It's About Time: Examining the Effect of Interviewer-Quoted Survey Completion Time Estimates on Survey Efficiency

Christopher D. Ward[1], **Bess Welch**[2], **Amie Conley**[3], **Philip J. Smith**[4], and **Stacie Greby**[5]

[1]**Institution:** NORC at the University of Chicago

[2]**Institution:** NORC at the University of Chicago

[3]**Institution:** NORC at the University of Chicago

[4]**Institution:** Centers for Disease Control and Prevention

[5]**Institution:** Centers for Disease Control and Prevention

Abstract

Declining response rates may introduce bias into survey results and increase costs. Two national surveys, the National Immunization Survey (NIS) and the NIS-Teen, were used to study the impact of survey length, as stated by the interviewer, and inclusion of a topic of interest to respondents on response rates. The two studies included comparisons of the standard survey instruments to revised, condensed instruments. The NIS study also included variations of the standard survey with sections considered of interest for parental respondents, the Parental Concerns Module (PCM), which contained questions about parents' thoughts and beliefs about vaccinations. The outcomes of interest were differences in the response rates and resulting survey costs in each of the study conditions. The shortened instruments resulted in higher response rates compared to both the standard instruments and the instruments including the PCM and reduced the overall time needed to complete an interview. Based on these results, the NIS and NIS-Teen questionnaires were both shortened.

Background

Response rates for telephone surveys have been declining for years. De Leeuw and de Heer (2002), Curtin et al. (2005), Galea and Tracy (2007) and Groves (2006) have described this in multiple studies spanning a range of survey topics and countries, and respondents' perception of interview burden has been found to be associated with nonresponse (Hansen 2007). Declining response rates can directly increase the costs for administering a survey and raise concerns about potential bias (Groves and Couper 1998).

Research has shown that increasing response rates, even by small amounts, can yield significant cost savings in data collection (Groves and Couper 1998). Beyond the standard data collection modifications designed to yield better response rates – shortening the questionnaire, use of incentives, longer field periods, advance letters, and increasing the number of callbacks – what else can be done? Respondent engagement with the survey topic has been found to be related to survey completion (Groves et al. 2000). Groves et al. (2004) found that when respondents are interested in the topic at hand, they are more likely to

respond, thereby leading to an increase in response rates. Two experiments based on the NIS and NIS-Teen surveys were designed to reduce nonresponse by testing a shorter survey instruments and varying the inclusion and placement of an interesting survey topic, the Parental Concerns Module (PCM). The PCM included questions on parents' perceptions and attitudes about vaccines. NIS interviewers previously reported respondents found PCM questions interesting. A primary goal of the studies was to determine the extent to which shorter surveys could yield higher response rates and lower costs.

Methods

The National Immunization Survey (NIS) has been conducted by the Centers for Disease Control and Prevention (CDC) since 1994 to estimate vaccination coverage for routine childhood vaccinations. The NIS-Teen has been conducted since 2006 as a complementary add-on survey to the NIS to estimate vaccination coverage for adolescents. The NIS-Teen employs the same screening mechanism as the NIS. Households eligible for both surveys complete the NIS first followed by the NIS-Teen. Both surveys use random digit dialing (RDD) telephone survey methodology to identify adults in households containing age-eligible children and adolescents. Detailed methods for both surveys have been described previously (CDC 2016; Smith et al. 2001, 2005). The NIS landline response rates have been stable over time, whereas the NIS-Teen landline rates have decreased modestly over time (see Tables G.1 and G.2 in CDC 2016). However, since cellular telephones have increasingly replaced landlines as the main source of communication, a cell frame was added to the NIS and NIS-Teen in 2011. Response rates for the cell frame are lower than that of the landline frame. The declining response rates have led to increases in survey costs and reduced efficiency.

Versions of the 2011 NIS and the 2013 NIS-Teen survey instruments were adapted as control conditions in each respective study. Both studies assessed the impact of survey length, as stated by the interviewer, on the overall response rate by comparing the control versions with condensed versions. The NIS study also assessed the effect of inclusion and placement of the PCM on overall response rates.

Sample Selection and Data Collection Protocol

In contrast to the regular NIS and NIS-Teen surveys, which both employ RDD designs, the studies employed proprietary list samples of households likely to contain children. For the NIS study, sample sizes were set to identify a total of 1,950 eligible households per study condition to detect a 3.0 percentage-point difference in the interview completion rate. For the NIS-Teen study, sample sizes were set to identify 1,230 households per study condition to achieve sufficient power to detect a 3.8 percentage-point difference in the interview completion rate.

To minimize the introduction of interviewer effects, new interviewers were hired for each study. Protocols used for interviewer training, incentive eligibility, and call scheduling rules for both studies were identical to those employed in regular NIS and NIS-Teen surveys. The NIS study was conducted from August 23, 2011, through December 4, 2011. The NIS-Teen study was conducted from June 27, 2013, through August 31, 2013.

Instrument Development and Study Design

NIS Study

Figure 1 outlines the structure of the four study conditions. The study design consisted of an experiment with a control condition, which was adapted from the third-quarter NIS instrument, and three treatment groups. Only households with a child 19 to 35 months of age were eligible for the survey.

Control Condition—The NIS questionnaire began with an eligibility screener. After screening, respondents were asked whether they had a child's vaccination record (shot card) available. Respondents with a shot card were given a 20-minute time estimate for survey completion and asked a series of questions about their child's vaccination history (Section A). Respondents without a shot card were given a 15-minute time estimate and asked a shorter set of vaccination questions (Section B). The survey included demographic questions (Section C), provider consent (Section D), health insurance questions (Section E), and the PCM. The control condition was treated as a single condition whether the respondent answered Section A or Section B.

Treatment Conditions—The three study conditions excluded Section A and included a shortened version of Section B containing questions about influenza vaccinations. The study conditions also varied placement and inclusion of the engaging topic (PCM). The Early Engagement condition placed the PCM directly after the shortened Section B; the Late Engagement condition included the shortened Section B and then mirrored the Control condition order and kept the PCM at the end; and the Short condition included the shortened Section B and excluded the PCM entirely. The stated time estimates for the Early and Late conditions were 15 minutes, while the time estimate for the Short condition was 10 minutes.

NIS-Teen Study

Figure 2 outlines the structure of the four NIS-Teen study conditions. Only those households with at least one child between 13 and 17 years of age were eligible for the survey.

Control Condition—The questionnaire began with a screener to determine eligibility, followed by up to four interview sections. After screening, respondents were asked whether they had a child's vaccination record (shot card) available. Those with a shot card were asked vaccination history questions in Section A and given a 20-minute time estimate for survey completion. Respondents without a shot card were routed to Section B, which contained a shorter set of vaccination questions, and were given a 15-minute time estimate. As with the NIS control condition, this was treated as a single condition whether the respondent answered Section A or Section B. All other sections of the instrument were identical for all respondents in the control condition.

Treatment Conditions—Three treatment conditions were included in the NIS-Teen study. All three treatment conditions excluded Section A. The three treatment conditions varied the length of Section B and the order of questions therein. The Long-Reversed condition contained the Section B control-condition questions but reordered them, so questions about

more recent adolescent vaccinations were asked first. The Medium-Reversed condition contained an abbreviated version of Section B in the same order as the Long-Reversed condition. Finally, the Short condition removed almost all Section B questions, asking NIS-Teen respondents to report on influenza vaccinations and abbreviated questions on tetanus, meningitis, and Human papillomavirus vaccinations. The stated time estimate for the Long-Reversed condition was 15 minutes, the Medium-Reversed condition was 12 minutes, and the Short condition was 10 minutes.

The hypothesis was examined using several kinds of response rates. Resolution rate was defined as the proportion of listed sample telephone numbers whose status as working or non-working could be determined. The screener completion rate was the proportion of households that screened as eligible or ineligible for the study. Completion of the interview was the proportion of households reaching the end of Section C: Demographics in each survey. The overall Council of American Survey Research Organizations (CASRO) response rate, American Association for Public Opinion Research (AAPOR) RR3 (AAPOR 2016), is the number of completed interviews divided by the number of eligible units in the sample with a proportional allocation assumption. Pairwise tests of proportions were conducted between the control condition and each of the treatment conditions. Chi-squared tests were completed to determine whether key performance rates varied by study condition. Finally, a comparison was made of the hours per completed interview (HPC) by condition to determine the extent to which each of the experimental conditions could yield lower costs. Weights were not produced for these studies, so only unweighted sample sizes were reported.

Results

Overall Response Rates

In the NIS study, 6,138 interviews were completed: 1,471 in the Control (488 completed Section A and 983 completed Section B); 1,525 in the Early Engagement; 1,531 in the Late Engagement; and 1,611 in the Short condition. In the NIS-Teen study, 1,257 interviews were completed in the Control (472 completed Section A and 785 completed Section B) 1,316 in the Long-Reversed; 1,413 in the Medium-Reversed and 1,550 in the Short condition for a total of 5,536 interviews. Both studies achieved sufficient completed interviews to meet the power requirements to allow for testing of differences in the interview completion rates by study condition.

The NIS study interview completion rates revealed that lower interviewer-quoted time estimates and shorter instruments correlated with higher interview completion rates (Table 1). The Short condition, with a 10-minute time estimate, had the highest interview completion rate, 86.38 percent. This rate was significantly higher than the rates for the other three study conditions. The Late and Early Engagement conditions, both with a 15-minute time estimate, had interview completion rates of 81.83 percent and 80.22 percent, respectively. The Control condition interview completion rate of 75.05 percent was significantly lower than that of the other three study conditions in pairwise comparisons. No pairwise comparison of resolution rate or screener completion rate between the Control condition and each of the remaining conditions was significant. This is an expected result,

because respondents in each experimental condition had the same experience until the time estimate was stated at the end of the screener.

Inclusion and placement of the PCM were manipulated to test the impact of an engaging topic on response rates. We compared retention between the Early and Late Engagement conditions, which varied only by the location of the PCM. The Early Engagement condition had an interview completion rate of 80.22, whereas the Late Engagement condition had an interview completion rate of 81.83. These rates were not significantly different from one another ($\alpha = 0.05$, not reported in table).

The NIS-Teen Study interview completion rates also indicated that lower interviewer-quoted time estimates and shorter instruments were related to higher interview completion rates (Table 2). The Short condition, with a 10-minute time estimate, had the highest interview completion rate at 82.1 percent. This rate was significantly higher than the rate for the Control condition (71.8 percent), but not the Medium condition (78.6 percent, $\alpha = 0.05$, not reported in table). The final interview completion rate of the Medium condition, with a 12-minute time estimate, was also significantly higher (78.6 percent) than that of the Control condition. The HPC was lowest for the Short condition (23.8 percent less than the Control) compared to the other conditions, and the Medium condition was also lower than the Control condition (12.8 percent less than the control), though not as low as that of the Short condition.

Discussion

Response rates varied by interviewer-quoted time estimate. In both studies, the shortest interview had the highest response rate, and the longest interview had the lowest response rate. Including an engaging topic, as operationalized in the NIS study, did not appear to increase the rate of interview completion. Placing the PCM early in the NIS instrument did not affect response rates. One possible reason that higher interview completion rates were not observed among early-PCM respondents is only respondents who were sufficiently cooperative reached the PCM, so any variation in completion rates we observed may be reduced by the baseline level of cooperation necessary to reach that portion of the survey. It was observed, however, that the length of the interview had a greater impact on likelihood of completion than did the inclusion of an engaging topic earlier in the interview.

Both experiments were subject to limitations. The studies employed list frames, which may limit the generalizability of these studies to the NIS, the NIS-Teen, or other RDD studies. Respondents on the list frames may be systematically different from those on the NIS frame or on other national RDD frames. Although the NIS study used the PCM as an engaging topic, there was not an objective assessment of whether the PCM was perceived as engaging. Previous NIS interviewers reported that respondents liked the PCM and found it interesting, but this interest may not have been enough to affect respondent behavior and impact response rates. There was a potential for change in nonresponse bias of the experimental and control conditions that was not assessed in this study. No evidence of change in overall bias was detected for the NIS before and after the short questionnaire was implemented in 2012 (Singleton et al. 2015).

These studies found that reducing the length of the reported time in the survey instruments by just a few minutes can increase response rates, thereby decreasing data collection costs. The Short conditions were instituted in the NIS in 2012 and in the NIS-Teen in 2014. The interview completion rates for both surveys increased with the implementation of the shorter instruments and survey administration costs decreased. These modifications have translated into significant benefits in cost savings and while maintaining the robustness of the surveys.

References

American Association for Public Opinion Research. Standard efinitions: final dispositions of case codes and outcome rates for surveys. 9. American Association for Public Opinion Research; Deerfield, IL: 2016. Available at http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Centers for Disease Control and Prevention. National Immunization Survey: a user's guide for the 2015 public-use data file. 2016. Available at https://www.cdc.gov/vaccines/imz-managers/nis/downloads/nis-puf15-dug.pdf

Curtin R, Presser S, Singer E. Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly. 2005; 69(1):87–98.

De Leeuw, E., de Heer, W. Trends in household survey nonresponse: a longitudinal and international perspective. In: Groves, RM.Dillman, DA.Eltinge, JL., Little, RJA., editors. Survey nonresponse. Wiley; New York: 2002.

Galea S, Tracy M. Participation rates in epidemiologic studies. Annals of Epidemiology. 2007; 17(9): 643–653. [PubMed: 17553702]

Groves RM. Nonresponse rates and nonresponse error in household surveys. Public Opinion Quarterly. 2006; 70(5):646–675.

Groves, RM., Couper, MP. Nonresponse in household interview surveys. Wiley; New York: 1998. How survey design features affect participation.

Groves RM, Presser S, Dipko S. The role of topic of interest in survey participation decisions. Public Opinion Quarterly. 2004; 68(1):2–31.

Groves RM, Singer E, Corning A. Leverage-saliency theory of survey participation: description and an illustration. Public Opinion Quarterly. 2000; 64(3):299–308. [PubMed: 11114270]

Hansen K. The effects of incentives, interview length, and interviewer characteristics on response rates in a CATI-study. International Journal of Public Opinion Research. 2007; 19(1):112–121.

Singleton, J., Elam-Evans, L., Khare, M., Yankey, D., Hill, H., Wolter, K., Pineau, VV. Estimating change in telephone survey bias in an era of declining response rates and transition to wireless telephones – evidence from the National Immunization Survey (NIS), 1995–2013; Presented at the Annual American Association of Public Opinion Research Conference; Hollywood, FL. 2015.

Smith P, Hoaglin D, Battaglia M, Khare M, Barker L. Statistical methodology of the National Immunization Survey, 1994–2002. Vital and Health Statistics. 2005; 2(138)

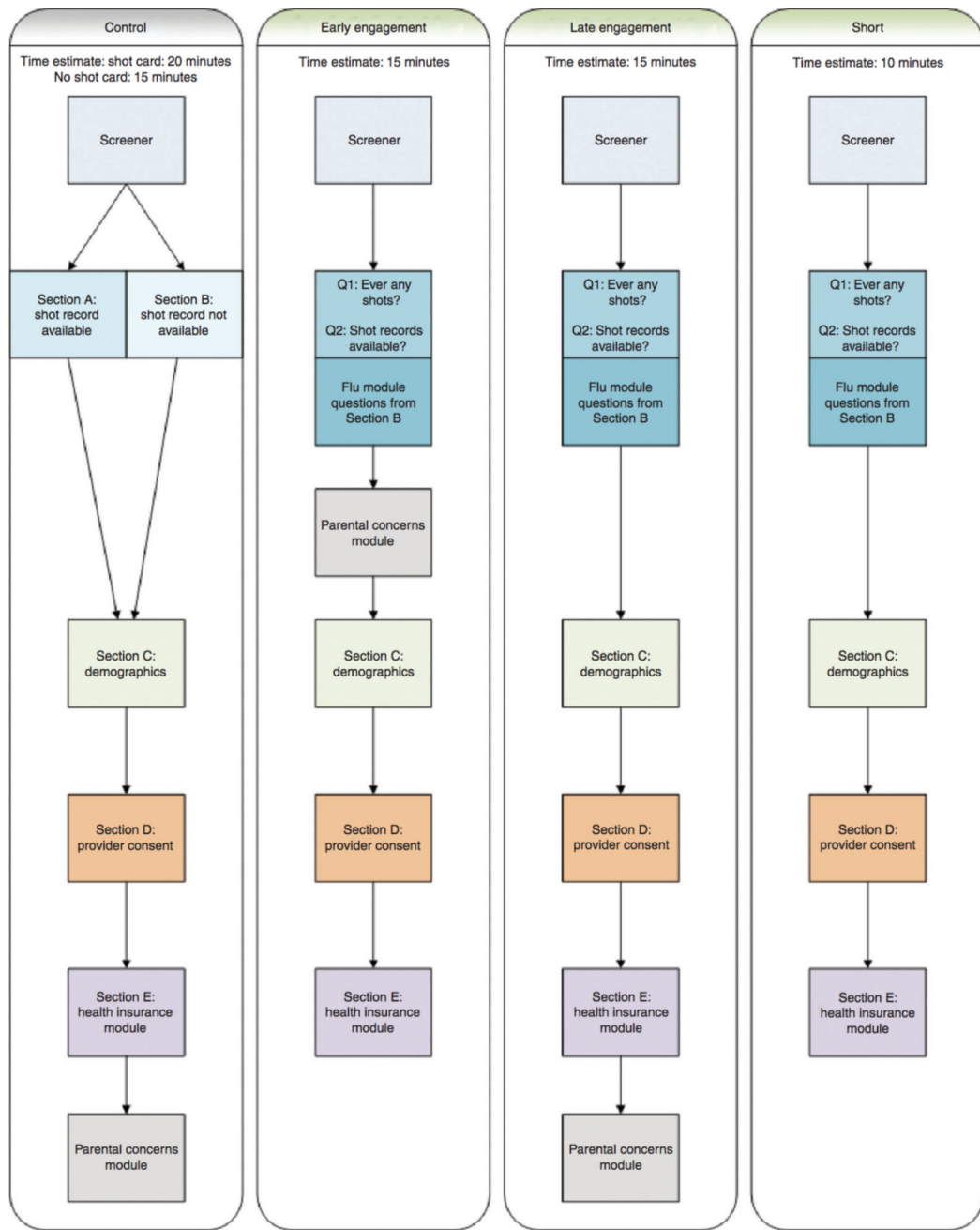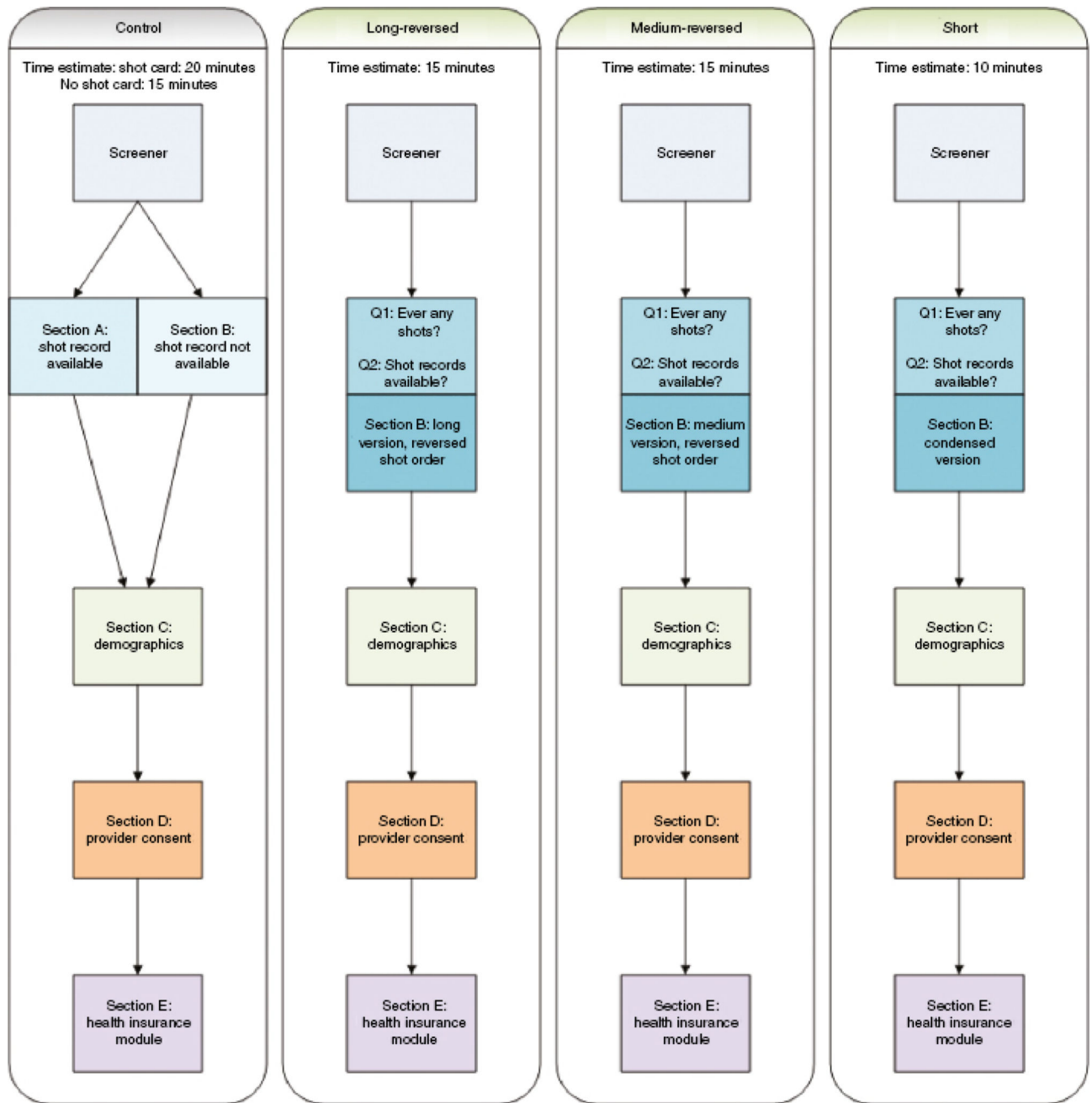**Figure 1.**
Design of the NIS questionnaire redesign study.

Figure 2.
Design of the NIS-Teen questionnaire redesign study.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Key performance rates: NIS Study (time estimate).

Instrument completion rate by time Estimate: NIS-Child questionnaire redesign study (time estimate in minutes)[a]	Control (20 or 15 minutes)	Early engagement (15 minutes)	Late engagement (15 minutes)	Short (10 minutes)	Significance[f]			
					b	c	d	e
Resolution rate	53.54%	53.89%	53.68%	54.05%	n.s.	n.s.	n.s.	n.s.
Screener completion rate	78.04%	77.77%	78.10%	77.82%	n.s.	n.s.	n.s.	n.s.
Interview completion rate	75.05%	80.22%	81.83%	86.38%	0.001	0.001	0.001	0.001
CASRO response rate	31.36%	33.62%	34.31%	36.34%	0.05	0.05	0.01	0.001
Percent reduction in HPC	–	8.28%	11.59%	22.30%				

[a] See Figure 1 for the questionnaire structure of each condition.

[b] Statistical significance reported from chi-squared test of proportions among all four conditions.

[c] Statistical significance reported from pairwise comparison on Early and Control conditions.

[d] Statistical significance reported from pairwise comparison on Late and Control conditions.

[e] Statistical significance reported from pairwise comparison on Short and Control conditions.

[f] Levels of significance reported: $p < 0.05$, $p < 0.01$, $p < 0.001$.

Table 2

Key performance rates:: NIS-Teen study rates by condition and time estimate.

Instrument completion rate by time estimate:: NIS-Teen questionnaire redesign study (time estimate in minutes)[a]				Significance[f]			
	Control (20 or 15 minutes)	Long (15 minutes)	Medium (15 minutes)	Short (10 minutes)	b	c	d	e
Resolution rate	38.24%	39.31%	38.64%	38.97%	n.s.	n.s.	n.s.	n.s.
Screener completion rate	73.01%	73.26%	74.29%	73.33%	n.s.	n.s.	n.s.	n.s.
Interview completion rate	71.83%	73.03%	78.63%	82.05%	0.001	n.s.	0.001	0.001
CASRO response rate	20.05%	21.03%	22.57%	23.45%	0.001	n.s.	0.01	0.001
Percent reduction in HPC	–	4.9%	12.8%	23.8%				

[a]See Figure 2 for the questionnaire structure of each condition.

[b]Statistical significance reported from chi-squared test.

[c]Statistical significance reported from pairwise comparison on Long and Control conditions.

[d]Statistical significance reported from pairwise comparison on Medium and Control conditions.

[e]Statistical significance reported from pairwise comparison on

[f]Levels of significance reported: $p < 0.05$, $p < 0.01$, $p < 0.001$.