Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates

Daikeler, Jessica; Bosnjak, Michael; Lozar Manfreda, Katja

This is a pre-copyedited, author-produced PDF of an article accepted for publication in
Journal of Survey Statistics and Methodology, Vol. 8, 2020, No. 3, pp. 513-539
following peer review.
The version of record is available online at: https://doi.org/10.1093/jssam/smz008
Page numbers have been adjusted to the publishers version, whereby this postprint is fully quotable.

Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates

Jessica Daikeler*
Michael Bošnjak
Katja Lozar Manfreda

Do web surveys still yield lower response rates compared with other survey modes? To answer this question, we replicated and extended a meta-analysis done in 2008 which found that, based on 45 experimental comparisons, web surveys had an 11 percentage points lower response rate compared with other survey modes. Fundamental changes in internet accessibility and use since the publication of the original meta-analysis would suggest that people's propensity to participate in web surveys has changed considerably in the meantime. However, in our replication and extension study, which comprised 114 experimental comparisons between web and other survey modes, we found almost no change: web surveys still yielded lower response rates than other modes (a difference of 12 percentage points in response rates). Furthermore, we found that prenotifications, the sample recruitment strategy, the survey's solicitation mode, the type of target population, the number of contact attempts, and the country in which the survey was conducted moderated the magnitude of the response rate differences. These findings have substantial implications for web survey methodology and operations.

KEYWORDS: Meta-analysis; Online survey; Replication and extension; Response rate difference; Response rates; Web survey usage.

JESSICA DAIKELER is with the GESIS, Leibniz Institute for the Social Sciences, B2, 1, 68159 Mannheim, Germany. MICHAEL BOŠNJAK is a Professor with the ZPID, Leibniz Institute for Psychology Information, Universitätsring 15, 54296 Trier, Germany. KATJA LOZAR MANFREDA is an Associate Professor with the Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5, 1000 Ljubljana, Slovenia
∗ Address correspondence to Jessica Daikeler, GESIS, Leibniz Institute for the Social Sciences, B2, 1, 68159 Mannheim, Germany; E-mail: jessica.daikeler@gesis.org.

1. INTRODUCTION

The use of online surveys is on the rise; in 2007, for the first time, online surveys constituted the majority of all quantitative survey modes implemented worldwide. According to ESOMAR's latest Global Market Research Report (2018, p. 139), web survey use has more than doubled compared with 2007. Underlying this widespread growth is the transformation of the web surveys from an initially novel to a well-established mode of survey implementation. The broad discussion on online data quality has pointed out positive data quality aspects of the web mode (e.g., an increased level of reporting of sensitive information) (Kreuter, Presser, and Tourangeau 2008; Sakshaug, Yan, and Tourangeau 2010) and time-sensitive aspects (Chang and Krosnick 2009). On the other hand, it has also revealed several shortcomings of web surveys such as question skipping, speeding, response inconsistency, satisficing (Heerwegh and Loosveldt 2008; Kim, Dykema, Stevenson, Black, and Moberg 2018), and representativeness issues (Cornesse and Bosnjak 2018). Web surveys are especially useful when surveying specific populations with high internet coverage such as students, customers, and employees with email addresses. (Cernat, Couper, and Ofstedal 2016; Patrick, Couper, Laetz, Schulenberg, O'Malley, et al. 2018). For these populations, the nonresponse bias problem is usually low. For the general population, however, internet users and noninternet users are not randomly distributed (Chang and Krosnick 2009; Blom, Burton, Booker, Cernat, Fairbrother, et al. 2015), and this presents a challenge to many online surveys.

Although the quality aspects of web surveys that deserve further attention are numerous, the present study limits the discussion to response rates as an indicator of nonresponse error.

2. BACKGROUND

Experimental studies comparing the response rates of web surveys with those of other survey modes have reported higher response rates for traditional survey modes (Fricker, Galesic, Tourangeau, and Yan 2005; Kirchner and Felderer 2016). By contrast, a substantial body of literature has emphasized the advantages of web surveys over traditional modes (Greene, Speizer, and Wiitala 2008; Boyle, Morrison, MacDonald, Duncan, and Rose 2016). Although these are individual experimental studies, several systematic reviews of response rate comparisons have also been conducted. For instance, Shih and Fan (2008) carried out a meta-analysis comparing only the response rates of web surveys and mail surveys and found, on average, that mail surveys had higher response rates than web surveys. However, the most comprehensive research synthesis to date on the response rate difference between web and other survey modes was conducted by Lozar Manfreda, Bosnjak, Berzelak, Haas, and Vehovar (2008). On average, the authors found an 11 percentage points lower response rate for web surveys than for other survey modes. Moreover, and even more importantly, they examined the study characteristics, also known as *moderators*, to determine which ones significantly influence this response rate difference. Their results revealed the following moderators of this difference: the sample recruitment base (i.e., a smaller response rate difference between web and other survey modes in the case of panel members as compared with one-time respondents); the solicitation or invitation mode chosen for web surveys (i.e., a higher response rate difference for postal mail solicitation compared with email solicitation); and the number of times contact is made with respondents (i.e., the more contacts made, the larger the response rate difference between modes).

We designed the present study as a replication and extension of Lozar Manfreda et al. (2008) previous research for two main reasons. First, we wanted to identify the benefits and limitations for web response rates compared with other survey modes; second, we wanted to determine whether Lozar Manfreda et al. (2008) findings are still applicable today. Several years have gone by since Lozar Manfreda et al. (2008) finalized their literature search in 2005, and during this time, the web survey field has faced many changes. Some of the limitations of web surveys have multiplied. First, there is greater sensitivity with respect to data security in today's world (Callegaro, Lozar Manfreda, and Vehovar 2015, p. 125). Second, there has been an increase in the diversity of internet browsers, (mobile) devices, and operating systems, which has caused problems of technical incompatibility (Couper and Peterson 2017). Third, there has been an increase in online over-surveying and spam emailing (Callegaro et al. 2015, p. 157). Fourth, there is a lower legitimacy of researchers who may carry out impersonal and quick web surveys (Groves, Fowler, Couper, Lepkowski, Singer, et al. 2011, p. 149; Callegaro et al. 2015, p. 171). On the other hand, new opportunities for web surveys have been developed due to (1) the increased web literacy of web respondents, which reduces technical limitations (Eshet-Alkalai and Chajut 2010); (2) higher internet coverage rates (e.g., World Bank 2017); (3) the availability of a variety of increasingly user-friendly devices with which to access the internet (e.g., touchscreens, Wi-Fi connections) (Al-Razgan, Al-Khalifa, Al-Shahrani, and AlAjmi 2012); (4) changes in internet access payments (from pay-per-minute to flat rates) (Aichele, Flickenger, Fonda, Forster, Howard, et al. 2013); (5) the fact that contacting people via other modes of communication has become more difficult, for example, due to the increasing number of households without landline telephones (Dillman, Smyth, and Christian 2014, p. 16).

Our second research objective—to determine whether Lozar Manfreda et al. (2008) findings are still applicable—is prompted by Shojania, Sampson, Ansari, Ji, Doucette, et al. (2007), who addressed in their research the question of how quickly systematic reviews go out of date and demonstrated that the median survival time was only 5.5 years. As a consequence, they recommended the regular updating of systematic reviews. Accordingly, this study aims to answer the following research questions (RQs):

RQ 1: Do web surveys yield lower, higher, or the same response rates as other survey modes?

To answer this question, we update the meta-analysis performed by Lozar Manfreda et al. (2008) with respect to possible changes over time. In addition, we aim to explore whether new studies have increased the explanatory power of the variables presumed to explain the variability of the response rate differences between web and other surveys modes, and to determine whether any other moderators also have an impact.

In the original meta-analysis performed by Lozar Manfreda et al. (2008), certain survey characteristics such as the number of contact attempts had an influence on the response rate differences between web and other survey modes. Therefore, the response rate differences were heterogeneous, and moderator explanation was reasonable. In our replication and extension of this study, we explore whether and to what extent the mean response rate difference varies and what moderator variables explain this variation. Hence, our second research question asks:

RQ2: Is the mean response rate difference heterogeneous?

The success of a survey, and thus the response rate, depends strongly on the survey settings and characteristics (Groves and Peytcheva 2008). We expect deviating effects depending on the modes to which web surveys were compared (e.g., mail, telephone, face-to-face, interactive voice response, touch tone). A paper-based questionnaire usually remains within reach of the respondent for a period of time and can therefore act as a reminder (Dillman et al. 2014, p. 382). In telephone surveys, the time of day that the call is placed plays a crucial role in whether the potential respondent is busy or unavailable to take the call at all (Tourangeau, Michael Brick, Lohr, and Li 2017). Email invitations and reminders for web questionnaires are more likely to be (un)intentionally overlooked (Petrovčič, Petrič, and Lozar Manfreda 2016). Incentives in online surveys can be confused with advertising and not taken seriously, especially if the survey sponsor is not a university or governmental organization. Additional effort must be made by researchers using modes other than the web for their surveys—for instance, mailing letters, making telephone calls, or even paying the respondent a personal visit. These additional efforts, if appreciated by respondents, may account for some of the greater legitimacy of these surveys compared with self-administered web surveys and may, therefore lead to higher response rates compared with email invitations or web questionnaires (Millar, Dillman, Messer, Genter, Williams, et al. 2011). Participation might also be higher if it is requested personally (via telephone or face-to-face), as potential respondents might find such personal requests harder to disregard. Moreover, compared with immediately answering survey questions on the phone, respondents have to be much more active when answering a web survey, especially if no email invitation is provided (Fricker et al. 2005; Greenlaw and Brown-Welty 2009). Nevertheless, surveys for specific target populations with certain characteristics (e.g., higher internet penetration, engagement with the survey topic) might work better online than surveys for the general population. Furthermore, with the success of the internet, the attitude of the population toward web surveys has changed over time. This change, reflected in the response rates, is why we expect an effect of the year of publication. Therefore, our research addresses whether study design or study circumstances have an effect on the response rate difference. These deliberations lead to our third research question:

RQ3: Do the sample recruitment base, solicitation mode, number of contacts, mode to be compared to, type of target population, type of sponsorship, use of incentives, and the year the studies were published impact the variation in the response rate difference?

Whereas our first and second aims in the present study are to update and to increase the statistical power of Lozar Manfreda et al. (2008) meta-analytical findings, which addressed the moderators listed in RQ3, our third aim is to extend their meta-analysis. Therefore, we consider three additional moderators: survey topic, prenotification (i.e., an advance contact with respondents to announce the survey), and survey country. These additions are possible due to the larger number of primary studies included. With regard to the survey topic, it can be assumed that some types of survey topics work better than others on the web. Specifically, web respondents are more likely to provide answers to sensitive questions (Kreuter et al. 2008). In addition, experimental evidence suggests that providing respondents with prenotifications has a consistently positive effect on response rates (Fan and Yan 2010). Here, we seek to determine whether prenotifications exert differential effects on the response rates of web surveys versus other survey modes. Receiving an email request to participate in a web survey may seem less legitimate to respondents, as sending such a request entails less effort on the part of researchers compared with requests via other channels such as telephone or postal mail. Legitimacy is further undermined by the high number of web surveys currently being conducted and the low level of trust in the online world (Dillman et al. 2014, p. 450). Therefore, we assume that the use of prenotifications in web surveys is less advantageous than in other survey modes, and we postulate that prenotification should increase the response rate difference. Interestingly, and to the best of our knowledge, no meta-analyses on response rates have included cross-national factors. This lack is all the more surprising because country specificities and cultural factors—for example, a country's internet coverage, mode-specific survey-taking climate, over-surveying, and openness to new technologies—play a role in the acceptance and conducting of web surveys (Lyberg and Dean 1992; Couper and De Leeuw 2003). Thus, we hypothesize that a variation in response rate differences between web and other survey modes exists across countries. These deliberations give rise to our fourth research question:

RQ4: Is the response rate difference influenced by (1) the use of prenotifications, (2) the survey topic, and (3) the country in which the survey is conducted?

Because we want to isolate the impact of the survey mode from other causes, we include in our meta-analysis only primary studies with experiments that compare web response rates to the response rates of other survey modes. The next section describes our research method. This is followed by the results section, in which we address the mean difference in response rates in web surveys versus other survey modes and the robustness of this difference and provide an analysis of the moderators. The article concludes with a discussion of our findings and the limitations of our study.

3. METHOD

To ensure a proper replication of the original study, response rate differences between web surveys and other survey modes were examined using meta-analytic techniques that closely followed those used in Lozar Manfreda et al. (2008). The present section briefly describes the meta-analytic methods, the eligibility criteria and search strategy, the coding of primary studies, and the statistical procedures.

Our systematic review and meta-analysis comprised four steps. First, we conducted a comprehensive literature search using specific search terms derived from a set of study eligibility criteria. Second, we reviewed the manuscripts identified by this literature search and screened out those that did not comply with our eligibility criteria. In the third step, we coded pertinent data in order to compute response rates, and we used the information on potential moderators to calculate effect sizes and perform the moderator analyses. In the final step, we carried out the meta-analytic statistical analyses. These four steps are explained in detail in the following sections.

3.1 Eligibility Criteria and Search Strategy

For our meta-analysis, we employed the same eligibility criteria as those used by Lozar Manfreda et al. (2008), as close adherence to these criteria was an important precondition for mapping possible changes over time.

Eligible studies had to meet the following criteria: (1) One of the survey modes used had to be a web-based survey (i.e., a survey in which a web questionnaire was used to gather responses from respondents online using various devices. (2) The web-based survey had to be compared with data from one or more other survey modes (e.g., email, mail, telephone, face-to-face, telefax). (3) Data had to be available on response rates of the web and other survey mode(s). (4) A split-sample experimental design had to have been employed with subjects from the same population who were randomly assigned to different modes. In other words, the eligible studies included a study design in which each respondent was randomly assigned to either the web mode or the comparison mode. (5) Subjects had to remain in the mode to which they were randomly assigned; in other words, studies in which subjects were permitted to switch modes were not eligible for inclusion. (6) The implementation of the comparison survey modes had to be identical, with the only difference being the mode used to answer the survey questionnaire. Hence, comparisons of surveys that used unequal incentives were excluded.

There is only one difference between the present criteria and those used in Lozar Manfreda et al. (2008) meta-analysis. In the original meta-analysis, primary studies that had the same number of contact attempts (regardless of the type of contact) were considered to be identical and were thus included in the meta-analysis. By contrast, we excluded experimental comparisons in which only one survey mode used prenotification (although the overall number of contact attempts might have been the same). This was not an option for the original meta-analysis because the number of studies was much smaller, and taking this approach would have led to a loss of statistical power. In addition, having a larger number of studies at our disposal allowed us to examine prenotification as a separate moderator. Consequently, we excluded seven effect sizes (Kaplowitz, Hadlock, and Levine 2001; Miller, Miller Kobayashi, Caldwell, Thurston, and Collett 2002; Grigorian, Sederstrom, and Hoffer 2004; Cole 2005), and in this respect, our study is not an exact replication of Lozar Manfreda et al. (2008) meta-analysis. Given the small number of excluded studies, we still consider this a valid approach to determining change over time. In addition, like Manfreda et al., we imposed no participant population, time period, or geographical restrictions.

As a first important step to ensure the quality of our meta-analysis, we performed a comprehensive literature search, applying the same search terms as those used in Lozar Manfreda et al. (2008) study (see see table A2 conference overview of the online supplementary material). To overcome the "publication bias" (Rosenthal 1979) problem, we employed several techniques. With the aid of a snowballing technique, we inspected the reference lists of the selected publications. However, to explicitly collect grey litertaure, we examined conference abstracts (see table A2 of the online supplementary material) from the years 2005 to 2016. The PRISMA flow diagram (Moher, Liberati, Tetzlaff and Altman, The PRISMA Group 2009) in figure 1 provides an overview of our search strategy, which was restricted to the literature in English. Finally, we included over 100 effect sizes in our meta-analysis (see figure 1).

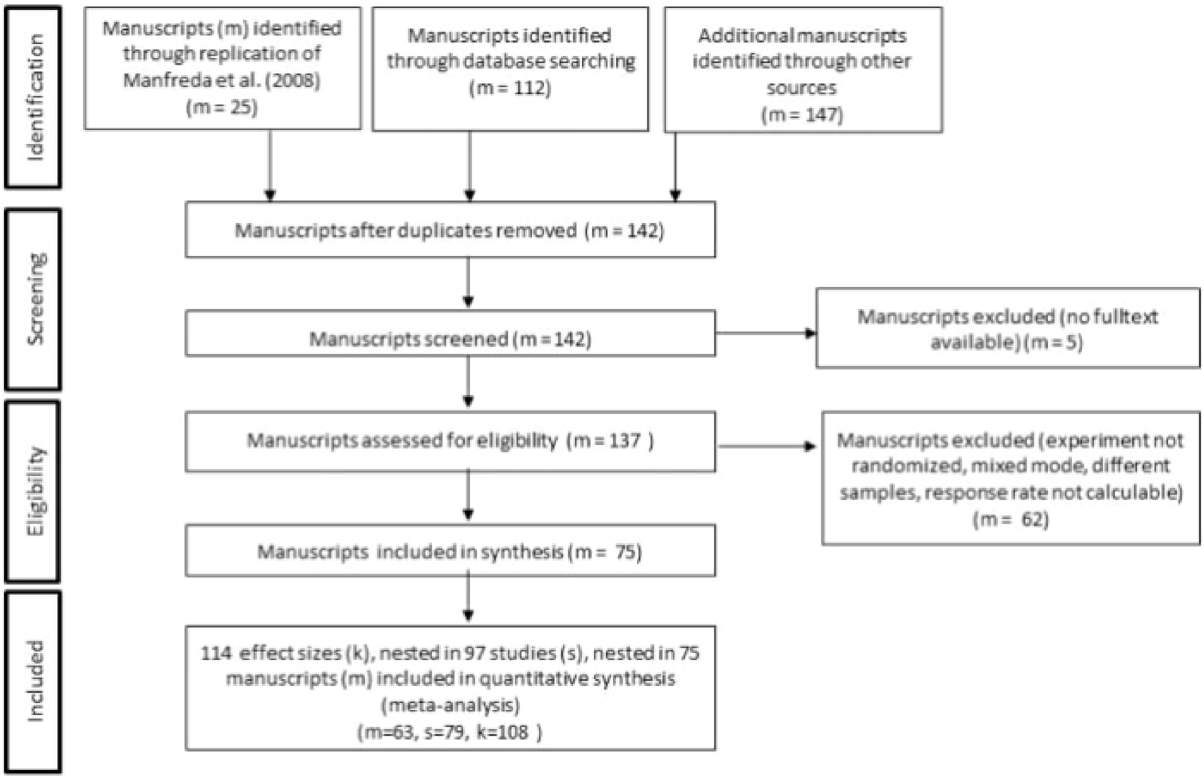

Figure 1. PRISMA 2009 Flow Diagram.
NOTE.— Meanings of different variables: m, manuscripts; s, studies; k, effect sizes. Adapted from Moher et al. (2009).

3.2 Coding Procedures

Coding was performed by two independent coders using the coding sheet (see table A3 of the online supplementary material). The solicitation mode used in the web mode was the only moderator coded for the web mode; all other moderators were applicable to both modes. The second coder was instructed by the first; coding samples were provided. The second coder coded a random sample of one third of the manuscripts, and the intercoder reliability showed a Krippendorff's alpha (Krippendorff 2004) of 0.92, indicating almost a 92 percent agreement between the two coders. As Krippendorff (2004) recommended an alpha value of 0.80 or higher, this is an excellent value.

3.3 Statistical Method

In line with the original meta-analysis by Lozar Manfreda et al. (2008), we calculated the response rate difference, which is our effect size, using raw frequency. Accordingly, we used the number of invited and eligible subjects compared with the number of actual respondents per mode. In most of the included studies, the effective initial sample size was calculated as the initial sample size minus undeliverable and noneligible units. However, raw frequencies are essential for calculating the confidence interval (CI) for each effect size. In those cases with insufficient data, we used the authors' definition of the "response rate" and calculated the raw frequencies. As the authors used the same response rate calculation logic and our effect size was the response rate difference between the two modes rather than the raw response rate, using the authors' definition of response rate and respondent was found to be adequate. In addition, although different survey projects may use different definitions of usable respondents (e.g., those who answered 90 percent of the items, 50 percent of the items, etc.), we relied on the authors' definitions of usable respondents and assumed that they used the same criteria for both modes under comparison. As we were interested only in differences, we found this strategy appropriate. We built a dummy variable based on whether the authors provided the raw frequencies or the response rates only. It showed no significant effect in moderating the average response rate difference.

Our effect size is the response rate difference (RD) between the web mode and the compared mode, which was calculated as follows:

$$RD = \frac{number\ of\ respondents\ web\ mode}{number\ of\ invited\ and\ eligible\ subjects\ web\ mode} - \frac{number\ of\ respondents\ other\ mode}{number\ of\ invited\ and\ eligible\ subjects\ other\ mode}$$

Thus, a positive RD indicates a higher response rate for the web mode, and a negative RD indicates a lower response rate for the web mode compared with the other survey mode.

In general, our statistical analysis comprised five steps (Lipsey and Wilson 2001). First, we computed the weighted mean response rate difference across all studies by weighting each effect size by the inverse of its variance. This variance component consisted of the study-level sampling error variance and an estimate of between-study variance (Borenstein, Hedges, Higgins, and Rothstein 2009). Section "A.1 Measurement Overview" of the online supplementary material provides a description and interpretation of typical meta-analytic measures and further references. Because inference should be made for a population of studies larger than the set of observed studies (Hedges and Vevea 1998), we used a random effects analysis. In the next step, we calculated the confidence interval for the mean effect size to indicate the degree of precision of the estimate and whether the mean effect size was statistically significant. In the third step, we performed a homogeneity analysis to assess whether the effect sizes came from the same population (random effects assumption). In the fourth step, we checked the robustness and quality of our findings by using a sensitivity analysis, an outlier analysis, and a publication bias check. The sensitivity analysis involved first calculating the effect size in a multilevel model by nesting the effect sizes in publications and then calculating the effect size separately for the old and the new studies. In the final analysis step, we conducted a mixed-effect model analysis for each moderator separately to determine which moderators had a significant influence on response rate differences. We used the R package "metafor" (version 1.9–9) for the analyses (Viechtbauer 2010).

4. RESULTS

4.1 Study Characteristics

Following our search strategy and eligibility criteria outlined previously, we identified seventy-five manuscripts (twenty-four from the previous study and fifty-one new manuscripts) that compared the response rates of web and other survey modes using split-sample randomized experimental designs. Because some of these manuscripts contained more than one response rate comparison, 114 response rate comparisons (k) (forty-four from the previous study and seventy new) were included in our study (see table A3 of the online supplementary material).

4.2 Mean Response Rate Difference: Web Surveys Versus Other Survey Modes

The sampling-error-weighted mean effect size estimate, computed across all 114 effect sizes under a random effects assumption, was −0.12 (95 percent CI, −0.16 to −0.09), which favors other survey modes over the web mode (table 1, first line). This result indicates that web surveys yielded, on average, a 12 percentage points lower response rate compared with other survey modes.

How did the response rate difference develop over time? The response rate difference in the Lozar Manfreda et al. (2008) study was about 11 percentage points (95 percent CI, −0.15 to −0.06) lower for web surveys, and this value increased slightly in the present analysis to 12 percentage points (95 percent CI, −0.16 to −0.09). This result emphasizes the tendency of a basically stagnant, albeit slightly increased, response rate difference over time, which is also depicted by a cumulative forest plot (figure 2) that chronologically describes the accumulation of evidence. The cumulative forest plot reveals two trends: First, the effect size becomes more precise over time (confidence intervals for the overall effect become smaller), which indicates a robust, time-invariant estimate that consistently favors other survey modes in terms of response rate differences. Second, the cumulative effect sizes have a slight tendency to the left, which indicates a rising response rate difference. Furthermore, when examining only the new effect sizes (2005–2016), on average, we detected a 15 percentage points lower response rate for web surveys (see table A5 of the online supplementary material). Thus, to answer the first research question, our results indicate that, overall, the response rate difference remained constant over time, with the tendency to increase nonsubstantially in favor of other survey modes.

Is the effect size heterogeneous? A homogeneity analysis for all effect sizes revealed a significant Q-score of 7,501 ($df = 114$, $p < 0.0001$), which indicates the heterogeneity of the effect size distribution under the random effects assumption. This finding called for a moderator analysis to investigate whether moderators influenced the response rate difference (see section 4.3 below).

Table 1. Summary of Random Effects Model: Replicated and Extended Moderators Predicting the Response Rate Difference Between Web and Other Survey Modes

Moderator variable		Meta-analytic summary statistics (random effects model)			Heterogeneity estimators			Mixed-Effect Meta Regression		
	Categories and number of cases	**n**	**Mean response difference(95 % CI)**	**T^2 (se)**	**$Q_{e\ total}$ (df/ p)**	**I^2**	**H^2**	**Model fit R2 %**	**Resid. Hg. Q_m (df)**	**Stat. Sig.**
	Random effects model without moderators	**114**	**−0.12 (−0.16 to −0.09)**	**0.03 (0.004)**	**7446 (112/ <.0001)**	**99.16**	**119.29**			
Moderator analysis: replication										
Compared mode	Email (10)	114	−0.13 (−0.26 to −0.01)	0.04 (0.005)	12,181.12 (109/<.001)	99.31	144.54	0.00	0.34 (3)	.95
	Mail (66)		−0.12 (−0.16 to −0.07)							
	Telephone (19)		−0.14 (−0.24 to −0.04)							
	Other (13)		−0.14 (−0.24 to −0.04)							
Sample recruitment strategy	Panel/ pre-recruited list (10)	114	−0.09 (−0.21 to 0.01)	0.03 (0.005)	14,404.06 (110/<.001)	99.30	142.04	7.14	9.95 (2)	.01
	One-Time Recruitment (34)		−0.16 (−0.22 to −0.11)							
	Existing List (70)		−0.08 (−0.12 to −0.05)							

Continued

Table 1. *Continued*

Moderator variable			Meta-analytic summary statistics (random effects model)		Heterogeneity estimators			Mixed-Effect Meta Regression		
Target population	Students (21)	110	−0.08 (−0.16 to −0.00)	0.03 (0.005)	13,201.72 (110/<.001)	99.26	135.83	3.42	6.48 (3)	.09
	Employees/ Members of Associations (34)		−0.08 (−0.15 to −0.02)							
	Business Respondents (11)		−0.12 (−0.24 to −0.01)							
	General Population (48)		−0.17 (−0.23 to −0.12)							
Type of sponsorship	Academic (78)	114	−0.13 (−0.17 to −0.09)	0.03 (0.005)	12,169.78 (110/<.001)	99.32	147.23	0.00	1.90 (3)	.28
	Governmental (27)		−0.12 (−0.19 to −0.05)							
	Commercial (9)		−0.04 (−0.17 to 0.09)							
Solicitation mode	Mail (61)	113	−0.11 (−0.18 to −0.03)	0.03 (0.00)	13,508.66 (110/<.001)	99.34	5.40	5.40	7.86 (2)	.02
	Email (40)		−0.06 (−0.12 to 0.00)							
	Other (13)		−0.08 (−0.20 to −0.04)							
Incentive	Both modes used incentives (40)	112	−0.15 (−0.21 to −0.09)	0.04 (0.00)	12,276.50 (109/<.001)	99.28	138.12	0.50	1.43 (1)	0.23
	No mode used incentives (69)		−0.10 (−0.14 to −0.06)							

Number of contacts	Continuous	80	−0.04 (−0.07 to −0.02)	0.04 (0.01)	8814.86 (77/<.001)	99.21	126.24	3.46	4.73 (2)	0.04
Publication year	Continuous	113	−0.01 (−0.01 to 0.00)	0.04 (0.00)	14,343.81 (111/<.001)	99.38	161.24	0.00	0.55 (1)	0.46
Moderator analysis: extension										
Survey topic	Public opinion (17)	113	−0.20 (−0.29 to −0.11)	0.03 (0.01)	10,839.21 (107/<.001)	99.3	138.7	0.14	4.08 (4)	0.39
	Professional issue (Job) (32)		−0.12 (−0.19 to −0.07)							
	Technology (12)		−0.14 (−0.21 to −0.02)							
	Lifestyle (19)		−0.14 (−0.24 to −0.03)							
	Other (33)		−0.09 (−0.15 to −0.02)							
Prenotification of survey	Both modes (63)	108	−0.14 (−0.18 to −0.10)	0.03 (0.03)	11,607.04 (105/<.001)	99.3	136.5	3.45	4.68 (1)	0.03
	No mode (45)		−0.05 (−0.10 to 0.00)							
Survey country	US (80)		−0.09 (−.013 to −0.05)	0.03 (0.01)	11,461.94 (96/<.001)	99.1	112.2	8.42	10.45 (2)	0.01
	US (80)		−0.16 (−0.28 to −0.04)							
	NL (6)		−0.16 (−0.31 to −0.09)							

Resid. Hg. = Residual heterogeneity, Stat. Sig. = Statistical significance.

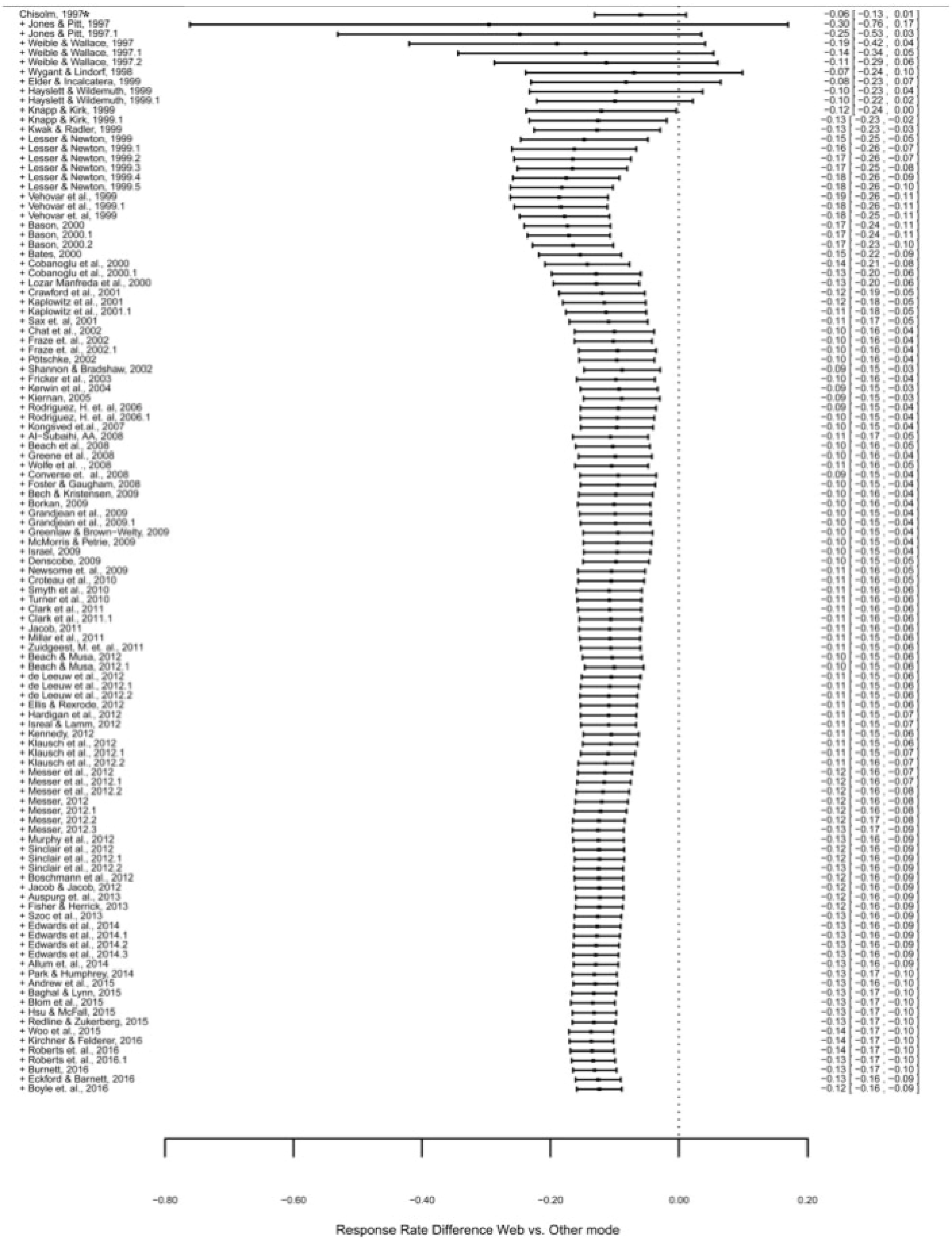

Figure 2. Cumulative Forest Plot.
*Fieldwork year (where given)

Before conducting this analysis, we addressed two questions regarding the validity of the findings: publication bias and robustness. Publication bias refers to the problem that significant results have a higher probability of being published and may distort the results. Sensitivity analyses did not identify publication bias in our data. We also performed several robustness checks such as excluding the outliers, performing separate analyses for the old and the newer studies, and applying a multilevel approach for effect sizes nested in articles. All the mean response rate differences pointed in the same direction, and no significant differences could be detected. This suggests a robust overall effect size in terms of magnitude and direction. A detailed description of the validity testing is provided in section A2 of the online supplementary material.

4.3 Moderator Analysis: Replication

This section presents the results for the moderators. First, the response rate difference was regressed on the survey mode to which web surveys were compared, the sample recruitment strategy, the target population, the type of sponsorship, the solicitation mode, the use of incentives, and the number of contacts—which are all the moderators included in the first meta-analysis (Lozar Manfreda et al. 2008). Second, we extended the original analysis by adding three new moderators: survey topic, prenotification, and survey country. Table 1 provides the results of the separate analyses that investigated the influence of moderators on the response rate difference between web and other survey modes. As indicated in the last column of table 1 (replication part), three of the six moderators—sample recruitment strategy, solicitation mode, and number of contacts—significantly explain the response rate difference ($p \leq 0.05$). All three moderators produced significant effects in the original meta-analysis, as well. The average response rate difference for panel members or respondents from an existing list was 9 percentage points lower for the web mode. This difference increased to 21 percentage points for one-time respondents (see figure 3). A second influential moderator was the solicitation mode: If participation was initially requested by a mode other than email, the response rate for web surveys was at least 8 percentage points lower. However, if respondents were asked by email to participate, this difference shrank, on

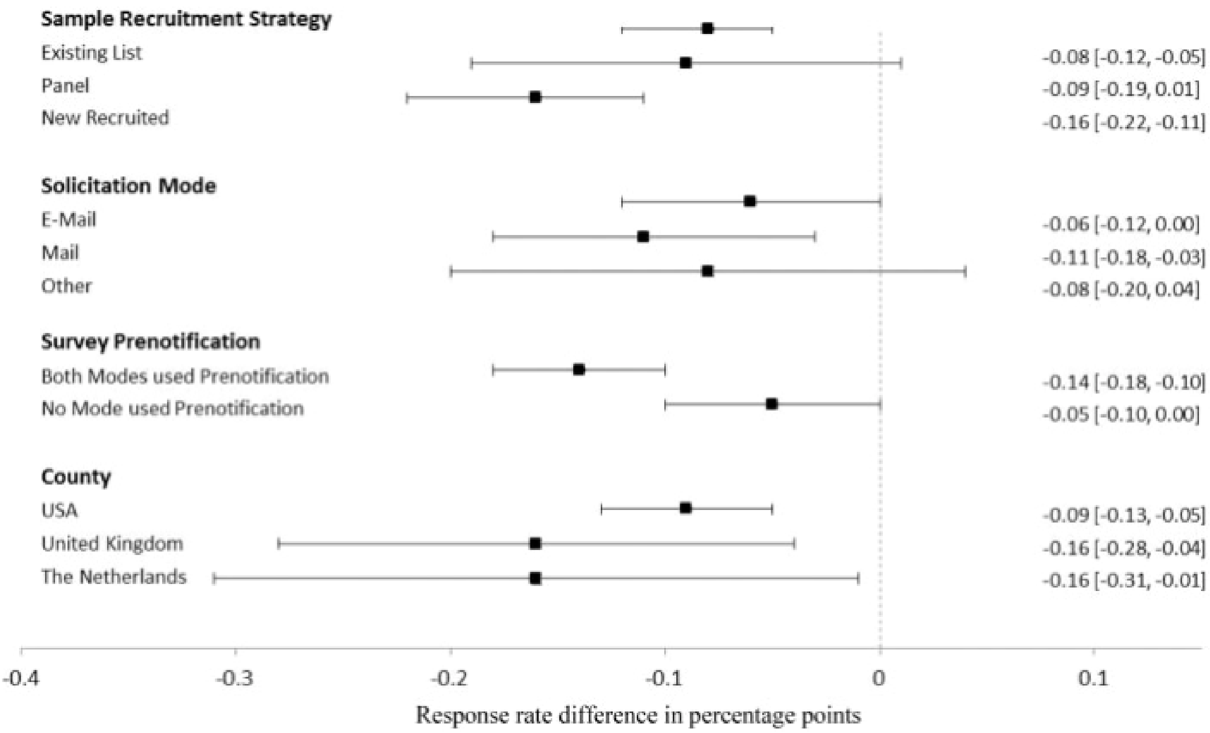

Figure 3. Forest Plot of Significant Categorical Moderators.
NOTE.— Number of contact attempts (continuous variable) was also significant.

average, to a difference of 6 percentage points. The final study characteristic that significantly influenced the response rate difference was the number of contact attempts. The results suggest that the larger the number of contacts was, the larger the response rate difference became (3 percentage points difference for each contact attempt). This result suggests that contact attempts are less effective in the web survey mode. With regard to the target population, our findings indicate that specific populations showed only a small difference in response rates (students and employees: 8 percentage points; business respondents: 12 percentage points), whereas the difference between the web mode and the compared mode in surveys of the general population increased distinctly ($p \leq 0.10$).

In summary and to answer the second research question, as in the original Lozar Manfreda et al. (2008) meta-analysis, the sample recruitment base, solicitation mode, and number of contacts were found to have a significant effect on explaining the response rate difference between web and other survey modes. Contrary to the original study, the type of target population was significant on the 10 percent level. However, the compared mode, the use of incentives, the type of sponsorship, and the year the studies were conducted did not significantly explain the response rate difference.

4.4 Moderator Analysis: Extension

This section presents an extension of the moderator analysis by including three new moderators that were not assessed in Lozar Manfreda et al. (2008) meta-analysis (see lower part of table 1). Significant effects were observed for two of these three moderators. First, when prenotifications were used, this strategy was more effective in other survey modes than in web surveys. The use of prenotifications increased the response difference to 15 percentage points (see figure 3). This result suggests that survey prenotifications are more effective in any mode other than the web survey mode. This result is in line with our expectations that an email prenotification for a web survey is perceived by target persons to be less important because it involves minor effort on the part of the researcher. The second significant new moderator is the country in which the survey was conducted. Response rates for web surveys in the United States were, on average, only 9 percentage points lower than for other survey modes; this figure rose to 16 percentage points for the United Kingdom and the Netherlands. We had to exclude other countries (Australia, Canada, Germany, Slovenia, and Sweden) from the analysis because less than five experiments in these countries were included in our meta-analysis, and the results would therefore have had little informative value. Providing a summary of the countries in geographical groups made little sense to us at this point, as attitudes to the World Wide Web cannot necessarily be delimited by geographical or continental borders. However, we tested geographically related and value-related

Table 2. Overview of Study Design Characteristics

Moderator Variable	Moderator had a significant influence on the response rate difference in the . . .	
	2008 meta-analysis	present meta-analysis
Compared mode	not significant	not significant
Sample recruitment strategy	significant	significant
Target population	not significant	significant
Type of sponsorship	not significant	not significant
Solicitation mode	significant	significant
Incentive	not significant	not significant
Number of contacts (categorical)	significant	significant
Publication year	not significant	not significant
Survey country	–	significant
Prenotification of Survey	–	significant
Survey Topic	–	not significant

(Hofstede 1980) categories, and the effects turned out to be very robust. As a result, the mode decision in the United States should favor web surveys, whereas a much higher response rate difference is to be expected in the United Kingdom and the Netherlands. Nevertheless, it is important to point out that a low response rate difference can result from a particularly good performance of the web mode or from a poor performance of the comparison mode.

To answer our fourth research question, our findings show that the use of prenotifications and the country in which the survey is conducted significantly impacted survey response rates, whereas the survey topic did not. For the latter, it should be noted that we could not classify survey topics on the basis of their sensitivity. Following the relevant literature, this classification would have been particularly useful, as online respondents have been found to be more willing to disclose information on sensitive topics (Kreuter et al. 2008). Table 2 provides an overview of all survey design characteristics and their development over time observed in the original metal-analysis and in our replication study.

5. DISCUSSION

Prior to the present study, the last meta-analysis on response rate differences between web surveys and other survey modes was conducted more than a decade ago (Lozar Manfreda et al. 2008). Since then, the status and relevance of the web mode has changed. We examined these changes by including in our meta-analysis over 100 experiments related to response rate differences between web surveys and other modes. Overall, we found a basically stagnant heterogeneous mean response rate difference of 12 percentage points. Consequently, by choosing a web survey mode, researchers run the risk of achieving lower response rates than in traditional modes. Two groups of reasons can be used to explain this finding: long-term generic and contextual. The first group includes the lower perceived legitimacy of web surveys. Respondents may consider researchers' efforts to be less substantial—for example, "merely" sending an email compared with more time-consuming contact by telephone, where a researcher calls respondents once or even several times. The greater effort on the part of the researcher and the personal contact make it more difficult for the respondent to refuse to participate. Furthermore, the literature suggests that respondents perceive web surveys to be less mandatory, and web survey requests via email are often overlooked or routed to spam filters before they are read (Dillman et al. 2014, p. 419). Contextual reasons include increased web over-surveying. Because web surveys are quicker and cheaper, they are often used for surveys with limited resources. Furthermore, they now constitute the most popular survey mode worldwide (ESOMAR 2018). Moreover, respondents may receive a large amount of spam emails and find it difficult to distinguish between those that are relevant and those that are not. Other contextual reasons may be the greater sensitivity about security and privacy on the internet (Marreiros, Tonin, and Vlassopoulos 2016), especially in Europe since the General Data Protection Regulation became applicable in May 2018 (European Commission 2018) and because of the great diversity of internet browsers, devices (including mobile), and operating systems that can cause technical incompatibility problems.

In addition to studying change over time, the second and third aims of the present study were to increase the statistical power of the moderator analyses and to identify further influencing factors, especially as the mean effect size is heterogeneous. In the original study, Lozar Manfreda et al. (2008) demonstrated how the sample recruitment base, solicitation mode, and number of contacts significantly influenced the response rate difference. Our research corroborates these findings and revealed other significant moderators. More specifically, we found that using a prenotification was more effective in all survey modes except the web mode, which confirms our assumption in this regard. People are more likely to overlook a prenotification via email than via traditional communication channels (Crawford, Couper, and Lamias 2001). One can argue that, in traditional survey modes, a researcher's investment in multiple contacts is perceived by the respondents to be an indication of the importance and legitimacy of the survey (Tuten 1997; Evans and Mathur 2005). Considerably more work is necessary to fully understand this phenomenon. Another significant predictor of the response rate difference is the survey country. Surveys conducted in the United States produce higher web response rates or lower response rates in other modes, which results in a lower response rate difference overall. This suggests that the nonresponse problem in web surveys is lowest in the United States. More research is needed to better understand the significant differences across countries and to determine the specific factors responsible for response rate differences at country level.

The present findings have substantial implications for the choice of survey mode. They offer cumulative evidence about the survey-environment factors that improve response rates in web surveys. To narrow the gap between response rates in web surveys and other survey modes, we therefore recommend forgoing the prenotification of web surveys and instead using email solicitation and between one and two contact attempts. In an ideal case, the sample consists of panel respondents from a specific population in the United States.

5.1 Limitations and Further Research

Changes in the web, and particularly in mobile technology, suggest that further meta-analyses should take into account different devices used to answer web questionnaires and the way in which they may be affecting response rates and, even more importantly, differences in nonresponse bias. Thus, replications of the present cumulative meta-analysis to further track changes over time should include a mobile devices dimension. The second limitation of the present study is that it does not account for the absolute response rate level. Although the response rate difference is small, it still ignores whether the absolute response rate was high (or low) in general across all modes. To gain further evidence about the absolute web response level and its moderators, we strongly recommend that meta-analytical research be carried out in this regard.

The third limitation of this meta-analysis is that we estimated a large number of moderator models. Our findings could therefore be affected by the possibility of capitalizing on chance (rejecting a true null hypothesis). This means that some of the moderators in this meta-analysis may have shown significant results only by chance. Although a Bonferroni correction could remedy this, it is not recommended for power reasons (Schmidt and Hunter 2014, p. 83; Polanin and Pigott 2015). The fourth limitation of our meta-analysis is the fact that we could not address the critical issue of breakoff rates, because the breakoff rates in the web surveys and the compared modes were only occasionally reported, and a meta-analytical consideration of this topic was not possible. As the literature indicates that breakoff rates in academic or governmental web surveys are higher than in other response modes and range between 14 and 35 percent (Musch and Reips 2000; Lozar Manfreda and Vehovar 2002; Peytchev 2009; McGonagle 2013), one cause for the response rate gap could be different breakoff probabilities. However, more research is needed on this issue.

Fifth, our study does not address whether the nonresponse rate is an indicator of nonresponse error and nonresponse bias. A low response rate does not necessarily lead to high nonresponse error, as the latter refers to the differences in the statistics between respondents and nonrespondents. Nonresponse error occurs if the nonrespondents—if they had responded—would have provided different answers than the actual respondents. Several studies have actually shown that low response rates do not necessarily indicate large nonresponse error (Keeter, Miller, Kohut, Groves, and Presser 2000; Groves and Peytcheva 2008). However, a high response rate usually minimizes the probability that nonrespondents affect survey results, which is why we believe that the present study adds new knowledge of relevance for understanding nonresponse error.

However, further meta-analytic research needs to be done to establish whether these findings hold for other measures of web survey data quality—first, representation-related errors and biases (as representativeness indicators) and second, measurement-related quality indicators (such as item nonresponse, consistency of answers, richness of responses to open-ended questions, speed of answering, acquiescence, social desirability, breakoff, and conditioning effects). With respect to data quality, if it can be shown that responses from web survey modes are comparable to the responses from other survey modes, the problem of lower response rates in web surveys would not be as critical, particularly when one takes into account that fewer resources are needed to conduct web surveys. The present study did not consider that web surveys are usually cheaper to conduct compared with traditional modes. One could argue that the money saved by conducting a web survey can be used to produce better data quality and reduce the response rate difference, for example, by incentivizing reluctant respondents.

Related to this, it should be emphasized that inspection of the cumulative forest plot (figure 3) reveals that, starting in 2002, the response rate difference did not change substantially. Therefore, experiments that simply compared the response rate difference across different survey modes could have stopped then. Instead, more effort should have been invested in exploring the mechanisms that induce web survey participation. Thus, further research should focus primarily on the value of the web mode: How can the value of online surveys be increased, taking account the variety of data quality indicators and the latest developments in mobile web surveys?

Supplementary Materials

Supplementary materials are available online at academic.oup.com/jssam.

Studies inserted in this meta-analysis

Al Baghal, T., and P. Lynn (2015), "Using Motivational Statements in Web-Instrument Design to Reduce Item-Missing Rates in a Mixed-Mode Context," *Public Opinion Quarterly*, 79, 568–579.

Allum, N., K. Auspurg, M. Blake, C. L. Booker, T. F. Crossley, J. D'Ardenne, and M. Fairbrother (2014), "Understanding Society Innovation Panel Wave 6: Results from Methodological Experiments," Understanding Society Working Paper Series, 2014-04, available at http://www.websm.org/uploadi/editor/1412860905Allum_2014_Understanding_Society_Innovation_Panel_Wave_6.pdf.

Al-Razgan, M. S., H. S. Al-Khalifa, M. D. Al-Shahrani, and H. H. AlAjmi (2012), "Touch-Based Mobile Phone Interface Guidelines and Design Recommendations for Elderly People: A Survey of the Literature," in *Neural Information Processing. ICONIP 2012. Lecture Notes in Computer Science*, ed. T. Huang, Z. Zeng, C. Li, and C.S. Leung, vol. 7666, 568–574, Berlin, Heidelberg: Springer.

Al-Subaihi, A. A. (2008), "Comparison of Web and Telephone Survey Response Rates in Saudi Arabia," *Electronic Journal of Business Research Methods*, 6, 123–132.

Auspurg, K., J. C. A. Burton, L. Cullinane Delavande Fumagalli, M. Iacovou, A. Jäckle, O. Kaminska, P. Lynn, et al. (2013), "Understanding Society Innovation Panel Wave 5: Results from Methodological Experiments," Understanding Society Working Paper Series, 2013-06, available at http://kops.uni-konstanz.de/bitstream/handle/123456789/26763/Auspurg_267631.pdf?sequence=2&isAllowed=y.

Bales, G. T., C. M. Hollowell, R. V. Patel, and G. S. Gerber (2000), "Internet and Postal Survey of Endourologic Practice Patterns Among American Urologists," *The Journal of Urology*, 163, 1779–1782.

Bason, J. (2000), "Comparison of Telephone, Mail, Web, and Ivr Surveys of Drug and Alcohol Use among University of Georgia Students," 55th Annual Conference of the American Association for Public Opinion Research, Washington D.C., available at http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=4A048F09579917B541C3393AD83C2B8E?doi=10.1.1.608.1652&rep=rep1&type=pdf.

Beach, S., D. Musa, P. Beeson, and C. Sparks (2008), "Mode Effects and Non-Response Bias in an Undergraduate Student Satisfaction Survey: Results from a Randomized Experiment Comparing Telephone and Web Administration," The American Association for (AAPOR) 63rd Annual Conference, 2008 & WAPOR 61th Annual Conference, New Orleans.

Bech, M., and M. B. Kristensen (2009), "Differential Response Rates in Postal and Web-Based Surveys among Older Respondents," *Survey Research Methods*, 3, 1–6.

Blom, A. G., J. Burton, C. Booker, A. Cernat, M. Fairbrother, V. Annette, O. Kaminska, et al. (2015), "Understanding Society Innovation Panel Wave 7: Results from Methodological Experiments," Essex, available at http://www.academia.edu/download/37875387/2015-03.pdf

Borkan, B. (2009), "The Mode Effect in Mixed-Mode Surveys: Mail and Web Surveys," *Social Science Computer Review*, 28, 371–380.

Boschman, J. S., H. F. van der Molen, M. H. W. Frings-Dresen, and J. K. Sluiter (2012), "Response Rate of Bricklayers and Supervisors on an Internet or a Paper-and-Pencil Questionnaire," *International Journal of Industrial Ergonomics*, 42, 178–182.

Boyle, K. J., M. Morrison, D. H. MacDonald, R. Duncan, and J. Rose (2016), "Investigating Internet and Mail Implementation of Stated-Preference Surveys While Controlling for Differences in Sample Frames," *Environmental and Resource Economics*, 64, 401–419.

Burnett, C. M. (2016), "Exploring the Difference in Participants' Factual Knowledge Between Online and in-Person Survey Modes," *Research & Politics*, 3, 1–7.

Cernat, A., M. P. Couper, and M. B. Ofstedal (2016), "Estimation of Mode Effects in the Health and Retirement Study Using Measurement Models," *Journal of Survey Statistics and Methodology*, 4, 501–524.

Chatt, C., M. Dennis, R. Li, A. Motta-Stanko, and P. Pulliam (2005), "Data Collection Mode Effects Controlling for Sample Origins in a Panel Survey: Telephone versus Internet," 60th Annual Conference of the American Association for Public Opinion Research, *Miami Beach, FL*. 2005.

Chisholm, J. (1999), "Using the Internet to measure customer satisfaction and loyalty," in Best Practices in Customer, ed. Ron Zemke and John A. Woods, pp. 305–317, New York: AMACOM.

Clark, M., M. Rogers, A. Foster, F. Dvorchak, F. Saadeh, J. Weaver, and V. Mor (2011), “A Randomized Trial of the Impact of Survey Design Characteristics on Response Rates among Nursing Home Providers,” *Evaluation Health Prof*, 34, 464–486.

Cobanoglu, C., B. Warde, and P. J. Moreo (2001), “A Comparison of Mail, Fax and Web-Based Survey Methods,” *International Journal of Market Research*, 43, 1–410.

Cole, S.-T. (2005), “Comparing Mail and Web-Based Survey Distribution Methods: Results of Surveys to Leisure Travel Retailers,” *Journal of Travel Research*, 43, 422–430.

Converse, P. D., E. W. Wolfe, X. Huang, and F. L. Oswald (2008), “Response Rates for Mixed-Mode Surveys Using Mail and E-Mail/Web,” *American Journal of Evaluation*, 29, 99–107.

Crawford, S. D., Mick P. Couper, and Mark J. Lamias (2001) “Web surveys: Perceptions of burden.” *Social science computer review*, 19.2, 146–162.

Crawford, S., S. McCabe, M. Couper, and C. Boyd (2002), “From Mail to Web: Improving Response Rates and Data Collection Efficiencies,” International Conference on Improving Surveys, pp. 25–28. available at http://www.websm.org/uploadi/editor/improving%20response%20rates%20and%20data%20collection.pdf

Croteau, A.-M., L. Dyer, and M. Miguel (2010), “Employee Reactions to Paper and Electronic Surveys: An Experimental Comparison,” *IEEE Transactions on Professional Communication*, 53, 249–259.

De Leeuw, E., G. Nicolaas, P. Campanelli, and J. Hox (2012), “Question or Mode Effects in Mixed-Mode Surveys: A Cross-Cultural Study in The Netherlands, Germany, and the UK,” The American Association for Public Opinion Research (AAPOR) 67th Annual Conference, 2012, Orlando Florida.

Denscombe, M. (2009), “Item Non-Response Rates: A Comparison of Online and Paper Questionnaires,” *International Journal of Social Research Methodology*, 12, 281–291.

Eckford, R. D., and D. L. Barnett (2016), “Comparing Paper-and-Pencil and Internet Survey Methods Conducted in a Combat-Deployed Environment,” *Military Psychology*, 28, 209–225.

Edwards, M. L., D. A. Dillman, and J. D. Smyth (2014), “An Experimental Test of the Effects of Survey Sponsorship on Internet and Mail Survey Response,” *Public Opinion Quarterly*, 78, 734–750.

Elder, A., and T. Incalcatera (2000), “Pushing the Envelope. Moving a Major Syndicated Study to the Web,” *Net Effects*, 3.

Ellis, J. M., and D. L. Rexrode (2012), “Addressed-Based Sampling—A Better Sample? Exploring the Benefits of Using Addressed-Based Sampling in a State-Wide Targeted Sub-Population,” http://www.websm.org/index.php?fl=2&lact=8&vir=1197 The American Association for Public Opinion Research (AAPOR) 67th Annual Conference, 2012.

Fisher, S. H., and R. Herrick (2013), “Old Versus New: The Comparative Efficiency of Mail and Internet Surveys of State Legislators,” *State Politics & Policy Quarterly*, 13, 147–163.

Foster, K. N., and M. Gaughan (2008), “Examining Response Rates and Patterns in a Multimode Experiment: A Study of Department Chairs/Heads in Stem Programs at Research Intensive Universities,” The American Association for Public Opinion Research (AAPOR) 63rd Annual Conference, Phoenix.

Fraze, S., K. Hardin, T. Brashears, J. Haygood, and J. Smith (2003), “The Effects of Delivery Mode upon Survey Response Rate and Perceived Attitudes of Texas Agri-Science Teachers,” *Journal of Agricultural Education*, 44, 27–37.

Fricker, S., M. Galesic, R. Tourangeau, and T. Yan (2005), “An Experimental Comparison of Web and Telephone Surveys,” *Public Opinion Quarterly*, 69, 370–392.

Grandjean, B. D., N. M. Nelson, and P. A. Taylor (2009), “Comparing an Internet Panel Survey to Mail and Phone Surveys on Willingness to Pay for Environmental Quality: A National Mode Test,” The American Association for Public Opinion Research (AAPOR) 64th Annual Conference, 2009, Hollywood.

Greene, J., H. Speizer, and W. Wiitala (2008), “Telephone and Web: Mixed-Mode Challenge,” *Health Services Research*, 43, 230–248.

Greenlaw, C., and S. Brown-Welty (2009), “*A Comparison of Web-Based and Paper-Based Survey Methods Testing Assumptions of Survey Mode and Response Cost*,” Evaluation Review 33, 464–480.

Grigorian, K., S. Sederstrom, and T. Hoffer (2004), "Web of Intrigue? Evaluating Effects on Response Rates of Between Web Saq, Cati, and Mail Saq Options in a National Panel Survey," American Association for Public Opinion Research, 2004, Phoenix.

Hardigan, P. C., C. T. Succar, and J. M. Fleisher (2012), "An Analysis of Response Rate and Economic Costs between Mail and Web-Based Surveys among Practicing Dentists: A Randomized Trial," *Journal of Community Health*, 37, 383–394.

Hayslett, M., and B. Wildemuth (2005), "Pixels or Pencils? The Relative Effectiveness of Web-Based Versus Paper Surveys," *Library & Information Science Research*, 26, 73–93.

Heerwegh, D., and G. Loosveldt (2008), "Face-to-Face Versus Web Surveying in a High-Internet-Coverage Population Differences in Response Quality," *Public Opinion Quarterly*, 72, 836–846.

Israel, G. D. "Using Mixed-Mode Contacts to Facilitate Participation in Public Agency Client Surveys," paper presented at the AAPOR Annual Conference, Boston, MA, USA.

Jacob, R. T. (2011), "An Experiment to Test the Feasibility and Quality of a Web-Based Questionnaire of Teachers," *Evaluation Review*, 35, 40–70.

Jones, M., G. Marsden, N. Mohd-Nasir, K. Boone, and G. Buchanan (1999), "Improving Web Interaction on Small Displays," *Computer Networks*, 31, 1129–1137.

Jones, R., and N. Pitt (1999), "Health Surveys in the Workplace: Comparison of Postal, Email and World Wide Web Methods," *Occupational Medicine (Oxford England)*, 49, 556–558.

Kaplowitz, M. D., T. D. Hadlock, and R. Levine (2001), "A Comparison of Web and Mail Survey Response Rates," *Public Opinion Quarterly*, 68, 94–101.

Keeter, S., C. Miller, A. Kohut, R. M. Groves, and S. Presser (2000), "Consequences of Reducing Nonresponse in a National Telephone Survey," *Public Opinion Quarterly*, 64, 125–148.

Kennedy, C., A. Ackermann, C. Turakhia, M. O. Emerson, and A. James (2012), "Mode Effects Measurement and Correction: A Case Study." The American Association for Public Opinion Research (AAPOR) 67th Annual Conference, available at http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2012/01_Kennedy_E2.pdf

Kerwin, J., P. D. Brick, K. Levin, D. Cantor, J. O'Brien, A. Wang, and S.-S. Stephanie (2004), "Web, Mail, and Mixed-Mode Data Collection in a Survey of Advanced Technology Program Applicants," 2004 Joint Statistical Meetings.

Kiernan, N. E., M. Kiernan, M. A. Oyler, and C. Gilles (2005), "Is a Web Survey as Effective as a Mail Survey? A Field Experiment among Computer Users," *American Journal of Evaluation*, 26, 245–252.

Kim, Y., J. Dykema, J. Stevenson, P. Black, and D. P. Moberg (2018), "Straightlining: Overview of Measurement, Comparison of Indicators, and Effects in Mail–Web Mixed-Mode Surveys," *Social Science Computer Review*, 37, 214–233.

Kirchner, A., and B. Felderer (2016), "The Effect of Nonresponse and Measurement Error on Wage Regression across Survey Modes: A Validation Study," in *Total Survey Error in Practice*, ed. Biemer, P. P., De Leeuw, E., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L. E., Tucker, N. C. & West, B. T., pp. 531–556, Hoboken: John Wiley & Sons.

Knapp, H., and S. Kirk (2003), "Using Pencil and Paper, Internet and Touch-Tone Phones for Self-Administered Surveys: Does Methodology Matter?," *Computers in Human Behavior*, 19, 117–134.

Kongsved, S. M., M. Basnov, K. Holm-Christensen, and N. H. Hjollund (2007), "Response Rate and Completeness of Questionnaires: A Randomized Study of Internet versus Paper-and-Pencil Versions," *Journal of Medical Internet Research*, 9, 1–10.

Kreuter, F., S. Presser, and R. Tourangeau (2008), "Social Desirability Bias in CATI, IVR, and Web Surveys," *Public Opinion Quarterly*, 72, 847–865.

Kwak, N., and B. Radler (2002), "A Comparison Between Mail and Web Surveys: Response Pattern, Respondent Profile, and Data Quality," *Journal of Official Statistics*, 18, 257–273.

Lesser, V., and L. Newton (2001), "Mail, Email and Web Surveys: A Cost and Response Rate Comparison in a Study of Undergraduate Research Activity," AAPOR Annual Conference, Montreal, Quebec.

Lozar Manfreda, K., and V. Vehovar (2002), "Survey Design Features Influencing Response Rates in Web Surveys," in The International Conference on Improving Surveys Proceedings, Citeseer, pp. 25–28.

Lozar Manfreda, K., V. Vehovar, and Z. Batagelj (2000), "Web versus Mail Questionnaire for an Institutional Survey". Paper presented at 'The Challenge of the Internet', organized by Association for Survey Computing, Latimer Conference Centre, UK, May 11–12, 2001.

McGonagle, K. A. (2013), "Survey Breakoffs in a Computer-Assisted Telephone Interview," in *Survey Research Methods*, NIH Public Access, 79–90, available at https://ojs.ub.uni-konstanz.de/srm/article/download/5126/4997

McMorris, B. J., R. S. Petrie, R. F. Catalano, C. B. Fleming, K. P. Haggerty, and R. D. Abbott (2009), "Use of Web and in-Person Survey Modes to Gather Data from Young Adults on Sex and Drug Use an Evaluation of Cost, Time, and Survey Error Based on A," *Evaluation Review*, 33, 138–158.

Mercer, A., A. Caporaso, D. Cantor, and R. Townsend (2015a), "How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys," Public Opinion Quarterly, 79, 105–129.

Mercer, A., K. McGeeney, N. Mathiowetz, R. Igielnik, and S. Keeter (2015b), "From Telephone to the Web: The Challenge of Mode of Interview Effects in Public Opinion Polls," The American Association for Public Opinion Research (AAPOR) 70th Annual Conference.

Messer, B. L. (2012), "Pushing Households to the Web: Experiments of a 'Web+Mail' Methodology for Conducting General Public Surveys." PhD. Dissertation, Washington State University, Pullman, 2012, available at https://research.wsulibs.wsu.edu:8443/xmlui/bitstream/handle/2376/4653/Messer_wsu_0251E_10588.pdf?sequence=1

Millar, M. M., D. A. Dillman, B. Messer, S. Genter, M. Williams, and T. Allen (2011), "Improving Response to Web and Mixed-Mode Surveys," *Public Opinion Quarterly*, 75, 249–269.

Newsome, J., K. Levin, P. D. Brick, P. Langetieg, M. Vigil, and M. Sebastiani (2009), "Multi-Mode Survey Administration: Does Offering Multiple Modes at Once Depress Response Rates?," 68th *Annual Meeting of the American Association for Public Opinion Research*, 2013, Boston.

Park, A., and A. Humphrey (2014), "Mixed-Mode Surveys of the General Population—Results from the European Social Survey Mixed-Mode Experiment," *NatCen Social Research*. Report. http://www.websm.org/uploadi/editor/1416295092Park_Humphrey_2014_Mixed_Mode_Survey_Of_The_General_Population.pdf

Patrick, M. E., M. P. Couper, V. B. Laetz, J. E. Schulenberg, P. M. O'Malley, L. D. Johnston, and R. A. Miech (2018), "A Sequential Mixed-Mode Experiment in the U.S. national Monitoring the Future Study," *Journal of Survey Statistics and Methodology*, 6, 72–97.

Pötschke, M. (2002), "Paper and Pencil or Online? Methodological Experiences from an Employee Survey," German Online Research Conference (GOR 2002), Stuttgart.

Roberts, C., D. Joye, and M.-E. Stähli (2016), "Mixing Modes of Data Collection in Swiss Social Surveys: Methodological Report of the Lives-FORS Mixed Mode Experiment," LIVES Working Papers, Lausanne.

Rodriguez, H. P., T. von Glahn, W. H. Rogers, H. Chang, G. Fanjiang, and D. G. Safran (2006), "Evaluating Patients' Experiences with Individual Physicians: A Randomized Trial of Mail, Internet, and Interactive Voice Response Telephone Administration of Surveys," *Medical Care*, 44, 167–174.

Sax, L. J., S. K. Gilmartin, and A. N. Bryant (2001), "Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys," *Research in Higher Education*, 44, 409–432.

Shannon, D. M., and C. C. Bradshaw (2002), "A Comparison of Response Rate, Response Time, and Costs of Mail and Electronic Surveys," *The Journal of Experimental Education*, 70, 179–192.

Sinclair, M., J. O'Toole, M. Malawaraarachchi, and K. Leder (2012), "Comparison of Response Rates and Cost-Effectiveness for a Community-Based Survey: Postal, Internet and Telephone Modes with Generic or Personalised Recruitment Approaches," *BMC Medical Research Methodology*, 12, 132.

Vehovar, V., K. Lozar Manfreda, and Z. Batagelj (2001), “Sensitivity of Electronic Commerce Measurement to the Survey Instrument,” *International Journal of Electronic Commerce*, 6, 31–51.

Viechtbauer, W. (2010), “Conducting Meta-Analyses in R with the Metafor Package,” *Journal of Statistical Software*, 36, 1–48.

Weible, R., and J. Wallace (1998), “Cyber Research: The Impact of the Internet on Data Collection,” *Marketing Research*, 10, 18–31.

Wolfe, E. W., P. D. Converse, and F. L. Oswald (2008), “Item-Level Nonresponse Rates in an Attitudinal Survey of Teachers Delivered via Mail and Web,” *Journal of Computer-Mediated Communication*, 14, 35–66.

Woo, Y., S. Kim, and M. P. Couper (2015), “Comparing a Cell Phone Survey and a Web Survey of University Students,” *Social Science Computer Review*, 33, 399–410.

Yeager, D. S., J. A. Krosnick, L. Chang, H. S. Javitz, M. S. Levendusky, A. Simpser, and R. Wang (2011), “Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples,” *Public Opinion Quarterly*, 75, 709–747.

Zuidgeest, M., M. Hendriks, L. Koopman, P. Spreeuwenberg, and J. Rademakers (2011), “A Comparison of a Postal Survey and Mixed-Mode Survey Using a Questionnaire on Patients’ Experiences with Breast Care,” *Journal of Medical Internet Research*, 13, 1–13.

References

Aichele, C., R. Flickenger, C. Fonda, J. Forster, I. Howard, T. Krag, and M. Zennaro (2013), *Wireless Networking in the Developing World*, Morrisville, NC: Limehouse Book Sprint Team.

Borenstein, M., L. V. Hedges, J. Higgins, and H. R. Rothstein (2009), *Introduction to Meta-Analysis*, Hoboken, CA: Wiley Online Library.

Callegaro, M., K. Lozar Manfreda, and V. Vehovar (2015), *Web Survey Methodology*, London: Sage.

Chang, L., and J. A. Krosnick (2009), “National Surveys via RDD Telephone Interviewing versus the Internet: Comparing Sample Representativeness and Response Quality,” *Public Opinion Quarterly*, 73, 641–678.

Cornesse, C., and M. Bosnjak (2018), “Is there an Association Between Survey Characteristics and Representativeness? A Meta-Analysis,” *Survey Research Methods*, 12, 1–13.

Couper, M. P., and E. D. De Leeuw (2003), “Nonresponse in Cross-Cultural and Cross-National Surveys,” *in Cross-cultural survey methods*, ed. Harkness, J.A., van de Vijver, F.J.R., Mohler, P., pp. 157–177, Hoboken, New Jersey: Wiley-Interscience.

Couper, M. P., and G. J. Peterson (2017), “Why Do Web Surveys Take Longer on Smartphones?,” *Social Science Computer Review*, 35, 357–377.

Dillman, D. A., J. D. Smyth, and L. M. Christian (2014), *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, Hoboken, New Jersey: John Wiley & Sons.

Eshet-Alkalai, Y., and E. Chajut (2010), “You Can Teach Old Dogs New Tricks: The Factors That Affect Changes Over Time in Digital Literacy,” *Journal of Information Technology Education: Research*, 9, 173–181.

ESOMAR (2018), “Global Market Research Report 2018: An ESOMAR Industry Report.”.

European Commission (2018), “2018 Reform of Eu Data Protection Rules.” https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en.

Evans, J. R., and A. Mathur (2005), “The Value of Online Surveys,” *Internet Research*, 15, 195–219.

Fan, W., and Z. Yan (2010), “Factors Affecting Response Rates of the Web Survey: A Systematic Review,” *Computers in Human Behavior*, 26, 132–139.

Groves, R. M., F. J. Fowler Jr, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau (2011), *Survey Methodology*, p. 561, Hoboken, New Jersey: John Wiley & Sons.

Groves, R. M., and E. Peytcheva (2008), “The Impact of Nonresponse Rates on Nonresponse Bias a Meta-Analysis,” *Public Opinion Quarterly*, 72, 167–189.

Hedges, L. V., and J. L. Vevea (1998), “Fixed-and Random-Effects Models in Meta-Analysis,” *Psychological Methods*, 3, 486.

Higgins, J., and S. Thompson (2002), “Quantifying Heterogeneity in a Meta-Analysis,” *Statistics in Medicine*, 21, 1539–1558.

Hofstede, G. (1980), *Culture’s Consequences*, Beverly Hills: Sage Publications.

Hsu, J., and B. McFall (2015), “Mode Effects in Mixed-Mode Economic Surveys: Insights from a Randomized Experiment.” *Finance and Economics Discussion Series 2015-008*. Washington, DC: Board of Governors of the Federal Reserve System.

Krippendorff, K. (2004), “Reliability in Content Analysis: Some Common Misconceptions and Recommendations,” *Human Communication Research*, 30, 411–433.

Light, R., and D. B. Pillemer (1986), *Summing Up: The Science of Reviewing Research*, Cambridge, MA: Harvard University Press.

Lipsey, M. W., and D. B. Wilson (2001), “Analysis Issues and Strategies,” in *Practical Meta-Analysis*, ed. M. Lipsey and D. Wilson, pp. 105–128, Thousand Oaks, CA: SAGE Publications, Inc.

Lozar Manfreda, K., M. Bosnjak, J. Berzelak, I. Haas, and V. Vehovar (2008), “Web Surveys Versus Other Survey Modes: A Meta-Analysis Comparing Response Rates,” *Journal of the Market Research Society*, 50, 79–104.

Lyberg, L., and P. Dean (1992), “Methods for Reducing Nonresponse Rates: A Review,” paper presented at the AAPOR Annual Conference, St. Petersburg, FL.

Marreiros, H., M. Tonin, and M. Vlassopoulos (2016), “Now that You Mention It’: A Survey Experiment on Information, Salience and Online Privacy,” CESifo Working Paper Series No. 5756, Munich, available at https://www.econstor.eu/bitstream/10419/128459/1/cesifo1_wp5756.pdf.

Messer, B. L., M. L. Edwards, and D. A. Dillman (2012), “Determinants of Item Nonresponse to Web and Mail Respondents in Three Address-Based Mixed-Mode Surveys of the General Public,” *Survey Practice*, 5, 1–8.

Miller, T., M. Miller Kobayashi, E. Caldwell, S. Thurston, and B. Collett (2002), “Citizen Surveys on the Web General Population Surveys of Community Opinion,” *Social Science Computer Review*, 20, 124–136.

Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman, and The PRISMA Group (2009), “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement,” *Annals of Internal Medicine*, 151, 264–269.

Musch, J., and U.-D. Reips (2000), “A Brief History of Web Experimenting,” in *Psychological Experiments on the Internet*, ed. H. Birnbaum, pp. 61–87, San Diego: Elsevier.

Petrovčič, A., G. Petrič, and K. Lozar Manfreda (2016), “The Effect of Email Invitation Elements on Response Rate in a Web Survey Within an Online Community,” *Computers in Human Behavior*, 56, 320–329.

Peytchev, A. (2009), “Survey Breakoff,” *Public Opinion Quarterly*, 73, 74–97.

Polanin, J. R., and T. D. Pigott (2015), “The Use of Meta-Analytic Statistical Significance Testing,” *Research Synthesis Methods*, 6, 63–73.

Redline, C. D., and A. Zukerberg (2015), “Hope Springs Eternal: Will a Probability Sample of Schools and Principals Respond by Web and Provide Email Addresses?,” paper presented at the AAPOR Annual Conference, Anaheim, CA.

Rosenthal, R. (1979), “The File Drawer Problem and Tolerance for Null Results,” *Psychological Bulletin*, 86, 638.

Sakshaug, J. W., T. Yan, and R. Tourangeau (2010), “Nonresponse Error, Measurement Error, and Mode of Data Collection Tradeoffs in a Multi-Mode Survey of Sensitive and Non-Sensitive Items,” *Public Opinion Quarterly*, 74, 907–933.

Schmidt, F. L., and J. E. Hunter (2014), *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, Thousand Oaks: Sage publications.

Shih, T.-H., and X. Fan (2008), “Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis,” *Field Methods*, 20, 249–271.

Shojania, K. G., M. Sampson, M. T. Ansari, J. Ji, S. Doucette, and D. Moher (2007), “How Quickly Do Systematic Reviews Go Out of Date? A Survival Analysis,” *Annals of Internal Medicine*, 147, 224–233.

Smyth, J. D., D. A. Dillman, L. M. Christian, and A. C. O’Neill (2010), “Using the Internet to Survey Small Towns and Communities: Limitations and Possibilities in the Early 21st Century,” *American Behavioral Scientist*, 53, 1423–1448.

Szoc, R., J. Pflieger, F. Barlas, and R. Thomas (2013), “Great Expectations: Changing Mode of Survey Data Collection in Military Populations,” *Federal Committee on Statistical Methodology annual meeting, Washington, DC*. 2013, available at https://nces.ed.gov/FCSM/pdf/C4_Szoc_2013FCSM.pdf.

Tourangeau, R., J. Michael Brick, S. Lohr, and J. Li (2017), “Adaptive and Responsive Survey Designs: A Review and Assessment,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180, 203–223.

Tuten, T. L. (1997), “Getting a Foot in the Electronic Door: Understanding Why People Read or Delete Electronic Mail,” Technical, ZUMA.

World Bank, I. (2017), “Individuals Using the Internet (% of Population),” World Telecommunication/ICT Development Report and Database, 2018.

Wygant, S., and R. Lindorf (1999), “Surveying Collegiate Net Surfers—Web Methodology or Mythology,” *Quirk’s Marketing Research Review*. https://www.quirks.com/articles/surveying-collegiate-net-surfers-web-methodology-or-mythology