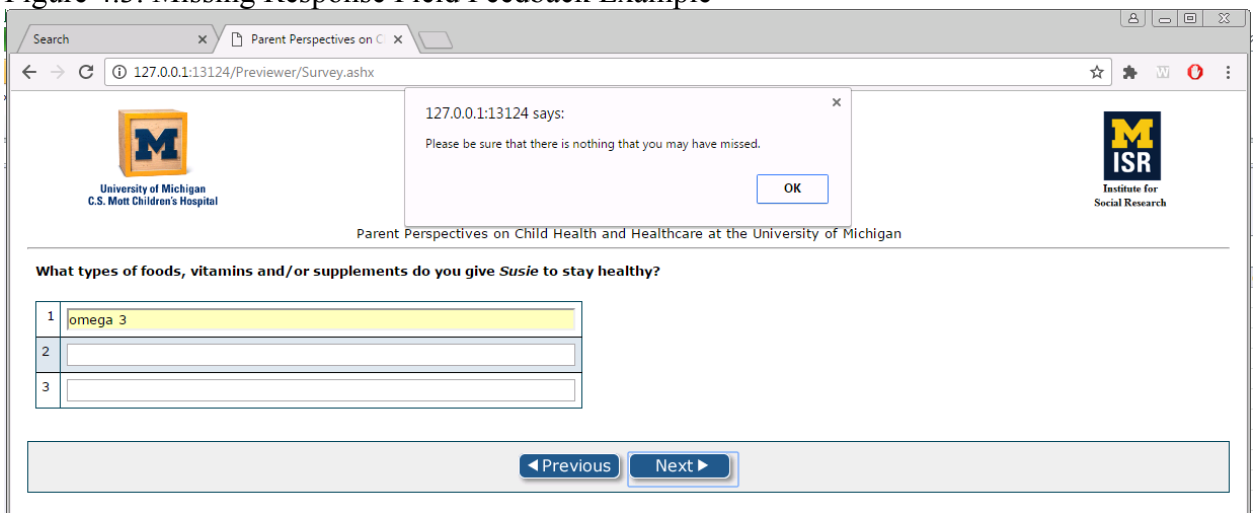
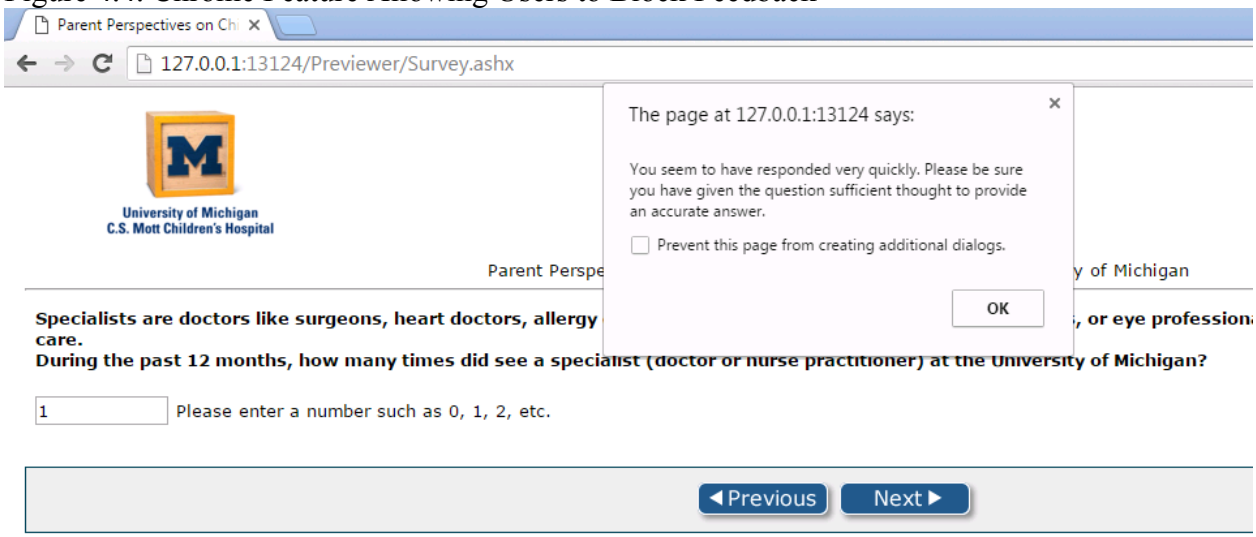


Figure 4.3. Missing Response Field Feedback Example



Appendix 4.2. Chrome Feature Allowing Users to Block Feedback

Figure 4.4. Chrome Feature Allowing Users to Block Feedback



Appendix 4.3. Regression Models Estimating the Effect of Feedback and Controlling for Respondents Who Reported Blocking Feedback and Chrome Use

Please note that the tables below include results for the models shown in the results section (1) along with the models controlling for those who reported blocking feedback (2) and Chrome browser use (3) for comparison purposes.

Table 4.16. The Effect of Feedback for Speeding on Previous Items in the Probability of Speeding in Subsequent Medical Visit Questions

	Visits to a Specialist			Visits to the ER		
VARIABLES	1	2	3	1	2	3
Feedback	0.44	0.47	-0.21	0.20	0.16	-0.18
	(0.618)	(0.632)	(0.826)	(0.276)	(0.279)	(0.350)
Commitment	-0.80	-0.80	-0.85	-0.20	-0.20	-0.20
	(0.631)	(0.631)	(0.637)	(0.265)	(0.265)	(0.265)
CRCs	1.30**	1.30**	1.33**	1.90***	1.90***	1.90***
	(0.502)	(0.502)	(0.505)	(0.328)	(0.328)	(0.328)
Chrome			-0.22			-0.01
			(0.345)			(0.237)
Feedback * Chrome			1.14			0.81^
			(0.932)			(0.470)
Blocked feedback		-0.37			0.89	
		(1.479)			(1.088)	
Constant	0.69	0.69	0.83^	1.10***	1.10***	1.10***
	(0.433)	(0.433)	(0.487)	(0.183)	(0.183)	(0.215)
Observations	205	205	205	687	687	687

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Visits to the Hospital		
VARIABLES	1	2	3
Feedback	-0.37^	-0.49*	-0.24
	(0.211)	(0.219)	(0.295)
Commitment	-0.04	-0.04	-0.05
	(0.201)	(0.201)	(0.202)
CRCs	3.64***	3.64***	3.70***
	(0.265)	(0.265)	(0.268)
Chrome			0.44*
			(0.187)
Feedback * Chrome			-0.23
			(0.360)
Blocked feedback		1.37**	
		(0.513)	
Constant	-0.97***	-0.97***	-1.20***
	(0.138)	(0.138)	(0.174)
Observations	1,097	1,097	1,097

Table 4.17. The Effect of Feedback on the Likelihood of a Match – Controlling for Blocking Feedback and Chrome Use

	Visits to a Specialist			Visits to the ER		
VARIABLES	1	2	3	1	2	3
Feedback	0.03	-0.08	0.15	0.79*	0.72^	0.26
	(0.722)	(0.749)	(1.004)	(0.381)	(0.381)	(0.516)
Commitment	-0.59	-0.59	-0.48	-0.52	-0.52	-0.50
	(0.691)	(0.691)	(0.699)	(0.341)	(0.341)	(0.344)
CRCs	0.43	0.43	0.37	0.25	0.25	0.25
	(0.561)	(0.561)	(0.565)	(0.291)	(0.291)	(0.294)
Chrome			0.45			-0.80**
			(0.322)			(0.258)
Feedback * Chrome			-0.35			0.92
			(1.079)			(0.659)
Blocked feedback		0.96			-	
		(1.509)				
Constant	-0.29	-0.29	-0.59	2.03***	2.03***	2.47***
	(0.441)	(0.441)	(0.495)	(0.251)	(0.251)	(0.300)
Observations	191	191	191	665	657	665

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Visits to the Hospital		
VARIABLES	1	2	3
Feedback	0.54	0.47	-0.02
	(0.466)	(0.466)	(0.512)
Commitment	-0.01	-0.01	-0.02
	(0.411)	(0.411)	(0.411)
CRCs	-0.48	-0.48	-0.47
	(0.360)	(0.360)	(0.361)
Chrome			0.17
			(0.291)
Feedback * Chrome			1.95^
			(1.115)
Blocked feedback			
Constant	2.93***	2.93***	2.85***
	(0.285)	(0.285)	(0.316)
Observations	1,075	1,058	1,075

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Table 4.18. The Effect of Feedback for Speeding on Previous Items in the Probability of Speeding in Subsequent Satisfaction Questions*

	History			Respect		
VARIABLES	1	2	3	1	2	3
Feedback	-4.67***	-5.01***	-5.01***	-4.14***	-4.20***	-3.99***
	(0.590)	(0.718)	(1.014)	(0.285)	(0.299)	(0.391)
Commitment	-0.03	-0.03	-0.03	-0.04	-0.04	-0.04
	(0.160)	(0.160)	(0.160)	(0.216)	(0.216)	(0.216)
Chrome			0.02			0.02
			(0.154)			(0.207)
Feedback * Chrome			0.56			-0.30
			(1.243)			(0.551)
Blocked feedback		2.04			0.74	
		(1.264)			(0.820)	
Constant	0.60***	0.60***	0.59***	1.89***	1.89***	1.88***
	(0.128)	(0.128)	(0.151)	(0.174)	(0.174)	(0.204)
Observations	922	922	922	991	991	991

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Recommend		
VARIABLES	1	2	3
Feedback	-3.97***	-4.05***	-4.17***
	(0.289)	(0.307)	(0.444)
Commitment	0.09	0.09	0.09
	(0.176)	(0.176)	(0.176)
Chrome			0.01
			(0.169)
Feedback * Chrome			0.37
			(0.574)
Blocked feedback		0.93	
		(0.826)	
Constant	1.42***	1.42***	1.41***
	(0.140)	(0.140)	(0.166)
Observations	1,115	1,115	1,115

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

* Please note that parameter estimates for Feedback*Chrome and blocking feedback could not be estimated due to an insufficient number of cases and the results are therefore not shown for two of the satisfaction items: Listened and Time.

Table 4.19. The Effect of Feedback on Reducing Straightlining and Acquiescence

	Acquiescence			Straightlining		
VARIABLES	1	2	3	1	2	3
Feedback	0.17^	0.14	0.24^	0.08	-0.00	0.32
	(0.096)	(0.098)	(0.135)	(0.164)	(0.168)	(0.232)
Commitment	-0.04	-0.04	-0.04	0.01	0.01	0.01
	(0.079)	(0.079)	(0.079)	(0.135)	(0.135)	(0.136)
Chrome			-0.00			0.33*
			(0.075)			(0.129)
Feedback * Chrome			-0.15			-0.47
			(0.183)			(0.315)
Blocked feedback		0.53			1.54^	
		(0.348)			(0.789)	
Constant	4.46***	4.46***	4.46***	0.06	0.06	-0.11
	(0.064)	(0.064)	(0.075)	(0.109)	(0.109)	(0.128)
Observations	1,159	1,159	1,159	1,159	1,159	1,159
R-squared	0.003	0.005	0.003			

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Table 4.20. The Effect of Feedback for Speeding on Previous Items in the Probability of Speeding in Subsequent Behavioral Frequency Questions*

	Smoking			Others smoking		
VARIABLES	1	2	3	1	2	3
Feedback	-4.20***	-4.28***	-4.08***	-1.57***	-1.58***	-1.03***
	(0.421)	(0.449)	(0.561)	(0.215)	(0.220)	(0.305)
Commitment	0.28	0.28	0.29	0.14	0.14	0.13
	(0.278)	(0.278)	(0.279)	(0.239)	(0.239)	(0.240)
Chrome			-0.12			0.21
			(0.268)			(0.232)
Feedback * Chrome			-0.27			-1.00*
			(0.825)			(0.398)
Blocked feedback		0.77			0.07	
		(1.125)			(0.608)	
Constant	0.84***	0.84***	0.89***	1.70***	1.70***	1.60***
	(0.222)	(0.222)	(0.257)	(0.189)	(0.189)	(0.218)
Observations	451	451	451	772	772	772

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

	Raise voice		
VARIABLES	1	2	3
Feedback	-2.22**	-2.16**	-2.09*
	(0.728)	(0.728)	(1.028)
Commitment	0.01	0.01	0.01
	(0.269)	(0.269)	(0.269)
Chrome			0.23

			(0.260)
Feedback * Chrome			-0.25
			(1.444)
Blocked feedback			
Constant	-2.45***	-2.45***	-2.57***
	(0.217)	(0.217)	(0.261)
Observations	1,040	1,027	1,040

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

* Please note that parameter estimates for Feedback*Chrome and blocking feedback could not be estimated due to an insufficient number of cases and the results are therefore not shown for one of the behavioral frequency items: Spanked.

Table 4.21. The Effect of Feedback for an Incomplete Date on the Likelihood of Reporting a Complete Date for Subsequent Questions*

	Last Visit to a Specialist		Last Visit to the ER		Last Visit to the Hospital	
VARIABLES	1	3	1	3	1	3
Feedback	2.83***	4.01***	2.09**	4.02**	1.79	18.42
	(0.546)	(0.856)	(0.667)	(1.228)	(1.118)	(3,252.026)
Commitment	-0.59	-0.67	-0.71	-0.93	-0.29	-0.25
	(0.464)	(0.473)	(0.655)	(0.696)	(1.323)	(1.332)
Chrome		0.68		0.95		0.39
		(0.469)		(0.686)		(1.135)
Feedback * Chrome		-2.41*		-3.44*		-17.95
		(1.093)		(1.471)		(3,252.026)
Constant	-0.90*	-1.21**	-0.51	-0.89	-1.10	-1.30
	(0.358)	(0.428)	(0.516)	(0.603)	(1.155)	(1.316)
Observations	143	143	68	68	24	24

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

* Please note that parameter estimates for blocking feedback could not be estimated due to an insufficient number of cases and the results are therefore not shown for any of the date items

Table 4.22. The Effect of Feedback on the Likelihood of a Matching Date to Subsequent Date Questions*

	Date of Last Visit to a Specialist	
VARIABLES	1	3
Feedback	-0.13	-0.14
	(0.759)	(0.975)
Commitment	-1.48^	-1.47^
	(0.862)	(0.863)
Chrome		0.20

		(0.882)
Feedback * Chrome		0.28
		(1.362)
Constant	0.56	0.43
	(0.627)	(0.835)
Observations	48	48

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

* Please note that a parameter estimate for blocking feedback could not be estimated due to an insufficient number of cases and the results are therefore not shown for this item

Table 4.23. The Effect of Feedback for Missing Response Fields on the Likelihood of Missing Response Fields in Subsequent Questions*

	Cut down		Improve child's health		
VARIABLES	1	3	1	2	3
Feedback	-1.15**	-0.72	-1.23***	-1.23***	-1.08***
	(0.353)	(0.497)	(0.185)	(0.187)	(0.259)
Commitment	0.08	0.08	0.10	0.10	0.10
	(0.390)	(0.390)	(0.220)	(0.220)	(0.220)
Chrome		0.14			-0.11
		(0.376)			(0.212)
Feedback * Chrome		-0.87			-0.28
		(0.653)			(0.335)
Blocked feedback				-0.10	
				(0.615)	
Constant	2.83***	2.76***	2.03***	2.03***	2.09***
	(0.310)	(0.356)	(0.175)	(0.175)	(0.208)
Observations	678	678	1,211	1,211	1,211

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

* Please note that a parameter estimate for blocking feedback could not be estimated due to an insufficient number of cases and the results are therefore not shown for Cutdown.

Table 4.24. The Effect of Feedback for Missing Response Fields on the Number of Mentions and Words to Open-ended Questions

	Number of mentions			Word count		
VARIABLES	1	2	3	1	2	3
Feedback	1.52***	1.54***	1.50***	6.37***	6.29***	6.49***
	(0.216)	(0.218)	(0.284)	(1.093)	(1.105)	(1.437)
Commitment	0.01	0.01	0.01	0.25	0.25	0.25
	(0.177)	(0.177)	(0.177)	(0.894)	(0.895)	(0.895)
Chrome			0.26			0.87
			(0.169)			(0.853)
Feedback *						
Chrome			0.09			-0.12
			(0.417)			(2.113)

Blocked feedback		-0.59			2.83	
		(1.172)			(5.925)	
Constant	3.04***	3.04***	2.91***	8.62***	8.62***	8.19***
	(0.143)	(0.143)	(0.165)	(0.721)	(0.721)	(0.837)
Observations	675	675	675	675	675	675
R-squared	0.075	0.075	0.079	0.054	0.054	0.055

Standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

Conclusion

The inability or unwillingness of respondents to provide accurate information presents a serious threat to the quality of survey measurement (Groves et al., 2009; Oksenberg, Vinokur, & Cannell, 1977b; Tourangeau, Rips, & Rasinski, 2000). This is of particular concern in web surveys because respondents may have little motivation to expend sufficient effort to provide accurate responses without the presence of an interviewer. In their pioneering work on survey research methods, Charles Cannell and colleagues at the University of Michigan in the 1970s and 80s demonstrated the promise of directly asking respondents to commit to providing complete and accurate answers. While promising, these studies were conducted decades ago, in interviewer administered modes, with limited measures of data quality. This dissertation consists of two experimental studies investigating the effectiveness of commitment as well as automatic feedback in promoting better data quality in online questionnaires.

The first study measures the effect of commitment – "yes" or "no" – in an online labor force survey. The experiment was embedded in a survey conducted by the Institute for Employment Research (Institut für Arbeitsmarkt und Berurfsforschung (IAB)) in Germany. The second study measures the effect of asking respondents to commit to engaging in five specific response behaviors that seem likely promote data quality including reading all of the questions carefully; trying to be as precise as possible; looking up information in records or on a calendar, if needed; providing as much information as possible, and; answering honestly. The survey was of the parents of child patients at University of Michigan (UM) Health System. The second study also examined the effect of providing feedback in response to behaviors that are associated with reduced data quality such as speeding, reporting an incomplete date, and leaving open-ended response fields blank. Cannell and his associates viewed feedback as complementary to

commitment (see Cannell, Miller, and Oksenberg (1981)) and examined it separately (Vinokur, Oksenberg, & Cannell, 1977) and along with commitment (Oksenberg et al., 1977b), finding stronger effects for feedback when combined with commitment. For both studies, administrative records were used to verify the accuracy of certain self-reported responses, in contrast to the indirect quality measures used in earlier evaluations of commitment.

Study 1 produced a number of promising effects for asking respondents to commit to providing complete, accurate, and honest answers. For the effect of commitment overall, regardless of whether respondents agreed or did not agree to the commitment, there was a significant reduction in item nonresponse overall and for reported income, specifically. Commitment group respondents were also more likely to report having checked records, were more accurate with their reported income in terms of reduced absolute error between reported income and income in the administrative records compared to the control group, and were less likely to report a rounded answer to the income question. There was also some evidence of more disclosure of socially undesirable answers in the commitment group.

While the response of the treatment group to the request for commitment in Study 1 was mostly positive, 4.8% did not agree and could be considered "not committed". The response behavior of the not committed was strikingly different from those who committed. While committed respondents had significantly less item nonresponse, not committed respondents had significantly more. There was no significant effect for the not committed on the accuracy of reported income, while committed respondents were significantly more accurate in terms of the absolute difference between reported income and income in the administrative records. Further, not committed respondents were also less likely to report checking records, consent to having their survey responses linked with administrative records or to being recontacted for a follow-up interview. In a couple of cases, the poorer response behavior of the not committed affected the results for commitment overall compared to the control group. First, there were marginally more break-offs for commitment overall compared to control, which could be attributed to the significantly higher proportion of break-offs for the not committed. Second, there was no effect in reducing straightlining for commitment overall because gains made from the

committed, who were marginally less likely to straightline, were cancelled out by marginally more straightlining among the not committed. However, the relatively poor response behavior of the not committed – because they were relatively few in number – did not diminish the overall effect of commitment for most measures.

In Study 2, asking respondents to commit to adhering to certain desired response behaviors produced more mixed results. For most measures, there was no overall effect of commitment. However, there was an effect of commitment when the response task was most difficult: commitment group respondents were significantly more likely to report a matching date of the last visit to the pediatrician and maintained a higher level of accuracy as the number of visits increased and as the time since the date of the most recent visit to a specialist increased. Checking records probably helped and, while some respondents clearly balked at the request to check information in records in Study 2, overall, significantly more respondents in the commitment group reported checking records. On the other hand, commitment group respondents were also more likely to overreport visits when there were zero visits in the records. Further, commitment group respondents had a significantly higher proportion of skipped items for medical visit and date questions and broke off at a higher rate than the control group. A possible downside to commitment is that respondents may feel compelled to report visits that did not actually occur, at least during the study's reference period, so as not to appear that they were not taking the task seriously or erring on the side of reporting as opposed to potentially underreporting. They may also opt to skip questions or not to complete the interview altogether if they feel that they are not able to provide complete and accurate answers, as requested in the commitment statement.

Yet, profound differences were observed in Study 2 based on how respondents responded to the request to commit to the desired response behaviors. Respondents were considered fully committed if they agreed to adhere to all of the requested behaviors, moderately committed if they agreed to four out of five (overwhelmingly, respondents in this category excluded looking up information in records or on a calendar), or least committed if they agreed to three or fewer or none of the behaviors. Compared to the least committed, fully committed respondents were significantly more accurate in the reported number of visits to the pediatrician and to a specialist in terms of the most direct

– but most stringent – measure of accuracy, exact matches with the records. They had significantly less item nonresponse overall and to the medical visit and date questions specifically, and provided significantly more mentions and longer responses to open-ended questions. While the fully committed were marginally more likely to straightline and acquiesce, contrary to expectation, this may well have had more to do with true differences between the fully committed and least committed on the topic of the questions than with the effect of commitment level: they may have actually been more satisfied with their child's pediatrician than were less committed respondents because respondents who were satisfied with the care their receives were more likely to fully commit. Results for socially desirable reporting were also not in the expected direction but this also may have had to do with true differences in parenting style between the fully and least committed, than lack of disclosure: it is possible that the kind of respondent who will fully commit to conscientious responding is also the kind of parent who is particularly conscientious about taking steps to improve their child's health. Further, while all respondents in the commitment group were more likely to break-off than in the control group, those who did not break-off took significantly more time to complete the interview. This was particularly pronounced among the fully committed, who took the longest amount of time, on average, to respond.

However, in a few cases, commitment overall harmed quality and can be localized to those who took the commitment pact most seriously (i.e. the fully committed) – e.g. overreporting when there are zero visits to report for visits to a specialist, overreporting relatively infrequent ER and hospital visits, and some increased break-offs.

Results for the overall effect of commitment were more positive in Study 1 than in Study 2. It is possible that the effect of commitment was less pronounced in Study 2 because respondents for this study, the parents of child patients who were sufficiently motivated to log on to complete the survey, may have been a fairly conscientious or committed population to begin with – many of whom would have fully committed if assigned to the treatment group – leaving little room for commitment to make a difference (i.e. a ceiling effect). Or is it possible that the decomposed commitment was less effective by giving license to those who did not commit fully to not try as hard? Unlike the binary commitment request in the earlier studies and in Study 1, the

commitment request in the Study 2 was actually comprised of asking respondents to commit to engaging in five specific response behaviors that seem likely promote data quality. Respondents could commit to some but not all of these practices so commitment as implemented here was continuous or graded. It is possible that indicating anything but full commitment may, in the minds of some respondents, excuse them from having to follow through on behaviors they did not commit to or not to try very hard. Further research is needed to know the extent to which this happens as a result of asking respondents to commit to a list of response behaviors.

Results from the current research indicate that, in contrast to how Cannell and others have conceptualized commitment, it may not have a blanket effect, which improves response quality for all respondents across the board. In a break from previous studies on commitment, both Study 1 and Study 2 saw more people refuse to commit or to fully commit, in the case of Study 2. In the original studies by Cannell and his associates (Miller & Cannell, 1982; Oksenberg et al., 1977b; Oksenberg, Vinokur, & Cannell, 1977a) very few sample members refused the commitment request in interviewer-administered (face-to-face and telephone) data collection. This was also the case in Conrad et al.'s (in press) and Vannette's (2016) web-based studies. The sizeable number of respondents who did not commit or fully commit may be related to the web mode and the lack of social presence of an interviewer to not only increase the number of respondents who would commit but also motivate them to adhere to the behaviors to which they had committed. Or it may have to do with attitudes toward survey participation and effort fifty or more years after the original studies.[7]

As opposed to being binary, results from Study 2 also demonstrate that respondent commitment falls on a range, with some respondents willing to commit to all of the expected behaviors associated with high quality responses while others are willing to commit to most or only a few of the behaviors. In particular, checking records appears to be a behavior that a substantial proportion of respondents find to be above and beyond what they are willing to commit to in responding to a web survey: 27% of respondents

[7] The recent web studies involved paid volunteer samples (Conrad, Couper, Tourangeau, & Zhang, in press; Vannette, 2016) so may not be comparable to the current, probability web sample.

who committed to all the behaviors except for one, 98% (295 out of 302) agreed to all except for checking records.

Further, results from Study 2 showed that there may be some potential downsides associated with commitment even for respondents who are fully committed. Fully committed respondents were more likely to overreport visits when there were zero in the records for visits to a specialist and relatively infrequent ER and hospital visits. Along with the least committed, they were also more likely to break-off and to skip questions compared to the control group. These results suggest that an unintended effect of commitment may be to compel those who take the pact the most seriously to report events that did not actually occur (at least during the study's reference period) so as not to appear that they were not taking the task seriously. Commitment may also encourage respondents to skip questions or discontinue the survey altogether rather than give inadequate answers, regardless of their professed commitment level.

Nonetheless, though perhaps more nuanced than in previous studies, results from the current studies show promise for commitment. Commitment in both studies motivated a substantial proportion of respondents to be as accurate and thoughtful as possible, and even to check records. Even though a number of respondents clearly balked at the request to check records – significantly more committed respondents in Study 1 and those who fully committed in Study 2 reported that they checked records. Thus, on balance, survey practitioners are likely to find the potential trade-offs involved with commitment a worthwhile price to pay, given the little to no operational cost of commitment, provided the data from those who do not commit is not discarded – for the benefits to data quality – particularly when the response task is difficult.

In Study 2, feedback was nested within the principal treatment of commitment for one group of respondents. Providing feedback produced a number of significant effects in reducing unwanted respondent behaviors including speeding, reporting incomplete dates and leaving response fields blank to open-ended questions. Feedback also improved data quality based on some measures. For example, respondents receiving feedback were more likely to report a matching number of visits to the ER compared to speeding respondents in the other groups that did not receive the feedback treatment and resulted in significantly more words and mentions to open-ended questions. However, this study

found no effect of feedback on reducing satisficing behaviors such as straightlining, unlike Zhang and Conrad (in press) and Kunz and Fuchs (2014a; 2014b) and found marginally more acquiescence. Consistent with Zhang and Conrad (in press) who found more socially desirable reporting among respondents receiving feedback interventions, the latter may reflect a possible downside of feedback, which is to increase self-consciousness by making it salient that respondents are being evaluated (Henderlong & Lepper, 2002) thereby increasing socially desirable reporting.

While statistical power was limited, there is also some evidence that feedback had more of an effect among the fully and moderately committed respondents, suggesting that the effect of feedback may be further enhanced for those who committed to all or most of the requested response behaviors. This is consistent with Cannell et al.'s (1981) original notion of commitment (and instructions) and feedback as complementary and reinforcing, particularly among those with a high level of commitment.

References

Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology*, *12*, 389–437.

Conrad, F. G., Couper, M. P., Tourangeau, R., & Zhang, C. (in press). Reducing speeding in web surveys by providing immediate feedback. *Cited with Permission.*

Conrad, F. G., & Zhang, C. (in press). Intervening to Reduce Satisficing in Web Surveys: Evidence from two experiments. *Social Science Computer Review*.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. Hoboken, New Jersey: John Wiley & Sons Inc.

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: a review and synthesis. *Psychological Bulletin*, *128*(5), 774.

Miller, P. V., & Cannell, C. F. (1982). A study of experimental techniques for telephone interviewing. *Public Opinion Quarterly*, *46*(2), 250–269.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977a). Effects of commitment to being a good respondent on interview performance. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 74–108). Department of Health, Education and Welfare.

Oksenberg, L., Vinokur, A., & Cannell, C. F. (1977b). The effects of instructions, commitment and feedback on reporting in personal interviews. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 133–199). Department of Health, Education and Welfare.

Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.

Vannette, D. L. (2016). The Effects of Asking Attention Filter Questions: Evidence from a 14 Country Survey Experiment. Presented at the Annual Conference of the American Association for Public Opinion Research, Austin, TX.

Vinokur, A., Oksenberg, L., & Cannell, C. F. (1977). Effects of feedback and reinforcement on the report of health information. In *Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977* (pp. 109–132). Department of Health, Education and Welfare.