

¹ Nonresponse and Sample Weighting in Organizational Surveying

Abstract

Post-stratification weighting is a common procedure used in public opinion polling applications to correct demographic constituency differences between samples and populations. Although common practice in public opinion polling, this form of data remediation is only lightly acknowledged as a procedural topic worthy of empirical investigation within the organizational surveying literature. The current paper induces survey nonresponse via data simulation across fictional constituent groups (aka organizational strata) and documents the impact of weighting on the accuracy of sample statistics. Our goal was to evaluate the effectiveness of weighting when confronted with *passive* versus *active* forms of nonresponse in an effort to: 1) interject this data-remediation procedure into the established organizational surveying literature focused on nonresponse, while also 2) exploring organizationally-relevant sampling scenarios that: a) benefit from, b) are “hurt” by, or c) are unaffected by post-stratification weighting. The results confirm that sampling contexts characterized by active nonresponse do benefit from application of sample weights, but only when accompanied by constituency differences in underlying population construct (e.g., surveyed *attitude*) standing. Alternatively, constituent member differences in population attitudes, when characterized by passive forms of nonresponse, exhibit no benefit from weighting (and in fact these scenarios are somewhat *hurt* by weighting). The simulations reinforce that, moving forward, it would be prudent for surveyors of all disciplinary backgrounds to mutually attend to the traditional focus of both traditions: public opinion polling (e.g., multiple possible methodological sources of error as well as post-stratification adjustment) and organizational surveying (e.g., *form* of nonresponse).

Keywords: Survey methodology, sample weighting, nonresponse, response rate

26 Nonresponse and Sample Weighting in Organizational Surveying

27 Akin to differential variable weighting (for instance: a) construct indicators within a
28 multi-item assessment scale [aka factor loadings], or b) predictors within a selection system
29 [aka regression weights]; e.g., per data matrix “columns”), sample weighting alters the
30 proportional contributions of *individual respondents* within a data set (e.g., data matrix
31 rows). Some respondents’ responses are assigned greater relative contribution and others
32 are assigned less. This practice is commonplace in the summary of general population
33 polling data reflecting, for example, elections and politics (e.g., Rivers & Bailey, 2009),
34 prevalence rates of psychological disorders (e.g., Kessler et al., 2009), or feelings of physical
35 safety (e.g., Quine & Morrell, 2008). It is also seemingly in the periphery of awareness and
36 interest within the published organizational surveying literature (see, for example, Kulas et
37 al., 2016; Landers & Behrend, 2015; Tett et al., 2014).

38 We speculate that this form of statistical remediation is gaining research interest in
39 the organizational surveying research domain, at least in part, because industrial
40 psychologists are keenly aware that response rates within organizational surveying
41 applications have been trending downward (see, for example, Anseel et al., 2010; Rogelberg
42 & Stanton, 2007). With lower response rates, surveyors are confronted with heightened
43 levels of scrutiny because, historically, a locally realized high response rate has been
44 positively associated with data quality (e.g., Anseel et al., 2010; Cycyota & Harrison, 2002,
45 2006; Frohlich, 2002). The orientation of this presentation, however, is that although
46 response rate is a commonly referenced proxy of survey quality, it is not response rate but
47 rather sample representativeness that should be the primary focus of concern for survey
48 specialists (see, for example, Cook et al., 2000; Krosnick, 1999). Representativeness can of
49 course be “hurt” by low response rates, but the relationship between these two survey
50 concepts is by no means exact (e.g., Curtin et al., 2000; Keeter et al., 2006; Kulas et al.,
51 2017). Stated differently, a high response rate is neither a sufficient nor necessary condition

52 for representative population sampling.¹

53 In the context of survey applications, population misrepresentation refers to a
54 discrepancy between estimated sample statistics and actual population parameters. Ideally,
55 such discrepancies arise from completely random sources (in which case resulting error is
56 less likely to be reasonably characterized as bias). In reality, however, discrepancies are
57 driven not only by purely random causes. There are several broader sampling methodology
58 factors that may be systematically driving the relative under- or over-selection of a
59 population segment (see, for example, Kulas et al., 2016), but the most commonly cited
60 contributor within the organizational sciences is non-response (e.g., invited individuals
61 simply either forget or consciously choose not to participate in the survey process, see, for
62 example, Rogelberg et al., 2000). Our presentation also focuses on this non-response
63 contributor to sample misrepresentation, but only because we aim to: 1) integrate the
64 organizational non-response and public-opinion post-stratification weighting literatures,
65 while also 2) highlighting the associations and dissociations between response rate and
66 misrepresentation (although we also note here that the current presentation and procedure
67 address additional sampling methodological sources of misrepresentation).²

¹ Statistical benefits exist that are commonly attributed to higher response rates, such as greater power. These benefits, however, do not originate from response rate, but rather its consequence: larger n . Our presentation reflects the fact that greater power (and/or, relatedly, smaller confidence intervals) may in fact foster a false sense of confidence regarding “data quality”. Primarily for this reason, we stress that the methodological/statistical concepts of response rate, sample size, and power should be fully disentangled from the principle of representativeness, and the importance of these dissociations drives the central theme of the current paper.

² Frequently presented as a separate consideration, *measurement error* is an additional contributor to population misrepresentation. The current focus is on deviations from a perfect sampling methodology as opposed to deviations from an ideal psychometric methodology. We do however note that future advancements within the broad domain of “survey error” would benefit from a unified perspective that encompasses error arising from both methodological sources: measurement and sampling strategy.

68 Nonresponse in Organizational Surveying

69 Within the organizational surveying domain, it is not uncommon for response rate
70 (RR) to be referenced as a proxy for survey data quality (see, for example, Baruch &
71 Holtom, 2008; Fan & Yan, 2010; Pedersen & Nielsen, 2016). Baruch (1999), for example,
72 states that, “...to have dependable, valid, and reliable results, we need a high RR from a
73 wide representation of the whole population under study” and that, “The level of RR is an
74 important, sometimes crucial, issue in relation to the validity of a paper’s results” (p. 422).
75 Fan and Yan (2010) similarly state that response rate is, “...the most widely used and
76 commonly computed statistic to indicate the quality of surveys” (p. 132). Pedersen and
77 Nielsen (2016) claim that a high survey response rate, “...diminishes sampling bias
78 concerns and promotes the validity of survey-based research findings” (p. 230). The general
79 consensus seems to be that there are three major (negative) consequences of low response
80 rates, including (a) yielding smaller sample size, which negatively impacts statistical power
81 and confidence intervals, (b) reducing the credibility of survey data, and (c) generating
82 biased samples that impair the generalizability of survey results (Biemer & Lyberg, 2003;
83 Luong & Rogelberg, 1998; Rogelberg et al., 2000).

84 To the likely frustration of those who associate response rate with survey data
85 quality, organizational survey response rates have, on average, been declining for decades.
86 Baruch (1999), for example, summarized response rates of 175 studies published in five
87 leading management and behavioral sciences journals in 1975, 1985, and 1995. His results
88 revealed an average response rate (across time periods) of 55.6% ($SD = 19.7\%$), but also a
89 trend within which response rates declined steadily from 64.4% to 55.7% to 48.4% over the
90 three time points. Nine years later, Baruch and Holtom (2008) conducted a follow-up
91 study of 1,607 studies published in 17 disciplinary-relevant journals in 2000 and 2005 but
92 found no substantial differences in response rates compared to those in 1995, suggesting
93 that the declining trend had perhaps reached a lower asymptote. However, a different

94 approach with similar goals (Anseel et al., 2010) analyzed 2,037 survey projects published
95 in 12 journals in Industrial and Organizational Psychology, Management, and Marketing
96 from 1995 to 2008 and did note a slight decline (overall $M = 52.3\%$) when controlling for
97 the use of response enhancing techniques.³ The most recent like-minded review focused on
98 the years 2010, 2015, and 2020 and noted a reversal of the trend, such that average
99 response rates had risen to 68% in 2020 (Holtom et al., 2022).

100 ***Form of Nonresponse***

101 Although high response rates are generally pursued as desirable within
102 organizational surveying applications, there has also been a broad acknowledgement that
103 not all forms of nonresponse should be considered equally worrisome. Rogelberg et al.
104 (2003), for example, propose a distinction between active and passive nonrespondents
105 based on intent and (in)action. According to Rogelberg et al. (2003), active
106 nonrespondents are those who intentionally refuse to participate in surveys, while passive
107 nonrespondents are those who fail to respond to surveys due to reasons such as forgetting
108 or misplacing invitations. Passive nonrespondents are thought to be similar to respondents
109 in both attitude (Rogelberg et al., 2003) as well as organizational citizenship behaviors
110 (OCBs, Spitzmüller et al., 2007), whereas active nonrespondents have been shown to
111 exhibit significantly lower organizational commitment and satisfaction, higher intention to
112 quit, lower conscientiousness, and lower OCBs than actual respondents (Rogelberg et al.,
113 2000, 2003; Spitzmüller et al., 2007). Additionally, it has been noted that selection of an
114 individual population element into a realized sample may in fact be predictable (because
115 of, for example, an increased likelihood of not responding when dissatisfied or disgruntled,

³ It is also possible that the declination had stabilized with mean response rates hovering around 50% after roughly the turn of the millennium ($M = 52.5\%$ for HRM studies from 2009 to 2013, Mellahi & Harris, 2016; $M = 52.0\%$ for management studies from 2000 to 2004, Werner et al., 2007). This stability, if authentic, may again possibly be accounted for by an increased contemporary emphasis on response enhancing strategies (Anseel et al., 2010; Fulton, 2016).

¹¹⁶ Taris & Schreurs, 2007).

¹¹⁷ The more commonly encountered form of organizational nonresponse appears to be
¹¹⁸ passive (Rogelberg et al., 2003; Rogelberg & Stanton, 2007), although subgroup rates may
¹¹⁹ evidence variability - men, for example, have a higher proclivity toward active nonresponse
¹²⁰ than do women (Luong & Rogelberg, 1998; Rogelberg et al., 2000; Spitzmüller et al., 2007).
¹²¹ The organizational surveying baseline default expectation is that, *on average*, roughly 15%
¹²² of nonrespondents should be expected to be accurately characterized as “active”
¹²³ (Rogelberg et al., 2003; Rogelberg & Stanton, 2007; Werner et al., 2007). It is this second,
¹²⁴ less frequently anticipated form of nonresponse that also carries the greater resulting threat
¹²⁵ of biased sample estimates (see, for example, Kulas et al., 2017; Rogelberg & Stanton,
¹²⁶ 2007). It is these biased estimates that are the desired target of remediation when applying
¹²⁷ sample weights.

¹²⁸ **Sample Weighting - a Brief Overview**

¹²⁹ Within public opinion polling contexts, when realized sample constituencies (e.g.,
¹³⁰ 44% male - by tradition from *carefully-constructed* and *randomly sampled* data frames)⁴
¹³¹ are compared against census estimates of population parameters (e.g., 49% male), weights
¹³² are applied to the sample in an effort to remediate the relative proportional under- or
¹³³ over-sampling. This is because, if the broader populations from which the under- or
¹³⁴ over-represented groups are sampled differ along surveyed dimensions (e.g., males, within
¹³⁵ the population, are *less likely to vote for Candidate X* than are women), then unweighted
¹³⁶ aggregate statistics (of, for example, projected voting results) will misrepresent the true
¹³⁷ population parameter. This remedial application of sample weights should also be

⁴ These important sampling concepts are very carefully attended to within public opinion polling contexts. Conversely, within organizational surveying traditions, these considerations are not commonly acknowledged, at least explicitly within the published literature. The weighting procedure presented in the current manuscript remediates bias regardless of the methodological source, but is dependent on accurate “census” population constituency estimates.

138 considered an option for researchers pursuing answers to analogous organizational pollings
 139 such as: “What is the mood of the employees?” This is because focused queries such as
 140 this are of course covertly complex - implicit in the question is a focus not on survey
 141 results, but rather the broader employee population. Acknowledging the appropriate object
 142 of attribution is of course important, because the next step (after gauging the mood of the
 143 surveyed respondents) is *doing something* about it. Weighting may be a procedural option
 144 for organizational surveyors to credibly transition a bit closer from, “What do the survey
 145 results say”? to “What do the employees feel”?

146 **Procedural application**

147 *Proportional weights* are the form of weights most directly relevant to organizational
 148 surveying applications that traditionally focus on nonresponse as the primary contributor
 149 to sample misrepresentation. These weights are ratios of the proportion of a population
 150 within a given stratum to the proportion of the sample within that same stratum:

$$\text{proportional weight}(\pi_k) = \frac{(N_k/N)}{(n_k/n)} \quad (1)$$

151 Over-sampling of elements of a stratum (k) results in proportional weights less than
 152 one, while under-sampling (relative to the population) results in proportional weights
 153 greater than one. The common procedure for weight estimation *when more than one*
 154 *stratum is specified* is an iterative process that may be referred to by multiple substantively
 155 synonymous terms: *rim weighting*, *iterative proportional fitting*, or *raking* (see, for example,
 156 Deming & Stephan, 1940). Regardless of label, the procedure guides the surveyor to:

- 157 1) Determine proportional weights for all levels within one stratum, and then assign
 158 these weights to cases.
- 159 2) Determine proportional weights for a second group (ratio of population percent to
 160 *current* sample percent [the current sample percent will be affected by the step 1

161 weighting procedure]). Multiply previous (step 1) weights by the proportional
162 weights for this second stratum and assign these new weights to cases.

- 163 3) Determine proportional weights for a third stratum (which will once again require
164 re-inspection of the *current* sample percent). Multiply the previous step 2 weights by
165 the third stratum proportional weights and assign to cases.
- 166 4) Iterate through steps 1, 2, and 3 (or more if more than three strata are considered)
167 until the weighted sample characteristics match the population characteristics to your
168 desired level of precision.

169 Possible strata relevant for organizational survey weighting include: branch, full-,
170 part-, or flex-time status, functional area, gender, geographic location, hierarchy,
171 remote-working categorization, salaried status, subsidiary, tenure, work shift, or any other
172 groupings especially suspected to plausibly possess a relatively disporportionate number of
173 active nonrespondents (through application of forecasting strategies such as those
174 advocated by, for example, Rogelberg and Stanton, 2007). Each of these strata may of
175 course also be the targeted focus of survey results feedback, but when *aggregating* results
176 across (or even within) strata, a consideration of the impact of nonresponse *has the*
177 *potential* to yield more accurate survey estimates. The explicit goal is therefore a closer
178 approximation of descriptive sample statistics to population parameters via statistical
179 remediation, and drives the current paper's focus on the interplay of four survey elements:
180 1) distribution of attitude within the larger population, 2) response rate, 3) nonresponse
181 form, and 4) remedial weighting.

182 *Research question 1:* What role does overall response rate play in population
183 misrepresentation?

184 *Research question 2:* What role does nonresponse form (passive versus active) play
185 in population misrepresentation?

186 *Research question 3:* What impact does the application of weights have on both

187 biased⁵ (e.g., misrepresentative) and unbiased sample estimates?

188 We view these questions as being analogous to similar questions asked and answered

189 with differential variable weighting within the broader applied psychological disciplines.

190 Just as, for example, there has been debate regarding the merits of differential versus unit

191 variable weighting in a selection context or aggregate scale score definition (e.g., Bobko et

192 al., 2007; Wainer, 1976), we propose that a similar consideration is appropriate with

193 persons, and therefore compare and contrast unit- versus variable-sample element

194 weighting.

195 Methods

196 We address our research questions within a simulated fictionalized context of

197 organizational surveying (wherein it is common to assess estimates of employee attitude or

198 perception; for example, commitment, culture/climate, engagement, satisfaction). We

199 began the simulations by establishing “populations”, each consisting of 10,000 respondents

200 characterized by demographic categorizations across gender (male and female) and

201 department (A and B). We therefore had four demographic groups (male-A, male-B,

202 female-A, and female-B). For these population respondents, we generated scaled continuous

203 responses (real numbers) ranging from values of 1 to 5, representing averaged aggregate

204 scale scores from a fictional multi-item survey with a common $1 \rightarrow 5$ Likert-type rating

205 scale.

206 In order to represent different proportions of relative constituency (for example,

207 more females than males or more department A workers than department B), we iterated

208 population characteristics at marginal levels (gender and department) starting at 20% (and

209 80%) with increments and corresponding decrements of 20%. For example, if males

⁵ We have to be careful about the use of the term “bias” - either very carefully distinguish between error and bias or just avoid use of the term altogether. Perhaps Dr. Robinson can help here.

210 accounted for 20% of the simulated population, then females were 80%; also if respondents
211 in Department A represented 60% of a population, then 40% were in Department B.
212 Marginal constituencies were therefore specified at all combinations (across the two
213 variables) of 20% and 80%, 40% and 60%, 60% and 40%, and 80% and 20%. This resulted
214 in population *cell* constituencies (e.g., men in department A) as low as 400 and as high as
215 6,400.

216 Additionally, each of these cell populations was characterized by an attitudinal
217 distribution in one of three different possible forms: normal, positively skewed, or
218 negatively skewed. These distributional forms were specified in an attempt to model
219 similarities and discrepancies in construct standing (e.g., commitment, satisfaction, or
220 engagement) across respondent groupings. The normal distribution exhibited, on average,
221 a mean of 3.0 whereas the skewed distributions were characterized by average means of 2.0
222 and 4.0, respectively. In total, eight crossings of distributional type across employee
223 categorization were specified (Table 1 presents the combinations of these distributions).
224 Note that these eight conditions are not exhaustive of all possible combinations of
225 constituent groups and attitudinal distribution - we limited the simulations to
226 combinations that we projected to collectively be most efficiently informative.

227 Individual attitudes were randomly sampled from population distributions at the
228 cell level (e.g., Department A Males) without replacement. Response rates
229 (methodologically these could also be conceptualized as *sampling* rates) were controlled at
230 the marginal level using 10% increments ranging from 60% to 90%, and these were fully
231 iterated. Our cell-level response rates therefore ranged from 36% to 81% - a range of rates
232 chosen because they are reasonable expectations according to the organizational surveying
233 literature (e.g., Mellahi & Harris, 2016; Werner et al., 2007). We therefore investigated
234 error within the aggregate mean (e.g., grand mean or total sample mean) attributable to
235 different likelihoods of sample inclusion from constituent groups of different relative size

236 and representing populations of different attitudinal distribution, but at response rates
237 reasonably expected to exist in real-world organizational surveying contexts.

238 It should be noted here that there are several collective patterns of response that
239 are intended to represent sampling scenarios reflecting *passive* nonresponse across groups,
240 regardless of response rate. These are the scenarios in which all subgroups exhibit the same
241 response rate (e.g., 36%, 36%, 36%, and 36%). All other combinations of response rate are
242 intended operationalizations of active forms of nonresponse (e.g., not *as reasonably*
243 characterized as missing at random), although the degree to which a sampling scenario
244 should be reasonably considered to be reflecting active nonresponse is intended to increase
245 incrementally across response rate conditions.

246 In an attempt to capture this “degree of active nonresponse”, we calculated a simple
247 index of response rate discrepancy (SD; presented in Table 2). The “least” active
248 nonresponse scenarios are characterized by two subgroups with identical response rates and
249 two having a slightly different response rate (e.g., Dept A Males = 36%, Dept A Females =
250 36%, Dept B Males = 42%, and Dept B Females = 42%; see the second row of Table 2, the
251 SD index = .034)⁶. Also here note that three of our eight Table 1 conditions represent
252 scenarios where the presence of active nonrespondents is not expected to result in bias
253 (e.g., regardless of patterns of nonresponse, the unweighted sample mean is expected to
254 yield an unbiased estimate of the population mean). These are Table 1 conditions one
255 through three, where attitudinal distributions are of *the same form* across groups,

⁶ This method of simplifying the presentation of our response rate conditions is fully independent of consideration of population constituency and distributional form. That is, the amount of bias present in a sample estimate is expected to be quite different for Condition 7 with response rates of 48%, 48%, 72%, 72% versus 48%, 72%, 48%, 72%, even though the crude response rate index ($SD = 0.139$) is the same for both scenarios. There is additional information within these simulations (the effect of a *combination* of response rate and population form on degree of bias) that is therefore not captured via this simple SD index.

256 regardless of any individual group response rate discrepancy from others'.

257 These operationalizations of passive and active forms of nonresponse differ from
258 other investigations with similar goals. Kulas et al. (2017), for example, directly tie
259 probabilities of sample inclusion to an individual's held attitude (the likelihood of sample
260 inclusion is fully dependent on the population member's attitude). Conversely, the
261 probability of sample inclusion in the current investigation is dependent only on *group*
262 membership (with some of these groups occasionally being characterized by unique
263 attitude distributional forms). Essentially, Kulas et al. (2017) operationalize active
264 nonresponse at the person-level whereas the current paper does so at the group level. This
265 may be a more appropriate procedural specification with regard to the implications of
266 these simulations, as organizational surveyors are more likely to have an inclination of a
267 group's collective attitude or likelihood to respond (e.g., night shift workers, machine
268 operators) than they are of any one individual employee.

269 Results

270 In total, we generated 327.68 million samples (4,096 unique combinations of
271 response rate and population constituency across gender and department, simulated 10,000
272 times each across our eight Table 1 conditions). Each of these samples was comprised of,
273 on average, $n = 5,625$, collectively representing an experiment-wide simulated n of 1.8432
274 trillion. For each individual simulation, weights were applied iteratively to the data at the
275 two marginal (variable) levels via raking, and were estimated via the *anesrake* package
276 (Pasek, 2018) in R version 4.2.2 (2022-10-31 ucrt).

277 We were most interested in comparing the extent to which unweighted (aggregated
278 responses without raking) and weighted (aggregated weighted responses) sample means
279 approximated the known population means across our controlled specifications of response
280 rate, nonresponse form, and attitudinal distribution. Population means were extracted
281 from each iteration, as the simulations specified a new population at each iteration. The

282 “misrepresentation” between sample and population was operationalized by calculating: 1)
283 the discrepancies between the population and both weighted and unweighted sample
284 means, as well as, 2) the averaged deviations of these discrepancies from the population
285 mean (discrepancy in the “mean” of the means is bias, dispersion about the “mean” of the
286 means is error). If the average weighted sample mean was closer to the true population
287 mean, relative to the unweighted one, then the weighting was deemed beneficial.⁷

288 Yet to do:

- 289 1. means and standard deviations (report at least in overall response rate section)
290 DONE (ALSO ANOVA)
- 291 2. find the polynomial regression DONE (Figure 1 and 3)
- 292 3. correct the .51 estimate in the Cattell section KULAS TO DO 1/5
- 293 4. find the figure with “red” and “grey” standard error bars (standard error section)

294 THIS WAS PREVIOUS VERSION - KULAS AMEND TEXT

295 **Role of overall response rate**

296 Research question 1 asked what role overall response rate plays in population
297 misrepresentation. This is presented most directly in Figure 1, with *moderate* response
298 rates exhibiting the greatest degrees of misrepresentation across our simulated conditions.
299 Note here again that conditions 1 through 3, which represent populations with similar
300 distributions of attitude, do not exhibit misrepresentation regardless of response rate
301 ($\bar{d}_{Cond1} = 0.01$, $sd_{Cond1} = 0.00$; $\bar{d}_{Cond2} = 0.01$, $sd_{Cond2} = 0.00$; $\bar{d}_{Cond3} = 0.01$, $sd_{Cond3} =$
302 0.00). These can be contrasted most particularly with conditions 6 ($\bar{d}_{Cond6} = 0.05$, sd_{Cond6}
303 = 0.04), 7 ($\bar{d}_{Cond7} = 0.07$, $sd_{Cond7} = 0.05$), and 8 ($\bar{d}_{Cond8} = 0.07$, $sd_{Cond8} = 0.05$), which
304 evidence considerable misrepresentation, particularly so at moderate response rates (the

⁷ Do we want to do a little more with the dispersion concept? Currently it's underreported in the Results (but stated here that it is something we look at). If so, do we say that the weighting was beneficial also if the dispersion (error) was relatively small?

305 greatest degree of misrepresentation occurs with response rates ranging from roughly 40%
 306 to 70%)⁸. Discrepancies in unweighted means between samples and populations -
 307 regardless of response rate - did broach statistical significance across the 8 conditions
 308 ($F_{(7,32,760)} = 2,938.50, p < .001$). Tukey's HSD revealed differences across all contrasts
 309 other than between Conditions 1, 2, and 3 and also Conditions 7 and 8. Focusing only on
 310 Conditions 4 through 8, the relationship between response rate and sample/population
 311 discrepancy was significant beyond the effect of condition ($\Delta R^2 = 0.00; F = 7,862.44$), and
 312 the polynomial response rate term further added to the discrepancy prediction ($\Delta R^2 =$
 313 0.02; $F = 2,503.61$).⁹

314 **Role of nonresponse form**

315 Research question 2 asked what role the *form* of nonresponse (passive versus active)
 316 plays in population misrepresentation. In terms of explaining the error that did emerge
 317 within unweighted means sampled from conditions 4 though 8, this error was largely
 318 attributable to form of nonresponse as operationalized by our SD index (See Figure 2).
 319 Figure 2 also adds context to the Figure 1 response rate relationships, with the most
 320 extreme misrepresentation paralleling circumstances of active nonresponse (e.g., to the
 321 “right” in Figure 2). The nature of the exact relationship was slightly nonlinear, being fit
 322 with quadratic functions within each condition (collapsing across conditions did exhibit
 323 slight within-array differences [which would affect the statistically perfect relationship ←

⁸ Note that a confound exists whereby extreme overall rates (e.g., .36/.81) are necessarily associated with passive forms of non-response as operationalized in the current paper. The “middle”-most response rates are those most likely to be characterized by a mixture of both passive and active forms of non-response.

⁹ **NEEDS FURTHER THOUGHT/EXPLANATION** Middle range more cases - for the lowest case, there's only 256 cases (all with the same response rate of 36%). That explains the “upward slope” on the left of the graphing spaces. Clarification here would state that *these graphs hint to form of nonresponse being important - the lowest and highest response rates are constrained such that all groups have the same/similar levels of response rate - this is our operationalization of passive nonresponse.*

³²⁴ suggests the R^2 was 1.0].¹⁰

³²⁵ The systematic patterns of heteroskedasticity of the Figure 2 scatterplots should
³²⁶ also be noted. There are *active nonresponse* scenarios in which no error is present (see, for
³²⁷ example, the lower right-hand portions of conditions 4 through 8 in Figure 2 where
³²⁸ discrepancy estimates of “0” appear all along the passive-active x-axis). These
³²⁹ circumstances are simulated conditions within which the response rates “parallel” the
³³⁰ **POPULATION(?)** distributional form ← important to clarify as it would help explain
³³¹ confusing analyses. For example, in Condition Eight, the distributional forms were:
³³² $PositiveSkew_{Male(A)}$, $PositiveSkew_{Male(B)}$, $NegativeSkew_{Female(A)}$,
³³³ $NegativeSkew_{Female(B)}$. Response rates that “mirror” distributional patterns in extreme
³³⁴ cases of active nonresponse ($SD = .156$; e.g., 54%_{Male(A)}, 54%_{Male(B)}, 81%_{Female(A)},
³³⁵ 81%_{Female(B)}) result in effectively zero error in the population mean approximation (average
³³⁶ discrepancy = 0.00, $SD = 0.00$). Alternatively, when the response rates are inverted for the
³³⁷ $SD=.156$ cases, (e.g., 54%_{Male_A}, 81%_{Male_B}, 54%_{Female_A}, 81%_{Female_B}), there is substantial
³³⁸ error in approximation (average discrepancy = 0.16, $SD = 0.03$).¹¹ Again, it is not merely
³³⁹ response rate or form that is associated with biased sample estimates, but rather the
³⁴⁰ nature of response rate relative to existing attitudinal differences.

³⁴¹ To partially address this moderation, the discrepancies between population
³⁴² constituency and sampling proportions were additionally evaluated through the lens of
³⁴³ Cattell’s profile similarity index (r_p , Cattell, 1949; Cattell et al., 1966). r_p is sensitive to

¹⁰ Need to find these analyses if retain - Figure 2 looks linear with heteroskedasticity (6/16/23). Talking with Yang 6/30/23 we DID run these analyses. Look in older versions of the paper for a description of the analyses. **Could also just cut this or move down**

¹¹ Need to redo this - .51 doesn’t appear on graph, highest should be . `combo` object - DV = PPSU;
 grouping vars = `Condition & SD` (SD isolates most extreme “active nonresponse” cases). 1/10/24 ## Not
 seeing any .36,.36,81,.81, will need to clarify/troubleshoot with Yang. THIS SECTION NEEDS TO BE
 RETHOUGHT

discrepancies in profile shape (pattern across profile components), elevation (average component score), and scatter (sum of individual components' deviation from the elevation estimate. Figure 3 demonstrates the pattern of unweighted sample mean deviation (from the population parameter) when this index is taken into consideration. Specifically, Figure 3 demonstrates a more pronounced *form of* nonresponse association when underlying attitudinal distributions evidence group differences, and in these scenarios, active nonresponse is shown to have a fairly large effect on error within the sample estimate (as well as systematically increasing degrees of heteroskedasticity paralleling the Cattell index; omnibus Breusch-Pagan = 3177.2, $p < .001$). The curvilinear nature of these (non conditions 1, 2, and 3) functions was again estimated via polynomial regression, with misrepresentation again exhibiting a linear association across the Cattell index ($\Delta R^2 = 0.24$, $p < .001$), but also a polynomial effect ($\Delta R^2 = 0.07$, $p < .001$).

Impact of weighting

Research question 3 was focused on the impact of weights on both biased (e.g., misrepresentative) and unbiased sample estimates¹². Figure 4 provides a broad summary of the results across the eight different attitudinal distribution conditions, presenting the average absolute discrepancy from the population mean for the weighted and unweighted sample estimates. Conditions one through three demonstrate that, on average, the unweighted sample mean provides a good (unbiased) estimate of the population mean when the distributional form does not differ across constituent groups (e.g., the distributions of attitudes are of similar functional forms and locations for all constituent groups). This is regardless of form or extent of nonresponse. Additionally, weighting remediates deviations about the true mean in all five attitudinally discrepant conditions, even when substantive relative error exists in the unweighted estimate (e.g., the rightmost bars in Figure 4). Although the *patterns* of unweighted sample mean discrepancies differed

¹² Come back to this phrasing after decision is made on RQ 3 wording (whether to avoid using the term bias or not).

369 across conditions, all eight conditions exhibited similar omnibus effect (weighting
370 ameliorating error wherever it arose [in the unweighted statistic]).

371 To further elaborate this point, consider, for example, Condition 4 as presented in
372 Table 1. Here, three groups are characterized by similar distributions of attitudes
373 (normally distributed) and one, Females from Department B, is characterized by negatively
374 skewed attitudes. The greatest unweighted error here arises from sampling scenarios in
375 which there are many Department B females (e.g., in our specifications, 6,400) and fewer
376 males and Department A females¹³, but the Department B females exhibit a much lower
377 response rate (e.g., 20%) than do other groups, who respond at a high rate (e.g., 80%).
378 That is, it is not merely response rate, but response rate within these identifiable groups,
379 and whether or not those response rate differences parallel underlying attitudinal
380 differences that drives sample misrepresentation.

381 ***Weighting and Sampling Error***

382 Mean square error is our second index for sample quality. It is a well-known
383 mathematical theorem that the application of weights increases (random) errors of
384 precision, which was also empirically true in the current study. For each condition in our
385 simulations, we calculated the standard deviations of 40.96 million unweighted and 40.96
386 million weighted samples means (4,096 possible population-sample combinations by 10,000
387 iterations), which yielded eight empirically-estimated standard errors of unweighted and
388 weighted sample means. Figure 4 visually presents these standard errors in eight pairs of
389 bars, demonstrating that the standard error of weighted sample means (red bar) tended to
390 be 16% to 18% larger than that of unweighted sample means (grey bar) regardless of
391 condition. These errors highlight the caveat that weighting should only be applied in the

¹³ Because of the “marginal” versus “cell” specifications of population constituencies, our most extreme example here necessarily results in 400 Department A males, 1,600 Department B males, and 1,600 Department A females. This was a decision based on keeping the population N’s at 10,000 and certainly more extreme population constituency combinations could be examined in future like-minded explorations.

392 active nonresponse case (e.g., although the aggregate effect of weighting with passive
393 nonresponse is error-minimizing, any one sampling condition is *more likely* to result in
394 greater deviation from the population parameter when weighting is applied to sample data
395 driven by passive nonresponse).

396 **Collective roles of response rate, form, and attitudinal distribution**

397 As an aggregate across sampling events, weighting always corrects sample bias,
398 when it is present in the unweighted estimate. However, the standard errors suggest that
399 for any *one* sampling event in the absence of bias, the likelihood that the sample mean
400 approximates the *mean* of sample means is (slightly) greater for the unweighted estimate.
401 When bias is present, however, (in the unweighted estimate) there is obviously no
402 advantage to “being closer” to this biased mean of means. That is, under some
403 circumstances, the mean of unweighted sample means does not center on the population
404 mean. The implications of this seem quite obvious: Weighting should only be applied if
405 bias is anticipated in the sample estimate. This may seem to be a picayune
406 recommendation, but we note here that this advocacy is not heeded in public opinion
407 polling applications, where the computation and application of weights are default
408 procedures (CITES? - perhaps AAPOR standards or personal communication with polling
409 agencies such as Gallop).

410 **Question for David - Can we look at the “crossing point?” (e.g., when
411 MSE becomes excessive)**

412 [perhaps David can derive/find a proof to parallel our results?] (Table 1
413 + ResponseRate1 + SDForm2 + Figure 4) Maybe try to combine Figures 2
414 and 3 (put SD on Figure 3 - color code)

415 Added population attitudes (1/20/23) - not sure if this clutters but more
416 consistent with flow of introduction

417

Discussion

418

We view nonresponse as a serious problem that should be addressed via repeated attempts to survey particularly reluctant or hard-to-reach respondents because nonresponse may be reasonably expected to be greatest in groups that are most unsatisfied [e.g., it may be typical for individuals representing these groups to have their responses diluted; see, for example, Taris and Schreurs (2007)]. However, several researchers have noted potentially misplaced relative emphasis on survey response rates, with Cook et al. (2000), Krosnick (1999), and Visser et al. (1996) articulating the point that representativeness of the sample is more important than response rate. We also believe that the goal in organizational surveying should be representativeness not exhaustiveness. Krosnick (1999) specifically comments that, even when probability sampling is employed, response rate does not necessarily implicate either good or poor sample representativeness. One aim of this paper is to reinforce this primary ‘representativeness’ orientation to those who may be otherwise inclined to focus on response rate as a sufficient index of quality (while also stressing sample weighting as a practice that can potentially remediate *misrepresentativeness*).

432

With the above in mind, we set out to answer three fairly straightforward questions:

433

What roles do 1) response rate and 2) form of nonresponse have on population misrepresentation, and 3) what impact does the application of weights have on the quality of sample estimates? The simulations demonstrate that the impact of mere response rate impact *depends* on the underlying distributions of population attitude. Conditions 1 through 3 (as well as all other conditions) are occasionally immune to response rate influence, depending on whether the pattern of nonresponse parallels the pattern of attitudinal distribution differences or not). Active forms of nonresponse can harm the unweighted sample estimate, but only when the pattern of active nonresponse is accompanied by differing distributions of attitudes within the active nonrespondent “populations” [this would appear to be a reasonable expectation based on the literature;

443 e.g., Rogelberg et al. (2000); Rogelberg et al. (2003); Spitzmüller et al. (2007)]. Weighting
444 “always” helps, as long as you capture the proper strata (which of course we were able to
445 do via controlled simulation), but also... Although the weighted mean proved an unbiased
446 estimate of the population mean across all simulations, in circumstances where no bias
447 existed in the unweighted estimate, the trade-off between bias-correction and random error
448 of precision (e.g., standard error) also needs to be acknowledged.

449 It may be noted here that the organizational surveying categorization of passive
450 versus active somewhat parallels the broader statistical focus on data that is missing at
451 random or completely at random (MAR or MCAR, see for example, Heitjan & Basu, 1996)
452 versus data not missing at random (MNAR, see for example, Enders, 2011). Imputation is
453 a common remediation technique for data MAR or MCAR whereas MNAR solutions may
454 involve strategies such as latent variable estimation procedures (Muthén et al., 1987). In
455 the context of organizational surveying, the current findings lead to a similar bifurcation of
456 remediation methods - post-stratification weighting is recommended only in the
457 circumstance of active nonresponse.

458 Previous presentations have noted that bias is sometimes associated with
459 nonresponse and other times it is not - this research has not been explicit in the specific
460 conditions that moderate this association, however. The current paper does make this
461 association explicit. It is not merely the form of nonresponse that determines whether or
462 not bias occurs, but also the underlying distributions that the response probabilities are
463 applied to. Some distributional patterns are immune to the biasing effects of active
464 nonresponse (see, for example, Conditions 1 through 3). Some patterns of active
465 nonresponse also result in no bias even when distributional patterns deviate substantially
466 (see, for example, Condition 8 where a 20%, 20%, 80%, 80% response rate pattern exhibits
467 no error). The target therefore should not be merely form of nonresponse but also
468 underlying attitudes. Regardless, however, weighting always remediates the error when it

469 occurs (and does not add error where it is absent).

470 The current findings are of course qualified by the uniqueness of our simulations,
471 most notably our ability to fully capture the correct population parameters (e.g., because
472 these were “created” by us, we were also able to identify these strata as the nonresponse
473 contributors). Even in the extreme conditions (e.g., a small “population” with a
474 correspondingly low response rate; see, for example, the lower-left hand corner of Figure 2),
475 the weighting algorithm was able to provide a bias correction. This is undoubtedly
476 attributable to our random sampling procedure (instead of, for example, sampling
477 conditionally from the population distributions), but here we do note that the raking
478 procedure is applied at the “margins” (e.g., variable level, not interaction level), although
479 our introduction of a biasing element is at the cell (interaction) level.

480 It has been stated that active nonresponse is relatively harmless unless the actively
481 nonrespondent group is relatively large [cites below]. The current study, however, suggests
482 that post-data-collection remediation. There may also be some important implications here
483 regarding sample (and population) size. Because organizational surveyors likely interface
484 with organizations of varying sizes (perhaps some of which are small- or medium-sized), the
485 implications of our simulations particularly in the small population conditions, were
486 highlighted. Findings specific to these conditions were: XXX, XXX, XXX.

487 There is of course no need to restrict weighting protocols to demographic groups -
488 organizational surveyors have a rich tradition of attending to drivers of nonresponse (see,
489 for example, Rogelberg & Stanton, 2007). Groupings of any sort can be the basis of
490 weighting (for example, pre-survey probing might assign probabilities of nonresponse, and
491 these probabilities can be retained post-administration as weighting guides).

492 It should also be pointed out that although the active nonrespondent group seems
493 to be a great concern, it will not seriously bias the results unless the proportion of active

494 nonrespondents is higher than 15% (Rogelberg et al., 2003; Rogelberg & Stanton, 2007;
495 Werner et al., 2007). “In this study we found that the active nonrespondent group was
496 relatively small (approximately 15%), but consistent in size with research conducted by .”
497 (Rogelberg et al., 2003, pp. 1110–1111). “Furthermore, consistent with Roth (1994) who
498 stated that when missingness is not random (as we found for active nonrespondents),
499 meaningful bias will only be introduced if the group is relatively large (which was not the
500 case in this study).” (Rogelberg et al., 2003, p. 1112).

501 “If the results show that the active nonrespondent group comprises a low proportion
502 of the population, fewer concerns for bias arise. If the proportion of active respondents is
503 greater than 15% of the group of individuals included in the interviews or focus groups
504 (this has been the average rate in other studies), generalizability may be compromised.”
505 (Rogelberg & Stanton, 2007, p. 201) * I believe there is an error here. The author want to
506 say that if the proportion of active nonrespondents is greater than 15% of the group .

507 “It has been suggested that it takes a response rate of 85% to conclude that
508 nonresponse error is not a threat (Dooeyl & Lindner, 2003). We agree that researchers
509 should provide both empirical and theoretical evidence refuting nonresponse bias whenever
510 the response rate is less than 85%.” (Werner et al., 2007, p. 293).

511 Note here however, the seeming disconnect between the reports of 15% active
512 nonresponse and declining response rates (trending toward 50%). Certainly with
513 decreasing overall response rates, the likely reasons would appear to be more active than
514 passive (e.g., it is difficult to entertain the idea that potential respondents are more likely
515 to forget to respond today than they were 40 years ago).

516 Integration of IT/IS systems within HR functions hopefully assists the
517 (un)likelihood that organizatinoal population frames are either deficient or
518 contaminated, although we note that this possibility (frame misspecification) is

519 much more plausible within organizations that do not have updated or
520 integrated HR IT/IS systems (perhaps, ironically, *smaller* organizations).

521 **Limitations**

522 The results are presented with at least three limitations: 1) our simulations are
523 comprehensive, iterating through all possible combinations of response rates - those
524 paralleling population distributions, those inversely mirroring population distributions, and
525 those "orthogonal to" population distributions, 2) the "SD" operationalization of passive to
526 active forms of nonresponse is a bit crude and insensitive to specific combinations of
527 response rates expected to manifest or not manifest in bias, and 3) substantial bias may be
528 present in the unweighted estimate even with only small proportions of active non-response
529 (e.g., only one or two groups exhibiting slightly different response rates, with the resulting
530 discrepancy [population versus sample mean] being quite large).

531 **Future Directions**

532 Our operationalization of passive nonresponse was based on realized subsample
533 differences in response rate. Of course it is plausible that consistent response rates (e.g.,
534 36%, 36%, 36%, 36%) could have corresponding *non-sampled* elements who represent active
535 non-response. Our methodology did not model these scenarios, but future like-minded
536 investigations may wish to do so.

537 A very practical implication of this study is that future organizational researchers
538 may find more success implementing strategic sampling strategies as opposed to (or in
539 addition to) pursuing response enhancement. That is, as a field, organizational researchers
540 have been focused on response-enhancing strategies that minimize the presence of
541 nonresponse. The current findings suggest that more careful adherence to random sampling
542 from carefully constructed population frames may provide a different route to the same
543 end-goal of sample representativeness.

544 Experimental methods within the psychological discipline have long been criticized

545 for heavy reliance on samples of convenience (for instance, student samples). Very little
546 progress has been made regarding the application of appropriate population sampling
547 procedures in experimentation. Certain non-experimental procedures (most notably
548 organizational surveying) hold paradoxical advantage over experimental procedures
549 primarily in this arena of sampling - particularly in consideration of population coverage,
550 which refers to the percent of a population that is reachable by the sampling procedure
551 (e.g., postal, intra-office, or internet invitation) and likelihood of having access to
552 population parameter estimates (e.g., strata constituencies). There is a rich tradition and
553 literature of public opinion polling procedures and techniques from which to draw. These
554 procedures, however, only hold advantage if the non-experimental methodologist
555 acknowledges the criticality of sample representativeness. The current paper provides one
556 corrective technique (post-stratification weighting) as an important focus for the
557 organizational surveyor who shares this primary interest in maximizing sample
558 representativeness.

559 We note the above “advantage” held by organizational surveyors because extensions
560 of the current protocol include investigating how inaccurate census estimates (and/or
561 grabbing the “wrong” group) affects the quality of sample estimates. That is, in our
562 controlled simulations, we were able to know population constituencies, because they were
563 set by us! In real-world applications, there is likely more error between the population
564 estimate and actual population constituency. Similarly, if the association between attitude
565 and group membership were to be controlled, there may be conditions identified whereby
566 weighting loses its efficacy (e.g., low “correlations” between attitude and group
567 membership). Future simulations should test boundary conditions for this type of error,
568 identifying at what point inaccuracy in the population constituency estimate appreciably
569 degrades the weighting procedure. Furthermore, it was demonstrated here that, when bias
570 exists, weighting corrects it. Weighting also, however, results in a larger mean square error
571 (MSE; expected spread of sample estimates around the population parameter). Feasibly

572 then, there is a point at which the decreased bias is accompanied by an unacceptably
573 inflated MSE. At which point does this occur? This is another fertile area for future
574 exploration.

575 Most potential issues with weighting are addressed through careful consideration of
576 the appropriate strata to take under consideration as well as ultimate level of aggregation
577 (what group constitutes the population of interest or focus of feedback; e.g., regional,
578 functional, or organizational?). We recommend the surveyor especially considers groups
579 that might have issues of active forms of nonresponse and collect those demographics so
580 weighting is an option. It is particularly in these contexts of ‘unsatisfied’ employees being
581 less likely to respond to surveys that pre-stratification consideration becomes critical (for
582 instance, if there is an inclination that attitudes may differ across, for example, night
583 versus day shift workers, it is important that shift be measured and incorporated as a
584 stratum prior to survey administration).

585 For Condition 5 (for example, low/high response rates with minority/majority
586 population constituencies). The lower-right to upper-left diagonal reflects response rates
587 that parallel population constituencies. The patterns across these stressors were consistent,
588 with the weighted sample means (red dots) providing unbiased estimates of the population
589 parameter, whereas the unweighted sample means (grey dots) tended to yield unbiased
590 estimates when the population constituencies were roughly equal (e.g., close to 50%/50%).

591 Figure 3 drills down this information further by extracting unweighted and weighted
592 estimates in one specific marginal population parameter combination (here, 60% males and
593 40% females; 40% in department A and 60% in department B). In doing so, the population
594 parameters were in control and sample parameters were set free (see dotted red rectangle
595 in Figure 2). Therefore, Figure 3 was then arranged in a fashion that allowed further
596 investigation into the interactive effect of marginal sample parameters (gender on the
597 x-axis and department on the y-axis) on the effectiveness of post-stratification weighting

598 reflected by the pattern of grey and red dots. **Huh? - find old version or delete**

599 Could be introducing more error if try to apply weights to correct constintuent
600 proportionalities with passive nonresponse.

601 PRIOR TO RQs: after chatting with Yang (10/31/19) these need to be
602 clarified a bit - reading 11/3 they make sense but need to be read very carefully.
603 Check with Yang on 1/26 Skype. 2/1 revisions seem ok to Kulas. Three moving
604 parts: underlying attitudinal distributions, response rate, and form of
605 nonresponse <- perhaps we should make these variables more explicit prior to
606 the procedure/results... .

References

- Anseel, F., Lievens, F., Schollaert, E., & Choragwicka, B. (2010). Response rates in organizational science, 1995–2008: A meta-analytic review and guidelines for survey researchers. *Journal of Business and Psychology*, 25(3), 335–349.
- Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. <https://github.com/crsh/papaja>
- Baruch, Y. (1999). Response rate in academic studies—a comparative analysis. *Human Relations*, 52(4), 421–438.
- Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. *Human Relations*, 61(8), 1139–1160.
- Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality* (Vol. 335). John Wiley & Sons.
- Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: A literature review, application to content validity, and meta-analysis. *Organizational Research Methods*, 10(4), 689–709.
- Cattell, R. B. (1949). R p and other coefficients of pattern similarity. *Psychometrika*, 14(4), 279–298.
- Cattell, R. B., Coulter, M. A., & Tsujioka, B. (1966). The taxonometric recognition of types and functional emergents. *Handbook of Multivariate Experimental Psychology*, 288–329.
- Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web-or internet-based surveys. *Educational and Psychological Measurement*, 60(6), 821–836.
- Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly*, 64(4), 413–428.
- Cycyota, C. S., & Harrison, D. A. (2002). Enhancing survey response rates at the executive level: Are employee-or consumer-level techniques effective? *Journal of*

- 634 Management, 28(2), 151–176.
- 635 Cycyota, C. S., & Harrison, D. A. (2006). What (not) to expect when surveying
636 executives: A meta-analysis of top manager response rates and techniques over
637 time. *Organizational Research Methods*, 9(2), 133–160.
- 638 Deming, W. E., & Stephan, F. F. (1940). On a least squares adjustment of a
639 sampled frequency table when the expected marginal totals are known. *The
640 Annals of Mathematical Statistics*, 11(4), 427–444.
- 641 Enders, C. K. (2011). Missing not at random models for latent growth curve
642 analyses. *Psychological Methods*, 16(1), 1–16.
- 643 Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A
644 systematic review. *Computers in Human Behavior*.
- 645 Frohlich, M. T. (2002). Techniques for improving response rates in OM survey
646 research. *Journal of Operations Management*, 20(1), 53–62.
- 647 Fulton, B. R. (2016). Organizations and survey research: Implementing response
648 enhancing strategies and conducting nonresponse analyses. *Sociological Methods
649 & Research*, 0049124115626169.
- 650 Heitjan, D. F., & Basu, S. (1996). Distinguishing “missing at random” and “missing
651 completely at random.” *The American Statistician*, 50(3), 207–213.
- 652 Holtom, B., Baruch, Y., Aguinis, H., & A Ballinger, G. (2022). Survey response
653 rates: Trends and a validity assessment framework. *Human Relations*, 75(8),
654 1560–1584.
- 655 Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the
656 impact of growing nonresponse on estimates from a national RDD telephone
657 survey. *International Journal of Public Opinion Quarterly*, 70(5), 759–779.
- 658 Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa,
659 S., Merikangas, K. R., Pennell, B.-E., Sampson, N. A., & Zaslavsky, A. M.
660 (2009). National comorbidity survey replication adolescent supplement (NCS-a):

- 661 II. Overview and design. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48(4), 380–385.
- 662
- 663 Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50(1), 537–567.
- 664
- 665 Kulas, J. T., Robinson, D. H., Kellar, D. Z., & Smith, J. A. (2017). Nonresponse in organizational surveying: Attitudinal distribution form and conditional response probabilities' impact on patterns of bias. *Public Opinion Quarterly*, 81(2), 401–421.
- 666
- 667
- 668
- 669 Kulas, J. T., Robinson, D. H., Smith, J. A., & Kellar, D. Z. (2016).
- 670 Post-stratification weighting in organizational surveys: A cross-disciplinary
- 671 tutorial. *Human Resource Management*.
- 672
- 673 Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary
- 674 distinctions between organizational, mechanical turk, and other convenience
- 675 samples. *Industrial and Organizational Psychology*, 8(2), 142–164.
- 676
- 677 Luong, A., & Rogelberg, S. G. (1998). How to increase your survey response rate.
- 678
- 679 The *Industrial-Organizational Psychologist*, 36(1), 61–65.
- 680
- 681 Mellahi, K., & Harris, L. C. (2016). Response rates in business and management
- 682 research: An overview of current practice and suggestions for future direction.
- 683
- 684 British *Journal of Management*, 27(2), 426–437.
- 685
- 686 Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with
- 687 data that are not missing completely at random. *Psychometrika*, 52(3), 431–462.
- 688
- 689 Pasek, J. (2018). *Anesrake: ANES raking implementation*.
- 690 <https://CRAN.R-project.org/package=anesrake>
- 691
- 692 Pedersen, M. J., & Nielsen, C. V. ek. (2016). Improving survey response rates in
- 693 online panels: Effects of low-cost incentives and cost-free text appeal
- 694 interventions. *Social Science Computer Review*, 34(2), 229–243.
- 695
- 696 Quine, S., & Morrell, S. (2008). Feeling safe in one's neighbourhood: Variation by
- 697

- location among older australians. *The Australian Journal of Rural Health*, 16, 115–116.
- Rivers, D., & Bailey, D. (2009). Inference from matched samples in the 2008 US national elections. *Proceedings of the Joint Statistical Meetings*, 1, 627–639.
- Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S., & Knight, W. E. (2003). Profiling active and passive nonrespondents to an organizational survey. *Journal of Applied Psychology*, 88(6), 1104.
- Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee attitude surveys: Examining the attitudes of noncompliant employees. *Journal of Applied Psychology*, 85(2), 284.
- Rogelberg, S. G., & Stanton, J. M. (2007). *Introduction: Understanding and dealing with organizational survey nonresponse*. Sage Publications Sage CA: Los Angeles, CA.
- Spitzmüller, C., Glenn, D. M., Sutton, M. M., Barr, C. D., & Rogelberg, S. G. (2007). Survey nonrespondents as bad soldiers: Examining the relationship between organizational citizenship and survey response behavior. *International Journal of Selection and Assessment*, 15(4), 449–459.
- Taris, T. W., & Schreurs, P. J. (2007). How may nonresponse affect findings in organizational surveys? The tendency-to-the-positive effect. *International Journal of Stress Management*, 14(3), 249.
- Tett, R., Brown, C., & Walser, B. (2014). The 2011 SIOP graduate program benchmarking survey part 7: Theses, dissertations, and performance expectations. *The Industrial-Organizational Psychologist*, 51(4), 62–73.
- Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for election forecasting? An evaluation of the columbus dispatch poll. *Public Opinion Quarterly*, 60(2), 181–227.
- Wainer, H. (1976). Estimating coefficients in linear models: It don't make no

- 715 nevermind. *Psychological Bulletin*, 83(2), 213.
- 716 Werner, S., Praxedes, M., & Kim, H.-G. (2007). The reporting of nonresponse
717 analyses in survey research. *Organizational Research Methods*, 10(2), 287–295.

Table 1*Attitudinal Distribution Conditions Specified in Current Paper*

Condition	Distributional Shape	mu	Male		Female		Bias Susceptibility
			Dept A	Dept B	Dept A	Dept B	
1	Normal	3	X	X	X	X	Low
	Positive Skew	2					
	Negative Skew	4					
2	Normal	3					Low
	Positive Skew	2	X	X	X	X	
	Negative Skew	4					
3	Normal	3					Low
	Positive Skew	2					
	Negative Skew	4	X	X	X	X	
4	Normal	3	X	X	X		Moderate
	Positive Skew	2					
	Negative Skew	4				X	
5	Normal	3	X	X			Moderate/High
	Positive Skew	2			X	X	
	Negative Skew	4					
6	Normal	3		X	X		Moderate/High
	Positive Skew	2	X				
	Negative Skew	4				X	
7	Normal	3					High
	Positive Skew	2	X		X		
	Negative Skew	4		X		X	
8	Normal	3					High
	Positive Skew	2	X	X			
	Negative Skew	4			X	X	

Table 2

Example Summarized Response Rate Conditions Represented in Figures 2 through 5

Example Response Rates (Any Combination)							Form (and degree) of Nonresponse
Male Dept A	Male Dept B	Female Dept A	Female Dept B	SD Index	Number of Conditions		
36%	36%	36%	36%	.000	256		Passive
36%	36%	42%	42%	.034	128		
48%	48%	54%	54%	.035	64		
42%	42%	49%	49%	.040	192		
48%	48%	56%	56%	.046	128		
56%	56%	64%	64%	.047	64		
54%	54%	63%	63%	.051	128		
63%	63%	72%	72%	.052	64		
36%	42%	42%	49%	.053	64		
42%	48%	49%	56%	.057	128		
49%	56%	56%	64%	.061	64		
48%	54%	56%	63%	.062	128		
56%	63%	64%	72%	.066	128		
36%	36%	48%	48%	.069	128		
64%	72%	72%	81%	.069	64		
42%	42%	56%	56%	.081	128		

Table 2 continued

NONRESPONSE AND SAMPLE WEIGHTING

35

Example Response Rates (Any Combination)

Male Dept A	Male Dept B	Female Dept A	Female Dept B	SD Index	Number of Conditions	Form (and degree) of Nonresponse
36%	42%	48%	56%	.085	128	
42%	49%	54%	63%	.089	128	
48%	48%	64%	64%	.092	128	
42%	48%	56%	64%	.096	128	
49%	56%	63%	72%	.098	128	
36%	36%	54%	54%	.104	192	
48%	54%	64%	72%	.106	128	
56%	63%	72%	81%	.109	128	
36%	48%	48%	64%	.115	64	
36%	42%	54%	63%	.120	128	
42%	42%	63%	63%	.121	64	
42%	54%	56%	72%	.123	128	
49%	63%	63%	81%	.131	64	
42%	48%	63%	72%	.137	128	
48%	48%	72%	72%	.139	64	
36%	48%	54%	72%	.150	128	

Table 2 continued

Example Response Rates (Any Combination)						
Male Dept A	Male Dept B	Female Dept A	Female Dept B	SD Index	Number of Conditions	Form (and degree) of Nonresponse
48%	54%	72%	81%	.154	128	
54%	54%	81%	81%	.156	64	
42%	54%	63%	81%	.164	128	
36%	54%	54%	81%	.186	64	Active

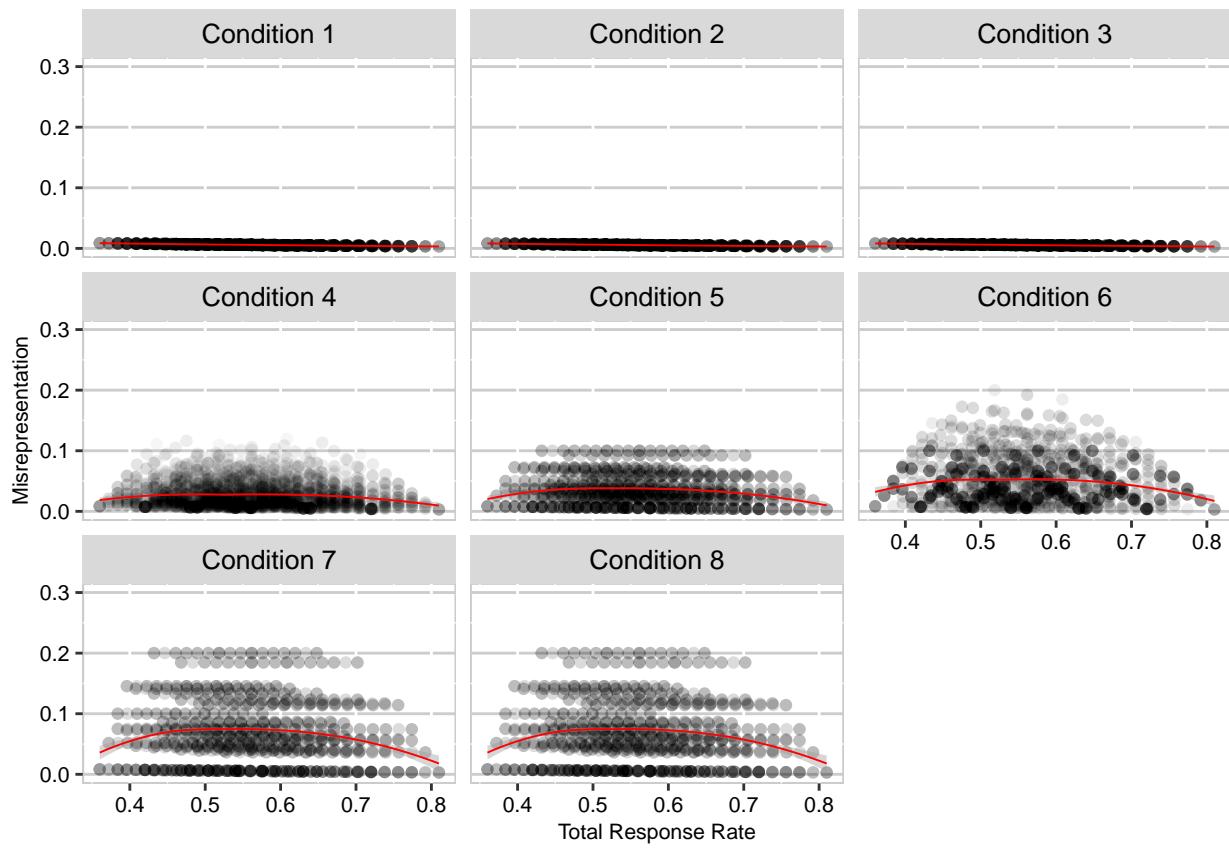


Figure 1

Relationship between total response rate and misrepresentation.

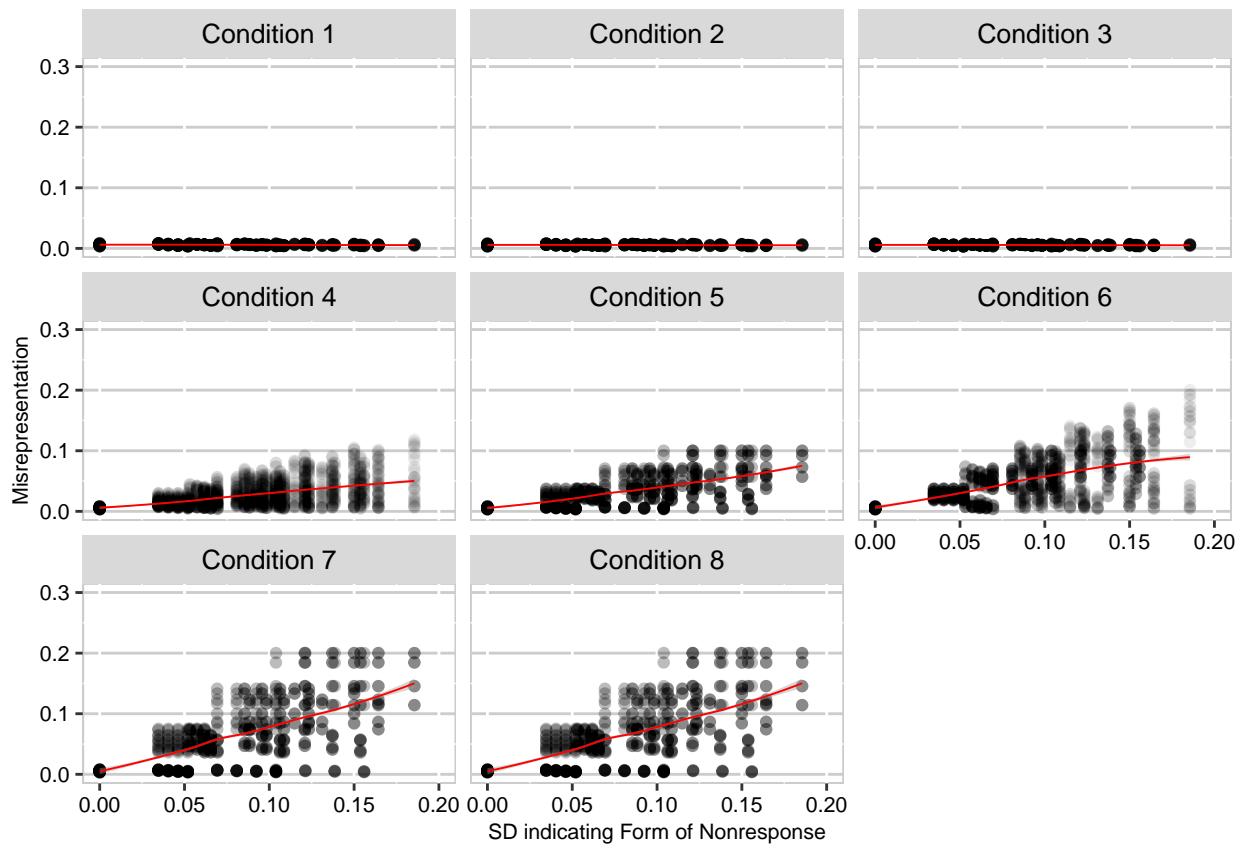


Figure 2

Relationship between nonresponse form and misrepresentation.

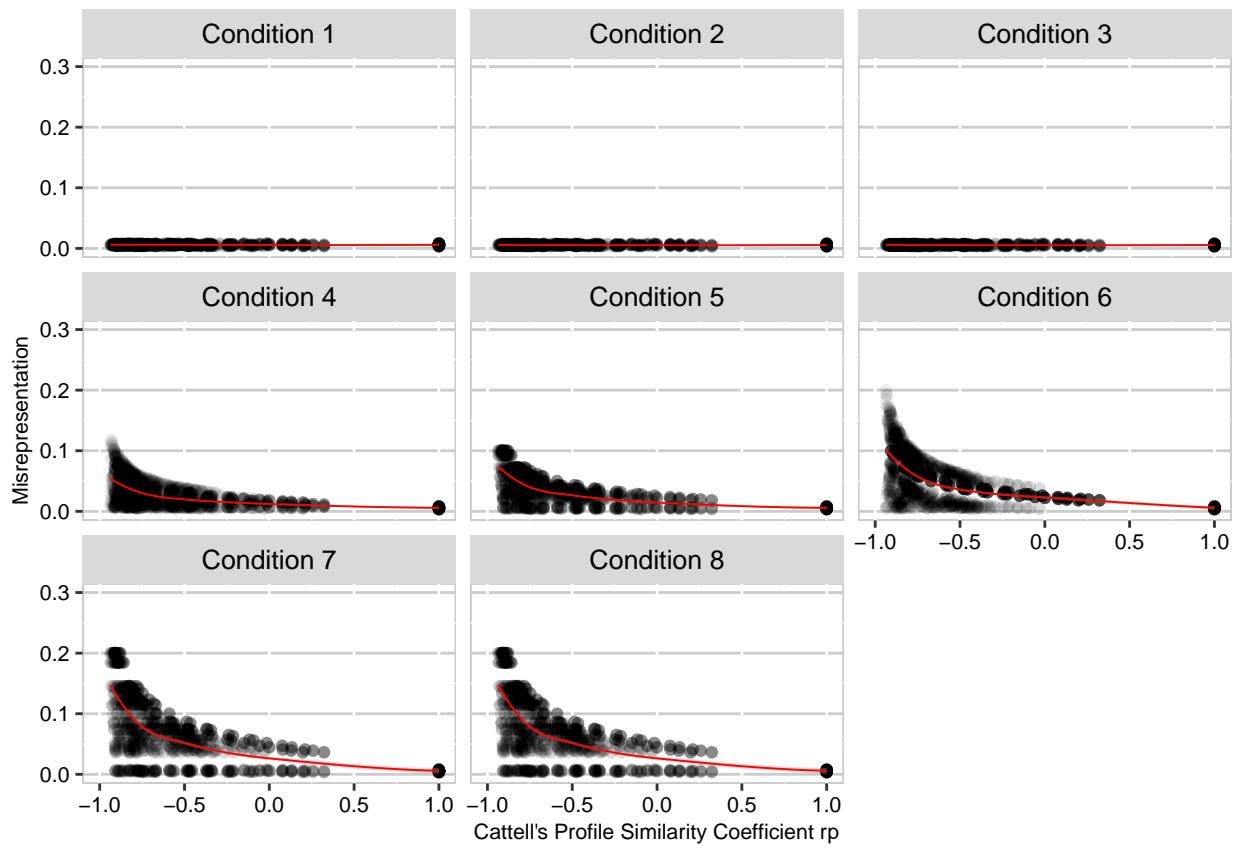
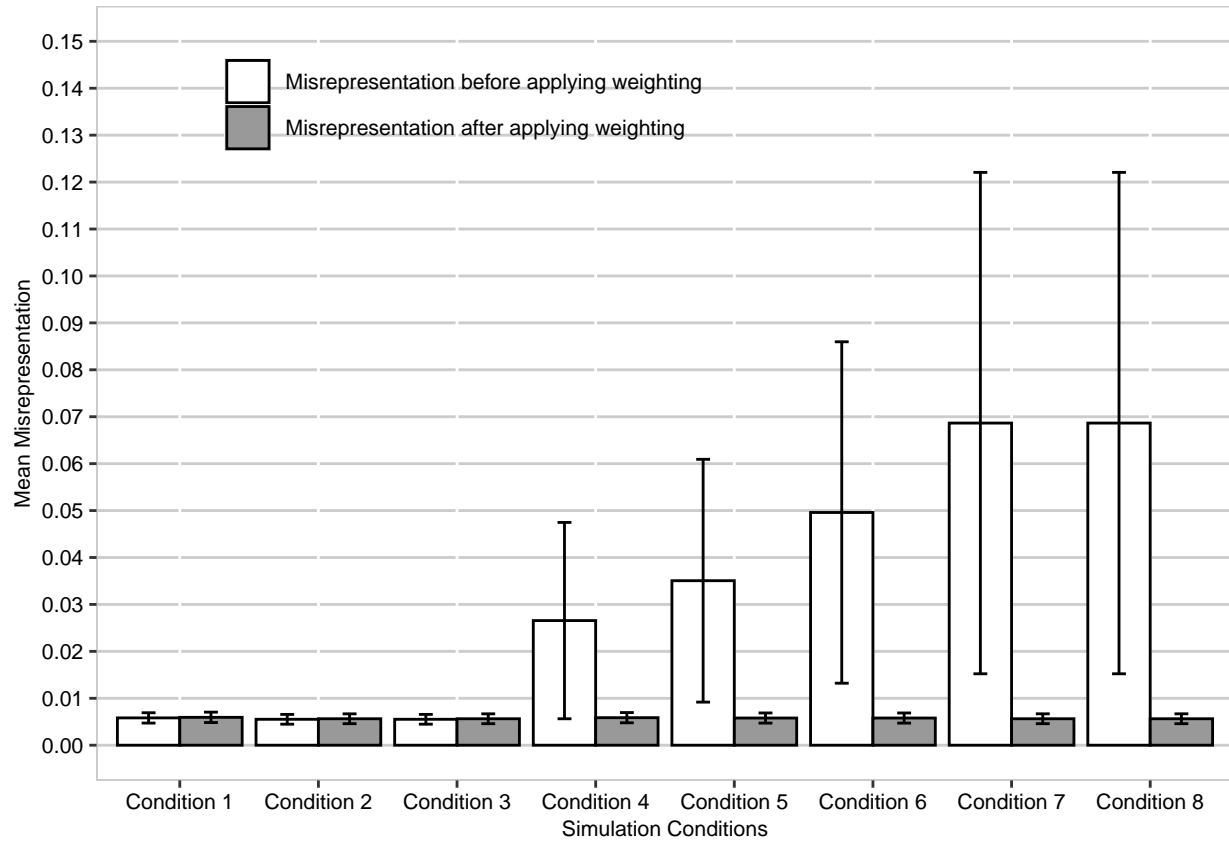


Figure 3

Effect of subgroup sampling rate match with distributional form on population misrepresentation.

**Figure 4**

Average absolute discrepancy (unweighted in white and weighted in grey) across the eight attitudinal conditions.