

2008/18 BACCALAUREATE AND BEYOND
(B&B:08/18) FULL-SCALE

OMB # 1850-0729 v. 13

Appendix C
Results of the B&B:08/18 Field Test Experiments

Submitted by
National Center for Education Statistics
U.S. Department of Education

February 2018

The field test of the 2008/18 Baccalaureate and Beyond Longitudinal Study (B&B:08/18) included two sets of experiments: the first set of experiments focused on survey participation and reduction of nonresponse error (Section C.1), while the second experiment focused on minimizing measurement error in survey items to further improve data quality (Section C.2). Full details of the experimental design were described and approved in B&B:08/18 Field Test (OMB# 1850-0729 v. 11-12) Supporting Statement Part B.

C.1 Evaluation of Data Collection Experiments

Decreasing response rates have been challenging survey researchers for many decades (e.g., Massey and Tourangeau 2012) since they increase the potential for nonresponse bias, increase survey cost, and potentially decrease sample sizes. Compared to surveys with completely standardized procedures, targeted or tailored designs have been used successfully to tackle nonresponse and attrition by increasing relevance and legitimacy of a study and reducing respondent burden (e.g., Groves and Heeringa 2006; Lynn 2017). Three data collection experiments for the B&B:08/18 field test were designed to investigate the effects of different aspects of tailoring: tailoring of contact materials, highlighting NCES as the survey source and signatory of e-mails (referred to as the sponsorship experiment), and offering a mini survey with an additional survey mode (i.e., offering the mini survey with/without a PAPI option). The B&B:08/18 field test data collection results provide insight in preparation for the full-scale study regarding the effectiveness of the various interventions in terms of rates of survey response, representativeness and data collection efficiency.

Survey response is investigated using response rates and résumé upload rates conditional on survey participation. Based on administrative frame data, B&B:08/18 staff conducted nonresponse bias analyses to assess representativeness for age, institutional sector of the NPSAS institution, region of the United States that the NPSAS institution is located in, and total enrollment counts. Efficiency is operationalized as number of the days between the start of the experiment and survey completion.¹ The analysis uses one-sided t-tests to assess whether survey response or the efficiency increases significantly in the experimental groups and two-sided t-tests to assess nonresponse bias. Table C.1 summarizes the indicators, their operationalization, and the analytic approaches.

Table C.1. Overview of indicators, operationalization and analytic approaches.

Indicator	Operationalization	Analytic Approach
Response		
Survey Completion	Response rates	t-test
Résumé Submission	Résumé upload rate	t-test
Representativeness of sample relative to population	Nonresponse bias <ul style="list-style-type: none"> • Set of indicators for all sample members • Comparison of estimates for respondents to estimates of full sample • 2 measures <ul style="list-style-type: none"> • Summary measures of absolute relative nonresponse bias • Count of significantly biased indicators (out of 21) 	descriptive t-test
Efficiency	Number of days from start of experiment to survey completion	t-test

¹ This analysis only includes respondents who completed the full survey online or via the telephone and exclude those respondents who completed a partial interview or via paper.

The overall response rate among the cases fielded ($n = 1,557$) in the B&B:08/18 field test is 61.3% ($n = 955$). This total includes respondents who completed the full interview ($n = 733$), those who completed part of the interview (partial completions) ($n = 18$), and respondents who completed the mini survey ($n = 204$).

a) *Experiment #1a: Tailoring of Contact Materials*

Background. The first experiment is aimed at increasing topic saliency, interest in the study and rewards of participating by communicating high personal relevance in the contact materials (for tailoring of advance materials and the theoretical motivation see Blau 1964; Cialdini 1984; Groves et al. 1992; Groves and McGonagle 2001; Groves et al. 2000; Lynn 2016; Tourangeau et al. 2010). To increase the personal relevance and motivation to participate in the B&B:08/18 field test B&B:08/18 staff customized the contact materials to reference the sample member's bachelor's degree major (tailored condition) in the experimental group. Letters in the control group included no such reference (standardized condition).²

To investigate the effects of tailoring, sample members with the available information about their bachelor's degree major from the B&B:08/09 field test ($n=1,100$) were randomly assigned to either the standardized condition ($n=633$), or to the tailored condition ($n=467$).³ Sample members for whom this information was not available were assigned to the standardized condition and were excluded from the subsequent analyses.

Results.

Response. Overall response rates in the two conditions were similar (tailoring: 71.7%, standardized: 72.0%, $t(1,098) = -0.11$, $p = .55$). Because the literature suggests that tailoring is more effective among reluctant sample members (Lynn 2016), B&B:08/18 staff calculated the effect of tailoring by whether the individuals had responded to the B&B:08/12 field test. Among B&B:08/12 field test nonrespondents, who are presumed to be less likely to respond to the B&B:08/18 field test, survey response increased by 6.4 percentage points, from 36.0% in the standardized condition to 42.4% percent in the tailored condition ($t(179) = 0.88$, $p = .19$). However, this finding, which is based on a small sample, was not statistically significant. Among B&B:08/12 field test respondents, this increase in response rates is only 1 percentage point from 78 percent to 79 percent ($t(920) = 0.36$, $p = .36$).

Representativeness. Overall, relative nonresponse bias in the tailored condition is lower both in terms of magnitude of the maximum relative bias as well as in terms of number of significantly biased indicators (i.e., 3 biased indicators under the standard condition (14.3%) compared to 0 biased indicators under the tailored condition (0.0%)). These findings suggest that tailoring leads to a more representative sample. There is little difference across both conditions for the average and the median absolute relative bias. Table C.2 summarizes the results.

² Examples include "B&B is interested in understanding how earning a *bachelor's degree* impacted your choices" in the *standardized* letter, and "B&B is interested in understanding how earning a *bachelor's degree in Engineering* impacted your choices" in the *tailored* letter.

³ The field test did not allow respondents to declare a double major.

Table C.2: Average, median, and maximum absolute relative bias, and percentage of significant deviations by experimental group

	Standardized materials	Tailored materials
Average absolute relative bias	8.38	9.45
Median absolute relative bias	5.53	5.53
Maximum absolute relative bias	44.47	39.40
Percentage of significantly biased indicators	14.29%	#

Rounds to zero.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Efficiency. While there is a positive trend such that respondents in the tailored condition (28.7 days) complete the survey on average one and a half days faster compared to respondents in the standardized condition (30.2 days), statistically significant findings were not detected ($t(753) = -0.70, p = .24$).

b) Experiment #1b: Emphasis of NCES as Source and Signatory of E-mails

Background. Past research has shown that individuals are “more likely to comply with a request if it comes from an authority” (Groves et al. 1992, p. 472). This is based on an increased sense of legitimacy for certain research (i.e., the government needs this information) and on trust, due to government employees facing high penalties when disclosing provided information (Dillman et al. 2014). A government sponsorship furthermore may increase the feeling of social responsibility and a sense of civic duty. Positive effects on response rates have been reported for organizations such as a university or government sponsors, compared to other, unknown organizations (e.g., Avdeyeva and Matland 2013; Edwards et al. 2014; Groves et al. 2012; Heberlein and Baumgartner 1978).

To investigate this effect, sample members were randomly assigned to receive e-mails from an “@rti.org” e-mail address, signed by the RTI study director (followed by the signature from the NCES study director), or to receive e-mails from an “@ed.gov” e-mail address signed by the NCES study director (followed by the signature of the RTI study director). This experiment started with the first e-mail reminder and continued throughout the end of data collection. The first condition is referred to as the “RTI” condition (n=670) and the latter condition as the “NCES” condition (n=662). This random assignment was crossed with the assignment of the tailoring experiment to ensure the ability to measure the independent effects of tailoring and sponsorship.

Results.

Response. Both groups achieved identical response rates at the end of data collection (54.8%, $t(1,330) = 0.02, p = .49$). The NCES condition did perform slightly better regarding the résumé upload rate (33.1%) compared to the RTI condition (30.5%) but this difference is not statistically significant ($t(728) = 0.74, p = .23$).

Representativeness. Investigating nonresponse bias, the results suggest that sending e-mails using an NCES address yields a more representative sample. Average, median, and maximum absolute relative bias are all lower in magnitude in the NCES condition resulting in only 1 biased indicator as opposed to 2 in the RTI condition (see Table C.3).

Table C.3: Average, median, and maximum absolute relative bias, and percentage of significant deviations by experimental group

	RTI condition	NCES condition
Average absolute relative bias	10.96	10.05
Median absolute relative bias	8.90	7.37
Maximum absolute relative bias	54.47	44.50
Percentage of significantly biased indicators	10.53%	4.76%

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Efficiency. No statistical difference was observed in the number of days to complete the interview between respondents in the NCES condition and those in the RTI condition. After the start of the experiment, respondents in the NCES condition took an average of 30.8 days to complete the interview, and respondents in the RTI condition took an average of 32.7 days.

c) Experiment #1c: Mini Survey

Background. Reducing the burden of the survey as a tool for nonresponse conversion, for example, by decreasing the survey length and offering alternative modes of completion, has been shown to successfully increase participation rates and increase representativeness in surveys (Biemer et al. 2016; Galesic and Bosnjak 2009; Groves and Couper 1998; Messer and Dillman 2011; Mowen and Cialdini 1980; Shettle and Mooney 2009).

To increase participation rates among the more reluctant sample members in the B&B:08/18 field test, sample members who failed to complete the survey by week 10 (out of 16) were offered a highly abbreviated version of the survey consisting of approximately 10 questions (referred to as the mini survey). Sample members were furthermore randomly assigned to either complete this mini survey in the original survey modes (mini-standard; n=404) or a condition that additionally allowed a completion using paper and pencil (mini-PAPI; n=402). Six sample members completed the mini survey before receiving the invitation to do so which is why those cases are excluded from the subsequent analyses reducing the analytic sample size to n= 401 in the mini-standard condition and n=399 in the mini-PAPI condition.

Results.

Response. The mini survey significantly increased overall response rate, relative to the usual interview, from 48.6% to 61.3% among the fielded cases ($t(1,556) = 7.19, p < .001$). As expected, of those sample members who had not completed the survey by week 10 of data collection, the mini-PAPI achieved a higher response rate (26.1%) compared to that in the mini-standard condition (23.4%). However, while the direction of this effect is as expected, statistically significant findings were not detected ($t(798) = 0.86, p = .20$). Conditional on participation in the mini survey, respondents in the mini-standard condition did upload their résumés at higher rates (35.1%) compared to those in the mini-PAPI condition (18.3%) ($t(196) = -2.73, p < .01$). The lower submission rate in the mini-PAPI condition rate is driven by the fact that none of the respondents who completed the survey via mail uploaded their résumés.⁴ Among the mini-PAPI respondents who completed the survey via the web, 25.7% uploaded their résumé, less than those in the mini-standard condition.

Representativeness. The mini-PAPI increased representativeness by reducing the magnitude of nonresponse bias across all three indicators, maximum, average, and median absolute relative nonresponse bias (see Table C4). Both conditions produce samples in which none of the indicators are significantly biased.

⁴ PAPI respondents were encouraged to upload their résumés online.

Table C.4: Average, median, and maximum absolute relative bias, and percentage of significant deviations by experimental group

	Mini-standard condition	Mini-PAPI condition
Average absolute relative bias	15.46	14.64
Median absolute relative bias	10.95	9.06
Maximum absolute relative bias	78.50	44.50
Percentage of significantly biased indicators	#	#

Rounds to zero.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Efficiency. B&B:08/18 staff adjusted the measure of efficiency capturing the time between when sample members received their mail invitation to complete the mini survey and when they completed the survey online to allow for mail transit time. Contrary to the expectations, the results suggest that respondents in the mini-PAPI condition completed the survey about three days later (23.0 days) compared to respondents in the mini-standard condition (19.9 days). These results are statistically significant ($t(166) = 1.66, p < .05$), and do not include respondents in the mini-PAPI version who complete the survey via mail ($n=30$).

Recommendations for the full-scale study

While there was no statistically significant increase in response rates or résumé submission rates⁵ in any of the data collection experiments, the field test results are suggestive of positive effects. Given that the interventions tested in the field test show no indication of negative effects,⁶ that the technical review panel members and the literature support these adjustments,⁷ and they are low-cost and easy to implement, the recommendations for the full-scale study are to: tailor the contact materials and reference the sample members' degree major(s), use NCES as the primary signatory and sender of the electronic communication materials, and use a sequential approach such as offering the mini survey followed by the mini-PAPI.

Using the B&B:08/18 field test to estimate response propensities this should yield an approximate overall response rate of 72%.⁸ To reduce the potential for nonresponse error and bias, contain costs, and achieve response rates of 75%, B&B:08/18 staff propose further modifications to the data collection protocols used in the full-scale study based on what lessons learned in other studies, such as the B&B:16/17 field test (see B&B:16/17 Appendix C, OMB# 1850-0926 v.3) or the B&B:08/12 (B&B:08/12 Data File Documentation <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2015141>).

C.2 Evaluation of Questionnaire Design Experiment

Experimental studies show that forced-choice formats yield consistently higher endorsement rates suggesting deeper cognitive processing and higher data quality (Smyth et al. 2006; Thomas et al. 2017). For this experiment, there are two grids where there are significant differences between the two forced-choice formats (Military Status and Result of Undergraduate costs). For the Military Status grid, those with the No/Yes forced-choice format had higher numbers of endorsement (0.12) compared to the

⁵ With the exception of the lower rate of résumé submission among mini-PAPI respondents.

⁶ With the exception of the lower rate of résumé submission among mini-PAPI respondents.

⁷ For tailoring, see Lynn 2016 and Tourangeau et al. 2010. For sponsorship, see Avdeyeva and Matland 2013; Edwards et al. 2014; and Groves et al. 2012. For mini-PAPI, see Biemer et al. 2016; Galesic and Bosnjak 2009; and Messer and Dillman 2011.

⁸ The B&B:08/18 field test was a purposive sample and is hence not entirely comparable to the full-scale sample.

Yes/No forced-choice format (0.04). However, the opposite finding was present for the result of undergraduate costs grid (No/Yes: 1.52; Yes/No: 1.78). These opposite results, and the lack of significant findings for the other four grids suggest that there is no acquiescence bias with the Yes/No grid.

Table C.6: Average number of affirmative responses for each grid by experimental group

	Check-All Format	Yes/No Forced-Choice Format	No/Yes Forced-Choice Format
Financial Aid	0.66	0.89	0.85
Reasons for Employment Change	0.19	0.28	0.28
Military Status	0.05	0.04	0.12
Current Household	1.28	1.32	1.26
Type of Retirement Accounts	1.32	1.40	1.43
Result of Undergraduate costs	1.33	1.78	1.52

NOTE: Results exclude telephone respondents.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Table C.7: Test statistics and *p*-values for average number of affirmative responses for each grid by experimental group

Grid Question	z-value	<i>p</i> -value
Financial Aid		
Check all that apply vs. Yes/No Forced-choice	2.93	.0034
Check all that apply vs. No/Yes Forced-choice	2.42	.0156
Yes/No Forced-choice vs. No/Yes Forced-choice	-0.47	.6390
Reasons for Employment Change		
Check all that apply vs. Yes/No Forced-choice	2.00	.0455
Check all that apply vs. No/Yes Forced-choice	2.08	.0375
Yes/No Forced-choice vs. No/Yes Forced-choice	0.12	.9067
Military Status		
Check all that apply vs. Yes/No Forced-choice	-0.88	.3805
Check all that apply vs. No/Yes Forced-choice	2.37	.0177
Yes/No Forced-choice vs. No/Yes Forced-choice	3.07	.0022
Current Household		
Check all that apply vs. Yes/No Forced-choice	0.36	.7155
Check all that apply vs. No/Yes Forced-choice	-0.24	.8138
Yes/No Forced-choice vs. No/Yes Forced-choice	-0.59	.5526
Type of Retirement Accounts		
Check all that apply vs. Yes/No Forced-choice	0.78	.4345
Check all that apply vs. No/Yes Forced-choice	1.07	.2832
Yes/No Forced-choice vs. No/Yes Forced-choice	0.31	.7580
Result of Undergraduate costs		
Check all that apply vs. Yes/No Forced-choice	3.88	.0001
Check all that apply vs. No/Yes Forced-choice	1.71	.0872
Yes/No Forced-choice vs. No/Yes Forced-choice	-2.12	.0344

NOTE: Results exclude telephone respondents. Significance tests based on Poisson models where the first group listed is the reference category.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Item Nonresponse. Unless there is an explicit checkbox for “None of the above” an unchecked box in a check-all format is hard to interpret as it might mean that (1) the response option does not apply, (2) the respondent missed the item in the list, or (3) that the respondent was unsure. To compare item nonresponse across all three formats, whether or not the entire grid lacked any responses was investigated, however, no difference was observed ($F = 0.75, p = .47$). Comparing item nonresponse across both forced-choice formats, the results in table C.8 show that both formats had similar item nonresponse rates across all six grids.

Table C.8: Average rates of item nonresponse for each grid by experimental group for the two forced-choice formats

	Yes/No Forced-Choice Format	No/Yes Forced-Choice Format	t-statistic	p-value
Financial Aid	17.5%	16.0%	-0.29	.7723
Reasons for Employment Change	7.9%	11.7%	0.50	.6221
Military Status	3.1%	2.0%	-0.76	.4491
Current Household	4.1%	3.7%	-0.27	.7901
Type of Retirement Accounts	10.4%	10.0%	-0.16	.8714
Result of Undergraduate costs	4.7%	2.3%	-1.43	.1543

NOTE: Results exclude telephone respondents. Significance tests based on simple linear regression models where the first group listed is the reference category.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Completion Time. The forced-choice formats took significantly longer for respondents to complete compared to the check all that apply format for all six grid questions. On average, a check all that apply format took 13.86 seconds compared to 16.43 seconds for the Yes/No format and 16.45 seconds for the No/Yes format (table C.9). The differences in times across all grids and overall for the two forced-choice formats were not statistically significant (table C.10). This suggests that the forced-choice formats encourage more cognitive processing than the check all that apply format and that there is not acquiescence bias occurring.

Table C.9: Average time (in seconds) spent on each grid and overall by experimental group

	Check-All Format	Yes/No Forced-Choice Format	No/Yes Forced-Choice Format
Financial Aid	15.35	29.59	29.92
Reasons for Employment Change	19.14	31.48	28.22
Military Status	8.51	9.62	9.66
Current Household	10.58	11.62	11.65
Type of Retirement Accounts	13.67	18.74	17.13
Result of Undergraduate costs	18.27	23.18	23.13
Overall	13.86	16.43	16.45

NOTE: Results exclude telephone respondents. Significance tests based on simple linear regression models where the first group listed is the reference category. To minimize the effect of extreme timing values on the results, outliers were identified and excluded from these analyses. First, each value was transformed by taking its natural logarithm. Second, within each survey screen, transformed values greater than the 75th percentile plus 1.5 times the interquartile range of the distribution or less than the 25th percentile minus 1.5 times the interquartile range of the distribution were removed.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Table C.10: Test statistics and *p*-values for average time (in seconds) for each grid by experimental group

Grid Question	<i>t</i>-value	<i>p</i>-value
Financial Aid		
Check all that apply vs. Yes/No Forced-choice	9.42	<.0001
Check all that apply vs. No/Yes Forced-choice	9.09	<.0001
Yes/No Forced-choice vs. No/Yes Forced-choice	-0.16	.8761
Reasons for Employment Change		
Check all that apply vs. Yes/No Forced-choice	3.20	.0017
Check all that apply vs. No/Yes Forced-choice	3.57	.0005
Yes/No Forced-choice vs. No/Yes Forced-choice	0.86	.3905
Military Status		
Check all that apply vs. Yes/No Forced-choice	3.10	.0020
Check all that apply vs. No/Yes Forced-choice	3.00	.0028
Yes/No Forced-choice vs. No/Yes Forced-choice	-0.12	.9014
Current Household		
Check all that apply vs. Yes/No Forced-choice	2.41	.0165
Check all that apply vs. No/Yes Forced-choice	2.31	.0210
Yes/No Forced-choice vs. No/Yes Forced-choice	-0.08	.9376
Type of Retirement Accounts		
Check all that apply vs. Yes/No Forced-choice	4.45	<.0001
Check all that apply vs. No/Yes Forced-choice	6.08	<.0001
Yes/No Forced-choice vs. No/Yes Forced-choice	1.77	.0765
Result of Undergraduate costs		
Check all that apply vs. Yes/No Forced-choice	5.84	<.0001
Check all that apply vs. No/Yes Forced-choice	5.80	<.0001
Yes/No Forced-choice vs. No/Yes Forced-choice	0.05	.9630
Overall		
Check all that apply vs. Yes/No Forced-choice	6.28	<.0001
Check all that apply vs. No/Yes Forced-choice	6.34	<.0001
Yes/No Forced-choice vs. No/Yes Forced-choice	0.05	.9602

NOTE: Results exclude telephone respondents. Significance tests based on simple linear regression models where the first group listed is the reference category. To minimize the effect of extreme timing values on the results, outliers were identified and excluded from these analyses. First, each value was transformed by taking its natural logarithm. Second, within each survey screen, transformed values greater than the 75th percentile plus 1.5 times the interquartile range of the distribution or less than the 25th percentile minus 1.5 times the interquartile range of the distribution were removed.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2008/18 Baccalaureate and Beyond (B&B:08/18) Field Test.

Recommendations for the full-scale study

Based on these results, implementing the Yes/No forced-choice format is recommended. There are generally higher endorsement rates with the forced-choice formats compared to the check all that apply format. The forced-choice format also had longer average completion times, indicating that they lead to more cognitive processing compared with the check-all-that-apply-format. As there are apparent no differences in data quality between the two forced-choice formats, implementing the format that is most common for respondents (i.e., Yes then No) is suggested. The only item that will remain a check all format is the item on household composition as this greatly simplifies the response task for this particular question. The option 'live alone' automatically implies that the other options do not apply and hence reduces respondent burden.

References

- Avdeyeva, O.A., and Matland, R.E. 2013. An Experimental Test of Mail Surveys as a Tool for Social Inquiry in Russia. *International Journal of Public Opinion Research*, 25(2), 173-194.
- Biemer, P., Murphy, J., Zimmer, S., Berry, C., Deng, G., and Lewis, K. 2016. A Test of Web/PAPI Protocols and Incentives for the Residential Energy Consumption Survey. Paper presented at the 2016 Annual Conference of the American Association for Public Opinion Research, Austin, TX (May 13, 2016).
- Blau, P.M. 1964. *Exchange and Power in Social Life*. New York, NY: Wiley.
- Cialdini, R.B. 1984. *Influence: The New Psychology of Modern Persuasion*. New York, NY: Quill.
- Callegaro, M., Murakami, H., Tepman, Z., and Henderson V. 2015. Yes-no Answers Versus Check-all in Self-administered Modes. *International Journal of Market Research*, 57(2): 203-223.
- Deming, W. E. (1953). On a Probability Mechanism to Attain an Economic Balance Between the Resultant Error of Response and the Bias of Nonresponse. *Journal of the American Statistical Association*, 48(264), 743-772.
- Dillman, D.A., Sinclair, M.D., and Clark, J.R. 1993. Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys. *Public Opinion Quarterly*, 57(3), 289-304.
- Dillman, D.A., Smyth, J.D., and Christian, L.M. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method 4th Edition*. John Wiley & Sons, Hoboken, NJ.
- Dutwin, D., Loft, J.D., Darling, J., Holbrook, A., Johnson, T., Langley, R.E., Lavrakas, P.J., Olson, K., Peytcheva, E., Stec, J., Triplett, T., and Zukerberg, A. 2014. Current Knowledge and Considerations Regarding Survey Refusals. AAPOR Task Force Report on Survey Refusal. Retrieved at http://www.aapor.org/Education-Resources/Reports/Current-Knowledge-and-Considerations-Regarding-Sur.aspx#_Toc393959577.
- Edwards, M.L., Dillman, D.A., and Smyth, J.D. 2014. An Experimental Test of the Effects of Survey Sponsorship on Internet and Mail Survey Response. *Public Opinion Quarterly*, 78(3), 734-750.
- Freedman, J.L., and Fraser, S.C. 1966. Compliance Without Pressure: The Foot-in-the-door Technique. *Journal of Personality and Social Psychology*, 4(2), 196-202.
- Galesic, M., and Bosnjak, M. 2009. Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*, 73(2), 349-360.
- Groves, R.M. 2006. Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70(5), 646-675.
- Groves, R.M., Cialdini, R., and Couper, M. 1992. Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 56(4), 475-495.

- Groves, R.M., and Couper, M. 1998. *Nonresponse in Household Interview Surveys*. Wiley.
- Groves, R.M., and McGonagle, K.A. 2001. A Theory-guided Interviewer Training Protocol Regarding Survey Participation. *Journal of Official Statistics*, 17, 249-65.
- Groves, R.M., Singer, E., and Corning, A. 2000. Leverage-Saliency Theory of Survey Participation. Description and Illustration. *Public Opinion Quarterly*, 64, 299-308.
- Groves, R.M., Presser, S., Tourangeau, R., West, B.T., Couper, M.P., Singer, E., and Toppe, C. 2012. Support for the Survey Sponsor and Nonresponse Bias. *Public Opinion Quarterly*, 76, 512-24.
- Heberlein, T.A., and Baumgartner, R. 1978. Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature. *American Sociological Review*, 43(4), 447-462.
- Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T.M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R.M., and Raghunathan, T.E. 2010. Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust For Non-Response: Examples from Multiple Surveys. *Journal of the Royal Statistical Society: Series A*, 173(2), 389-407.
- Lynn, P. 2016. Targeted Appeals for Participation in Letters to Panel Survey Members. *Public Opinion Quarterly*, 80(3), 771-782.
- Lynn, P. 2017. From Standardised to Targeted Survey Procedures for Tackling Non-response and Attrition. *Survey Research Methods*, 17(1), 93-103.
- Massey D.S., and Tourangeau R. 2012. Where Do We Go from Here? Nonresponse and Social Measurement. *The Annals of the American Academy of Political and Social Science*. 645(1), 222-236.
- Medway, R.L., and Tourangeau, R. 2015. Response Quality in Telephone Surveys. Do Prepaid Incentives Make a Difference? *Public Opinion Quarterly*, 79(2), 524-543.
- Messer, B.L., and Dillman, D.A. 2011. Surveying the General Public Over the Internet Using Address-Based Sampling and Mail Contact Procedures. *Public Opinion Quarterly*, 75(3), 429-457.
- Millar, M.M., and Dillman, D.A. 2011. Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly*, 75(2), 249-269.
- Mowen, J.C., and Cialdini, R.B. 1980. On Implementing the Door-in-the-Face Compliance Technique in a Business Context. *Journal of Marketing Research*, 17, 253-258.
- Schouten, B., Cobben, F. and Bethlehem, J. 2009. Indicators for the Representativeness of Survey Response. *Survey Methodology*, 35(1), 101-113.
- Shettle, C., and Mooney, G. 1999. Monetary Incentives in US Government Surveys. *Journal of Official Statistics*, 15(2), 231-250.
- Singer, E., and Ye, C. 2013. The Use and Effects of Incentives in Surveys. *Annals of the American Academy of Political and Social Science*, 645(1), 112-141.

- Smyth, J.D., Dillman, D.A., Christian, L.M., and Stern, M.J. 2006. Comparing Check-all and Forced-choice Question Formats in Web Surveys. *Public Opinion Quarterly*, 70(1), 66-77.
- Thomas, R.K., Barlas, F.M, Buttermore, N.R., and Smyth, J.D. 2017. Acquiescence Bias in Yes-No Grids? The Survey Says... No. Presented at the 2017 Annual Conference of the American Association for Public Opinion Research, New Orleans, LA (May 20, 2017).
- Tourangeau, R., Groves, R.M., and C.D. Redline. 2010. Sensitive Topics and Reluctant Respondents. Demonstrating a Link Between Nonresponse Bias and Measurement Error. *Public Opinion Quarterly* 74(3), 413-432.