

<sup>1</sup> 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

**25 Nonresponse and Sample Weighting in Organizational Surveying**

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

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

51 population sampling.<sup>1</sup>

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

---

<sup>1</sup> 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.

<sup>2</sup> Frequently presented as a separate consideration, *measurement error* is an additional contributor to population misrepresentation and is not addressed via the weighting procedure. The concern of weighting is 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.

## 66 Nonresponse in Organizational Surveying

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

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

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

### 98 ***Form of Nonresponse***

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

---

<sup>3</sup> 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).

The more commonly encountered form of organizational nonresponse appears to be active variability - men, for example, have a higher proclivity toward active nonresponse than women (Luong & Rogelberg, 1998; Rogelberg et al., 2000; Spitzmüller et al., 2007). In organizational surveying baseline default expectation is that, *on average*, roughly 15% of respondents should be expected to be accurately characterized as “active” (Rogelberg et al., 2003; Rogelberg & Stanton, 2007; Werner et al., 2007). It is this second, less frequently observed form of nonresponse that also carries the greater resulting threat of biased sample estimates (see, for example, Kulas et al., 2017; Rogelberg & Stanton, 2007). It is these estimates that are the desired target of remediation when applying sample weights.

## Sample Weighting - a Brief Overview

Within public opinion polling contexts, when realized sample constituencies (e.g., male - by tradition from *carefully-constructed* and *randomly sampled* data frames)<sup>4</sup> are tested against census estimates of population parameters (e.g., 49% male), weights are applied to the sample in an effort to remediate the relative proportional under- or over-sampling. This is because, if the broader populations from which the under- or over-represented groups are sampled differ along surveyed dimensions (e.g., males, within the population, are *less likely to vote for Candidate X* than are women), then unweighted aggregate statistics (of, for example, projected voting results) will misrepresent the true population parameter. This remedial application of sample weights should also be considered a potential solution for researchers pursuing answers to analogous organizational pollings such as: “What is the mood of the employees?” This is because focused queries such as this are often

<sup>4</sup> 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.

136 course covertly complex - implicit in the question is a focus not on survey results, but rather  
 137 the broader employee population. Acknowledging the appropriate object of attribution is of  
 138 course important, because the next step (after gauging the mood of the surveyed  
 139 respondents) is *doing something* about it. Weighting may be a procedural option for  
 140 organizational surveyors to credibly transition a bit closer from, “What do the survey results  
 141 say”? to “What do the employees feel”?

142 **Procedural application**

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

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

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

- 153 1) Determine proportional weights for all levels within one stratum, and then assign these  
 154 weights to cases.
- 155 2) Determine proportional weights for a second group (ratio of population percent to  
 156 *current* sample percent [the current sample percent will be affected by the step 1  
 157 weighting procedure]). Multiply previous (step 1) weights by the proportional weights  
 158 for this second stratum and assign these new weights to cases.

159        3) Determine proportional weights for a third stratum (which will once again require  
160            re-inspection of the *current* sample percent). Multiply the previous step 2 weights by  
161            the third stratum proportional weights and assign to cases.

162        4) Iterate through steps 1, 2, and 3 (or more if more than three strata are considered)  
163            until the weighted sample characteristics match the population characteristics to your  
164            desired level of precision.

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

178            *Research question 1:* What role does response rate play in population  
179            misrepresentation?

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

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

183 biased<sup>5</sup> and unbiased sample estimates?

184 We view these questions as being analogous to similar questions asked and answered  
185 regarding differential *variable* weighting within the broader applied psychological disciplines.  
186 Just as, for example, there has been debate regarding the merits of differential versus unit  
187 variable weighting in a selection context or aggregate scale score definition (e.g., Bobko et al.,  
188 2007; Wainer, 1976), we propose that a similar consideration is appropriate with persons,  
189 and therefore compare and contrast unit versus proportional sample weighting.

190 **Methods**

191 We address our research questions within a simulated fictionalized context of  
192 organizational surveying (wherein it is common to assess estimates of employee attitude or  
193 perception; for example, commitment, culture/climate, engagement, satisfaction). We began  
194 the simulations by establishing “populations”, each consisting of 10,000 respondents  
195 characterized by demographic categorizations across gender (male and female) and  
196 department (A and B). We therefore had four demographic groups (Male.A, Male.B,  
197 Female.A, and Female.B). For these population respondents, we generated scaled continuous  
198 responses (real numbers) ranging from values of 1 to 5, representing averaged aggregate scale  
199 scores from a fictional multi-item survey with a common 1 → 5 Likert-type rating scale.

200 In order to represent different proportions of relative constituency (for example, more  
201 females than males or more department A workers than department B), we iterated  
202 population characteristics at marginal levels (gender and department) starting at 20% (and  
203 80%) with increments and corresponding decrements of 20%. For example, if males  
204 accounted for 20% of the simulated population, then females were 80%; also if respondents in  
205 Department A represented 60% of a population, then 40% were in Department B. Marginal  
206 constituencies were therefore realized at all combinations (across the two variables) of 20%

---

<sup>5</sup> 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.

207 and 80%, 40% and 60%, 60% and 40%, and 80% and 20%. This resulted in population *cell*  
208 constituencies (e.g., Male.A, Female.A, Male.B, Female.B) as low as 400 and as high as 6,400  
209 - see Figure 1 for further clarification of our “cell” and “margin” terminology and variable  
210 specification.

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

222 Individual attitudes were randomly sampled from population distributions at the cell  
223 level (e.g., Male.A) without replacement. These response rates (methodologically these could  
224 alternatively be conceptualized as *sampling* rates) were specified at 10% increments ranging  
225 from 60% to 90%, and these were fully iterated across each of our four marginal groups  
226 (Males, Females, Departments A and B). Our cell-level response rates therefore ranged from  
227 36% to 81% - a range of rates specified because they are approximations of reasonable  
228 expectations according to the organizational surveying literature (e.g., Mellahi & Harris,  
229 2016; Werner et al., 2007). We therefore investigated error within the aggregate mean (e.g.,  
230 grand mean aka total sample mean) attributable to different likelihoods of sample inclusion  
231 from constituent groups of different relative size and representing populations of different  
232 attitudinal distribution, but at response rates reasonably expected to exist in real-world

233 organizational surveying contexts.

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

242 In an attempt to capture this “degree of active nonresponse”, we calculated a simple  
243 index of response rate discrepancy (SD; presented in Table 2). The “least” active  
244 nonresponse scenarios are characterized by two subgroups with identical response rates and  
245 two having a slightly different response rate (e.g., male.a = 36%, female.a = 36%, male.b =  
246 42%, and female.b<sup>6</sup> = 42%; see the second row of Table 2, the SD index = .034)<sup>7</sup>. Also here  
247 note that three of our eight Table 1 conditions represent scenarios where the presence of  
248 active nonrespondents is not expected to result in bias (e.g., regardless of patterns of  
249 nonresponse, the unweighted sample mean is expected to yield an unbiased estimate of the  
250 population mean). These are Table 1 conditions one through three, where attitudinal  
251 distributions are of *the same form* across groups, regardless of any individual group response

---

<sup>6</sup> “Lowercase” specification of simulation strata indicates sample constituencies (e.g., male.b) whereas uppercase implicates population (e.g., Male.B).

<sup>7</sup> 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.

252 rate discrepancy from others'.

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

265 **Results**

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

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

278 “Misrepresentation” between sample and population was operationalized as: 1) the  
 279 discrepancies between the population and both weighted and unweighted sample means, as  
 280 well as, 2) the averaged deviation of these discrepancies from the population mean  
 281 (discrepancy in the “mean” of the means is bias, dispersion about the “mean” of the means is  
 282 error). If the average weighted sample mean was closer to the true population mean, relative  
 283 to the unweighted one, then the weighting was deemed beneficial.<sup>8</sup>

284 **Unweighted effects**

285 **Role of response rate**

286 Research question 1 asked what overall effect response rate has on population  
 287 misrepresentation. This is presented most directly in Figure 2, with *moderate* response rates  
 288 exhibiting the greatest degrees of misrepresentation across our simulated conditions. Note  
 289 here again that conditions 1 through 3, which represent populations with similar  
 290 distributions of attitude, do not exhibit misrepresentation regardless of response rate ( $\bar{d}_{Cond1}$   
 291 = 0.01,  $sd_{Cond1} = 0.00$ ;  $\bar{d}_{Cond2} = 0.01$ ,  $sd_{Cond2} = 0.00$ ;  $\bar{d}_{Cond3} = 0.01$ ,  $sd_{Cond3} = 0.00$ ). These  
 292 can be contrasted most particularly with conditions 6 ( $\bar{d}_{Cond6} = 0.05$ ,  $sd_{Cond6} = 0.04$ ), 7  
 293 ( $\bar{d}_{Cond7} = 0.07$ ,  $sd_{Cond7} = 0.05$ ), and 8 ( $\bar{d}_{Cond8} = 0.07$ ,  $sd_{Cond8} = 0.05$ ), which evidence  
 294 considerable misrepresentation, particularly so at moderate response rates (the greatest  
 295 degree of misrepresentation occurs with response rates ranging from roughly 40% to 70%).<sup>9</sup>  
 296 Discrepancies in unweighted means between samples and populations - regardless of response  
 297 rate - did broach statistical significance across the 8 conditions ( $F_{(7,32,760)} = 2,938.50$ ,  $p <$   
 298 .001). Tukey’s HSD revealed differences across all contrasts other than between Conditions 1,

---

<sup>8</sup> 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? Probably need Dr. Robinson to weigh in on this one

<sup>9</sup> Note that a confound exists whereby extreme overall rates (e.g., .36/.81) are necessarily associated with more 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.

299 2, and 3 and also Conditions 7 and 8. Retaining only Conditions 4 through 8, the  
 300 relationship between response rate and sample/population discrepancy was significant  
 301 beyond the effect of condition ( $\Delta R^2 = 0.00; F = 7,862.44$ ), and a polynomial response rate  
 302 term further added to the discrepancy prediction ( $\Delta R^2 = 0.02; F = 2,503.61$ ).<sup>10</sup>

303 **Role of nonresponse form**

304 Research question 2 asked what role the *form* of nonresponse (passive versus active)  
 305 plays in population misrepresentation. In terms of explaining the error that did emerge  
 306 within unweighted means sampled from conditions 4 though 8, this error was largely  
 307 attributable to form of nonresponse as operationalized by our SD index (See Figure 3).  
 308 Figure 3 also adds context to the Figure 2 response rate relationships, with the most extreme  
 309 misrepresentation paralleling circumstances of active nonresponse (e.g., to the “right” in  
 310 Figure 3).

311 The systematic patterns of heteroskedasticity of the Figure 3 scatterplots should also  
 312 be noted. There are *active nonresponse* scenarios in which no error is present (see, for  
 313 example, the lower right-hand portions of conditions 4 through 8 in Figure 3 where  
 314 discrepancy estimates of “0” appear all along the passive-active x-axis). These circumstances  
 315 are simulated conditions within which the response rates “parallel” the *population*  
 316 *distributional form*. For example, in Condition Eight, the distributional forms across  
 317 populations were: *PositiveSkewMale(A)*, *PositiveSkewMale(B)*, *NegativeSkewFemale(A)*,  
 318 *NegativeSkewFemale(B)*. Response rates that “mirror” distributional patterns in extreme  
 319 cases of active nonresponse (e.g., SD = .156; 54%<sub>Male(A)</sub>, 54%<sub>Male(B)</sub>, 81%<sub>Female(A)</sub>,

---

<sup>10</sup> 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*. Also need clarification on hierarchical regression (what is meant by response rate - how was that specified in the regression).

320 81%<sub>Female(B)</sub>) result in effectively zero error in the population mean approximation (average  
 321 discrepancy = 0.00,  $SD = 0.00$ ). Alternatively, when the response rates are inverted for the  
 322 SD=.156 cases, (e.g., 54%<sub>Male\_A</sub>, 81%<sub>Male\_B</sub>, 54%<sub>Female\_A</sub>, 81%<sub>Female\_B</sub>), there is substantial  
 323 error in approximation (average discrepancy = 0.16,  $SD = 0.03$ ). Here, it is not merely  
 324 response rate or form that is associated with biased sample estimates, but rather the nature  
 325 of response rate relative to existing attitudinal differences.<sup>11</sup>

326 ***Need to work on this section***

327 To further expand upon this *attitudinal form/pattern of nonresponse* interplay, the  
 328 discrepancies between population constituency and sampling proportions were additionally  
 329 evaluated through the lens of Cattell's profile similarity index ( $r_p$ , Cattell, 1949; Cattell et  
 330 al., 1966).  $r_p$  is sensitive to discrepancies in profile shape (pattern across profile components),  
 331 elevation (average component score), and scatter (sum of individual components' deviation  
 332 from the elevation estimate. Here, the profile similarity index references the relationship  
 333 between the response rates (NEED YANG TO VERIFY - THINK THIS IS  
 334 SSmale;SSfemale;SSdepta;SSdeptb from `combo` object) and sample sizes  
 335 (cellrate.ma;cellrate.mb;cellrate.fa;cellrate.gb) across experimental *cells*. For example,  
 336 VERIFY BEFORE CLARIFYING HERE. Figure 4 demonstrates the pattern of unweighted  
 337 sample mean deviation (from the population parameter) when this index is taken into  
 338 consideration. Specifically, Figure 4 demonstrates a more pronounced *form of* nonresponse  
 339 association when underlying attitudinal distributions evidence group differences (e.g.,  
 340 incrementally across the 8 specified conditions), and in these scenarios, active nonresponse is  
 341 shown to have a fairly large effect on error within the sample estimate (as well as  
 342 systematically increasing degrees of heteroskedasticity paralleling the Cattell index; omnibus  
 343 Breusch-Pagan [across conditions] = 3177.2,  $p < .001$ ). The curvilinear nature of these

---

<sup>11</sup> Don't think this is correct - maybe frame: "sometimes the active non-response is non-troublesome - when it fully parallels the distributional proportions (?)" ← still confusing. Looked at with Yang 3/1/24 and still confused - maybe leave in for reviewers to note and question.

<sup>344</sup> functions was estimated via hierarchical polynomial regression (excluding conditions 1, 2,  
<sup>345</sup> and 3), with misrepresentation exhibiting a linear association across condition ( $R^2 = 0.15$ ,  $p$   
<sup>346</sup>  $< .001$ ) as well as incrementally across the Cattell index ( $\Delta R^2 = 0.24$ ,  $p < .001$ ), and also  
<sup>347</sup> exhibiting an incremental polynomial effect ( $\Delta R^2 = 0.07$ ,  $p < .001$ ).

<sup>348</sup> To further elaborate this point, consider, for example, Condition 4 as presented in  
<sup>349</sup> Table 1. Here, three groups are characterized by similar distributions of attitudes (normally  
<sup>350</sup> distributed) and one, Female.B, is characterized by negatively skewed attitudes. The  
<sup>351</sup> greatest unweighted error here arises from sampling scenarios in which there are many  
<sup>352</sup> Female.B (e.g., in our specifications, 6,400) and fewer males and Department A females<sup>12</sup>,  
<sup>353</sup> but the female.b exhibit a much lower response rate (e.g., 20%) than do other groups, who  
<sup>354</sup> respond at a high rate (e.g., 80%). That is, it is not merely response rate, but response rate  
<sup>355</sup> within these identifiable groups, and whether or not those response rate differences parallel  
<sup>356</sup> underlying attitudinal differences that drives sample misrepresentation.

### <sup>357</sup> Impact of weighting

<sup>358</sup> Research question 3 was focused on the impact of weights on both biased (e.g.,  
<sup>359</sup> misrepresentative) and unbiased sample estimates<sup>13</sup>. Figure 5 provides a broad summary of  
<sup>360</sup> the results across the eight different attitudinal distribution conditions, presenting the  
<sup>361</sup> average absolute discrepancy from the population mean for the weighted and unweighted  
<sup>362</sup> sample estimates. Conditions one through three demonstrate that, on average, the  
<sup>363</sup> unweighted sample mean provides a good (unbiased) estimate of the population mean when  
<sup>364</sup> the distributional form does not differ across constituent groups (e.g., the distributions of

---

<sup>12</sup> Because of the “marginal” versus “cell” specifications of population constituencies, our most extreme example here necessarily results in 400 Male.A’s, 1,600 Male.B’s, and 1,600 Female.A’s. 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.

<sup>13</sup> Come back to this phrasing after decision is made on RQ 3 wording (whether to avoid using the term bias or not).

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 5). Although the *patterns* of unweighted sample mean discrepancies differed across conditions, all eight conditions exhibited similar omnibus effect (weighting ameliorating error wherever it arose [in the unweighted statistic]).

### **Weighting and Sampling Error**

Mean square error is our second index for sample quality. It is well-known that the application of weights increases (random) errors of precision, which was also empirically true in the current study. For each condition in our simulations, we calculated the standard deviations of 40.96 million unweighted and 40.96 million weighted samples means (4,096 possible population-sample combinations by 10,000 iterations), which yielded eight empirically-estimated standard errors of unweighted and weighted sample means. Figure 5 visually presents these standard errors in eight pairs of bars, demonstrating that the standard error of weighted sample means tended to be 16% to 18% larger than that of unweighted sample means regardless of condition (excluding Conditions 1-3). These errors highlight the caveat that weighting should only be applied in the active nonresponse case (e.g., although the aggregate effect of weighting with passive nonresponse is error-minimizing, any one sampling condition is *more likely* to result in greater deviation from the population parameter when weighting is applied to sample data driven by passive nonresponse).

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

As an aggregate across sampling events, weighting always corrects sample bias when it is present in the unweighted estimate. However, the standard errors suggest that for any *one* sampling event in the absence of bias, the likelihood that the sample mean approximates the *mean* of sample means is (slightly) greater for the unweighted estimate. When bias is present (in the unweighted estimate), there is obviously no advantage to “being closer” to

392 this biased mean of means. That is, under some circumstances, the mean of unweighted  
393 sample means does not center on the population mean. The implications of this seem quite  
394 obvious: Weighting should only be applied if bias is anticipated in the sample estimate. This  
395 may seem to be a picayune recommendation, but we note here that this advocacy is not  
396 heeded in public opinion polling applications, where the computation and application of  
397 weights are default procedures (CITES? - perhaps AAPOR standards or personal  
398 communication with polling agencies such as Gallop).

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

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

404           

### Discussion

405           We view nonresponse as a serious problem that should be addressed via repeated  
406 attempts to survey particularly reluctant or hard-to-reach respondents because nonresponse  
407 may be reasonably expected to be greatest in groups that are most unsatisfied [e.g., it may  
408 be typical for individuals representing these groups to have their responses diluted; see, for  
409 example, Taris and Schreurs (2007)]. However, several researchers have noted potentially  
410 misplaced relative emphasis on survey response rates, with Cook et al. (2000), Krosnick  
411 (1999), and Visser et al. (1996) articulating the point that representativeness of the sample is  
412 more important than response rate. We also believe that the goal in organizational surveying  
413 should be representativeness not exhaustiveness. Krosnick (1999) specifically comments that,  
414 even when probability sampling is employed, response rate does not necessarily implicate  
415 either good or poor sample representativeness. One aim of this paper is to reinforce this  
416 primary ‘representativeness’ orientation to those who may be otherwise inclined to focus on  
417 response rate as a sufficient index of quality (while also stressing sample weighting as a

418 practice that can potentially remediate *misrepresentativeness*).

419 With the above in mind, we set out to answer three fairly straightforward questions:  
420 What roles do 1) response rate and 2) form of nonresponse have on population  
421 misrepresentation, and 3) what impact does the application of weights have on the quality of  
422 sample estimates? The simulations demonstrate that the impact of mere response rate  
423 impact *depends* on the underlying distributions of population attitude. Conditions 1  
424 through 3 (as well as all other conditions) are occasionally immune to response rate  
425 influence, depending on whether the pattern of nonresponse parallels the pattern of  
426 attitudinal distribution differences or not). Active forms of nonresponse can harm the  
427 unweighted sample estimate, but only when the pattern of active nonresponse is  
428 accompanied by differing distributions of attitudes within the active nonrespondent  
429 “populations” [this would appear to be a reasonable expectation based on the literature; e.g.,  
430 Rogelberg et al. (2000); Rogelberg et al. (2003); Spitzmüller et al. (2007)]. Weighting  
431 “always” helps, as long as you capture the proper strata (which of course we were able to do  
432 via controlled simulation), but also... Although the weighted mean proved an unbiased  
433 estimate of the population mean across all simulations, in circumstances where no bias  
434 existed in the unweighted estimate, the trade-off between bias-correction and random error  
435 of precision (e.g., standard error) also needs to be acknowledged.

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

444 of active nonresponse.

445 Previous presentations have noted that bias is sometimes associated with nonresponse  
446 and othertimes it is not - this research has not been explicit in the specific conditions that  
447 moderate this association, however. The current paper does make this association explicit. It  
448 is not merely the form of nonresponse that determines whether or not bias occurs, but also  
449 the underlying distributions that the response probabilities are applied to. Some  
450 distributional patterns are immune to the biasing effects of active nonresponse (see, for  
451 example, Conditions 1 through 3). Some patterns of active nonresponse also result in no bias  
452 even when distributional patterns deviate substantially (see, for example, Condition 8 where  
453 a 20%, 20%, 80%, 80% response rate pattern exhibits no error). The target therefore should  
454 not be merely form of nonresponse but also underlying attitudes. Regardless, however,  
455 weighting always remediates the error when it occurs (and does not add error where it is  
456 absent).

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

467 It has been stated that active nonresponse is relatively harmless unless the actively  
468 nonrespondent group is relatively large [cites below]. The current study, however, suggests  
469 that post-data-collection remediation. There may also be some important implications here

470 regarding sample (and population) size. Because organizational surveyors likely interface  
471 with organizations of varying sizes (perhaps some of which are small- or medium-sized), the  
472 implications of our simulations particularly in the small population conditions, were  
473 highlighted. Findings specific to these conditions were: XXX, XXX, XXX.

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

479 It should also be pointed out that although the active nonrespondent group seems to  
480 be a great concern, it will not seriously bias the results unless the proportion of active  
481 nonrespondents is higher than 15% (Rogelberg et al., 2003; Rogelberg & Stanton, 2007;  
482 Werner et al., 2007). "In this study we found that the active nonrespondent group was  
483 relatively small (approximately 15%), but consistent in size with research conducted by."  
484 (Rogelberg et al., 2003, pp. 1110-1111). "Furthermore, consistent with Roth (1994) who  
485 stated that when missingness is not random (as we found for active nonrespondents),  
486 meaningful bias will only be introduced if the group is relatively large (which was not the  
487 case in this study)." (Rogelberg et al., 2003, p. 1112).

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

494 "It has been suggested that it takes a response rate of 85% to conclude that

495 nonresponse error is not a threat (Dooely & Lindner, 2003). We agree that researchers  
496 should provide both empirical and theoretical evidence refuting nonresponse bias whenever  
497 the response rate is less than 85%.” (Werner et al., 2007, p. 293).

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

503 Integration of IT/IS systems within HR functions hopefully assists the  
504 (un)likelihood that organizational population frames are either deficient or  
505 contaminated, although we note that this possibility (frame misspecification) is  
506 much more plausible within organizations that do not have updated or integrated  
507 HR IT/IS systems (perhaps, ironically, *smaller* organizations).

## 508 **Limitations**

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

## 518 **Future Directions**

519 Our operationalization of passive nonresponse was based on realized subsample  
520 differences in response rate. Of course it is plausible that consistent response rates (e.g., 36%,

521 36%, 36%, 36%) could have corresponding *non-sampled* elements who represent active  
522 non-response. Our methodology did not model these scenarios, but future like-minded  
523 investigations may wish to do so.

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

531 Experimental methods within the psychological discipline have long been criticized  
532 for heavy reliance on samples of convenience (for instance, student samples). Very little  
533 progress has been made regarding the application of appropriate population sampling  
534 procedures in experimentation. Certain non-experimental procedures (most notably  
535 organizational surveying) hold paradoxical advantage over experimental procedures primarily  
536 in this arena of sampling - particularly in consideration of population coverage, which refers  
537 to the percent of a population that is reachable by the sampling procedure (e.g., postal,  
538 intra-office, or internet invitation) and likelihood of having access to population parameter  
539 estimates (e.g., strata constituencies). There is a rich tradition and literature of public  
540 opinion polling procedures and techniques from which to draw. These procedures, however,  
541 only hold advantage if the non-experimental methodologist acknowledges the criticality of  
542 sample representativeness. The current paper provides one corrective technique  
543 (post-stratification weighting) as an important focus for the organizational surveyor who  
544 shares this primary interest in maximizing sample representativeness.

545 We note the above “advantage” held by organizational surveyors because extensions  
546 of the current protocol include investigating how inaccurate census estimates (and/or

547 grabbing the “wrong” group) affects the quality of sample estimates. That is, in our  
548 controlled simulations, we were able to know population constituencies, because they were  
549 set by us! In real-world applications, there is likely more error between the population  
550 estimate and actual population constituency. Similarly, if the association between attitude  
551 and group membership were to be controlled, there may be conditions identified whereby  
552 weighting loses its efficacy (e.g., low “correlations” between attitude and group membership).  
553 Future simulations should test boundary conditions for this type of error, identifying at what  
554 point inaccuracy in the population constituency estimate appreciably degrades the weighting  
555 procedure. Furthermore, it was demonstrated here that, when bias exists, weighting corrects  
556 it. Weighting also, however, results in a larger mean square error (MSE; expected spread of  
557 sample estimates around the population parameter). Feasibly then, there is a point at which  
558 the decreased bias is accompanied by an unacceptably inflated MSE. At which point does  
559 this occur? This is another fertile area for future exploration.

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

570 For Condition 5 (for example, low/high response rates with minority/majority  
571 population constituencies). The lower-right to upper-left diagonal reflects response rates that  
572 parallel population constituencies. The patterns across these stressors were consistent, with

573 the weighted sample means (red dots) providing unbiased estimates of the population  
574 parameter, whereas the unweighted sample means (grey dots) tended to yield unbiased  
575 estimates when the population constituencies were roughly equal (e.g., close to 50%/50%).

576 Figure 3 drills down this information further by extracting unweighted and weighted  
577 estimates in one specific marginal population parameter combination (here, 60% males and  
578 40% females; 40% in department A and 60% in department B). In doing so, the population  
579 parameters were in control and sample parameters were set free (see dotted red rectangle in  
580 Figure 2). Therefore, Figure 3 was then arranged in a fashion that allowed further  
581 investigation into the interactive effect of marginal sample parameters (gender on the x-axis  
582 and department on the y-axis) on the effectiveness of post-stratification weighting reflected  
583 by the pattern of grey and red dots. **Huh? - find old version or delete**

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

586 Mention tradition of single-item indicators in public opinion polling versus multi-item  
587 scales in Psychological assessment?

588 PRIOR TO RQs: after chatting with Yang (10/31/19) these need to be clarified  
589 a bit - reading 11/3 they make sense but need to be read very carefully. Check  
590 with Yang on 1/26 Skype. 2/1 revisions seem ok to Kulas. Three moving parts:  
591 underlying attitudinal distributions, response rate, and form of nonresponse <-  
592 perhaps we should make these variables more explicit prior to the  
593 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*

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

- 648                   *Psychiatry*, 48(4), 380–385.
- 649                   Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50(1),  
650                   537–567.
- 651                   Kulas, J. T., Robinson, D. H., Kellar, D. Z., & Smith, J. A. (2017). Nonresponse in  
652                   organizational surveying: Attitudinal distribution form and conditional response  
653                   probabilities' impact on patterns of bias. *Public Opinion Quarterly*, 81(2),  
654                   401–421.
- 655                   Kulas, J. T., Robinson, D. H., Smith, J. A., & Kellar, D. Z. (2016).  
656                   Post-stratification weighting in organizational surveys: A cross-disciplinary  
657                   tutorial. *Human Resource Management*.
- 658                   Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary  
659                   distinctions between organizational, mechanical turk, and other convenience  
660                   samples. *Industrial and Organizational Psychology*, 8(2), 142–164.
- 661                   Luong, A., & Rogelberg, S. G. (1998). How to increase your survey response rate.  
662                   *The Industrial-Organizational Psychologist*, 36(1), 61–65.
- 663                   Mellahi, K., & Harris, L. C. (2016). Response rates in business and management  
664                   research: An overview of current practice and suggestions for future direction.  
665                   *British Journal of Management*, 27(2), 426–437.
- 666                   Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with  
667                   data that are not missing completely at random. *Psychometrika*, 52(3), 431–462.
- 668                   Pasek, J. (2018). *Anesrake: ANES raking implementation*.  
669                   <https://CRAN.R-project.org/package=anesrake>
- 670                   Pedersen, M. J., & Nielsen, C. V. ek. (2016). Improving survey response rates in  
671                   online panels: Effects of low-cost incentives and cost-free text appeal interventions.  
672                   *Social Science Computer Review*, 34(2), 229–243.
- 673                   Quine, S., & Morrell, S. (2008). Feeling safe in one's neighbourhood: Variation by  
674                   location among older australians. *The Australian Journal of Rural Health*, 16,

- 675 115–116.
- 676 Rivers, D., & Bailey, D. (2009). Inference from matched samples in the 2008 US  
677 national elections. *Proceedings of the Joint Statistical Meetings*, 1, 627–639.
- 678 Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S., &  
679 Knight, W. E. (2003). Profiling active and passive nonrespondents to an  
680 organizational survey. *Journal of Applied Psychology*, 88(6), 1104.
- 681 Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee  
682 attitude surveys: Examining the attitudes of noncompliant employees. *Journal of  
683 Applied Psychology*, 85(2), 284.
- 684 Rogelberg, S. G., & Stanton, J. M. (2007). *Introduction: Understanding and dealing  
685 with organizational survey nonresponse*. Sage Publications Sage CA: Los Angeles,  
686 CA.
- 687 Spitzmüller, C., Glenn, D. M., Sutton, M. M., Barr, C. D., & Rogelberg, S. G. (2007).  
688 Survey nonrespondents as bad soldiers: Examining the relationship between  
689 organizational citizenship and survey response behavior. *International Journal of  
690 Selection and Assessment*, 15(4), 449–459.
- 691 Taris, T. W., & Schreurs, P. J. (2007). How may nonresponse affect findings in  
692 organizational surveys? The tendency-to-the-positive effect. *International Journal  
693 of Stress Management*, 14(3), 249.
- 694 Tett, R., Brown, C., & Walser, B. (2014). The 2011 SIOP graduate program  
695 benchmarking survey part 7: Theses, dissertations, and performance expectations.  
696 *The Industrial-Organizational Psychologist*, 51(4), 62–73.
- 697 Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for  
698 election forecasting? An evaluation of the columbus dispatch poll. *Public Opinion  
699 Quarterly*, 60(2), 181–227.
- 700 Wainer, H. (1976). Estimating coefficients in linear models: It don't make no  
701 nevermind. *Psychological Bulletin*, 83(2), 213.

- 702 Werner, S., Praxedes, M., & Kim, H.-G. (2007). The reporting of nonresponse  
703 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

## 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

Population Specification (N = 10,000):

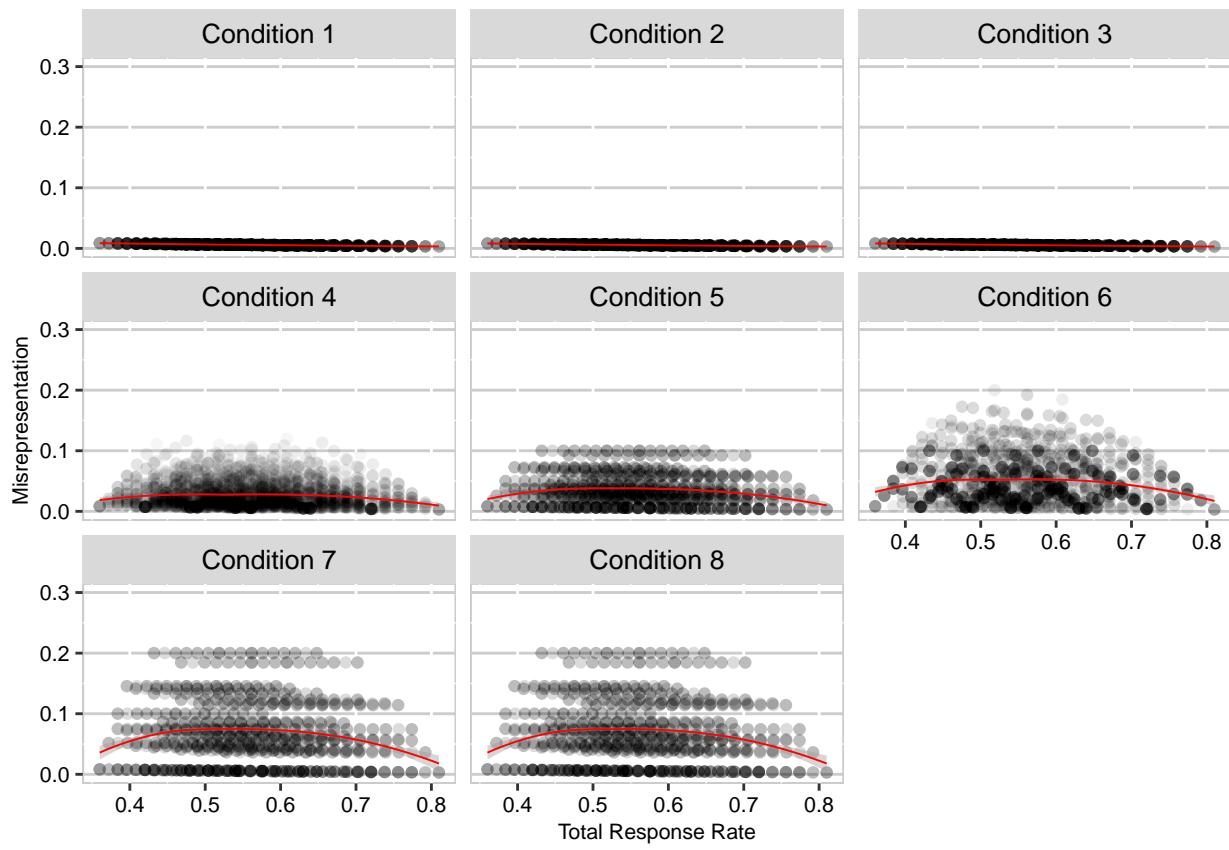
		Department		“Marginal” constituencies (department)
		A (4,000)	B (6,000)	
Gender	Male (2,000)	Male.A (800)	Male.B (1,200)	
	Female (8,000)	Female.A (3,200)	Female.B (4,800)	

“Marginal” constituencies (gender)

“Cell” constituencies

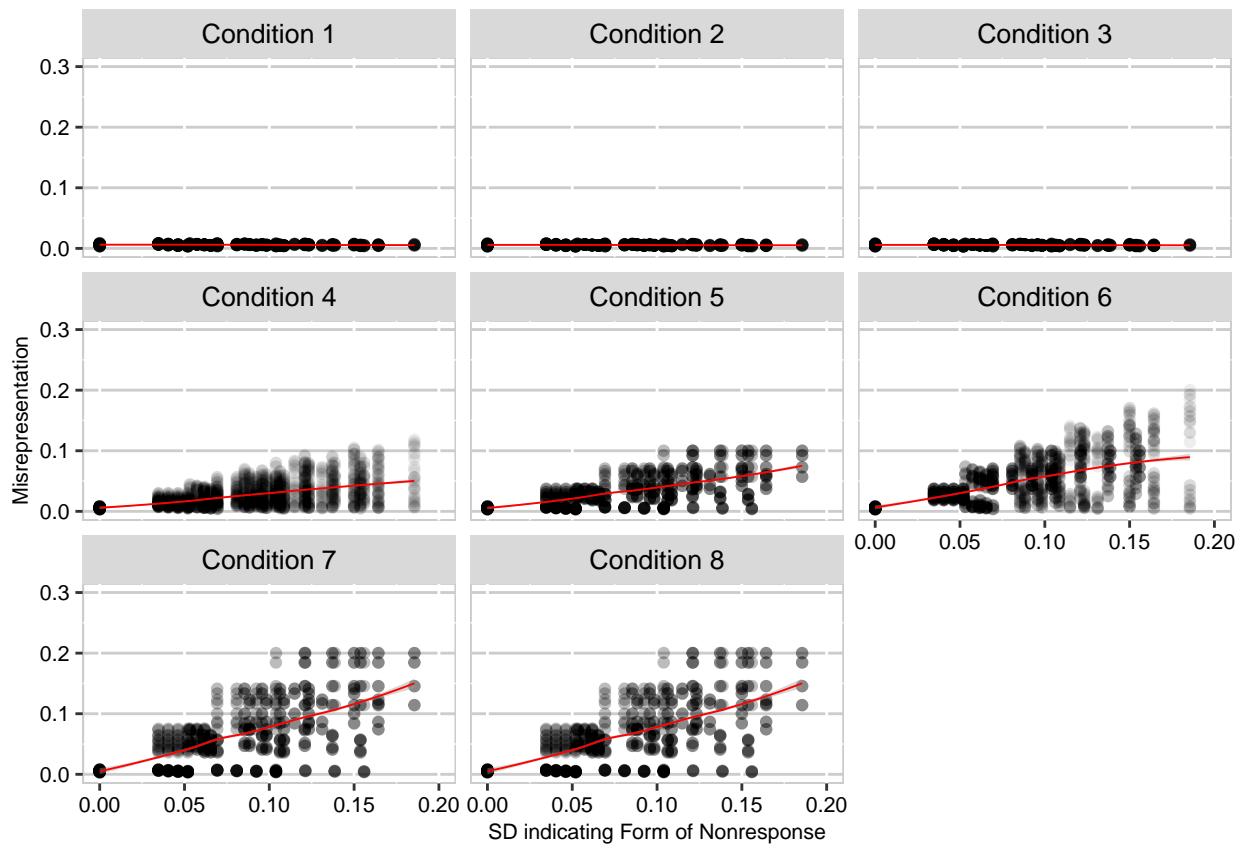
**Figure 1**

*Visual demonstrating terms used to describe population elements.*



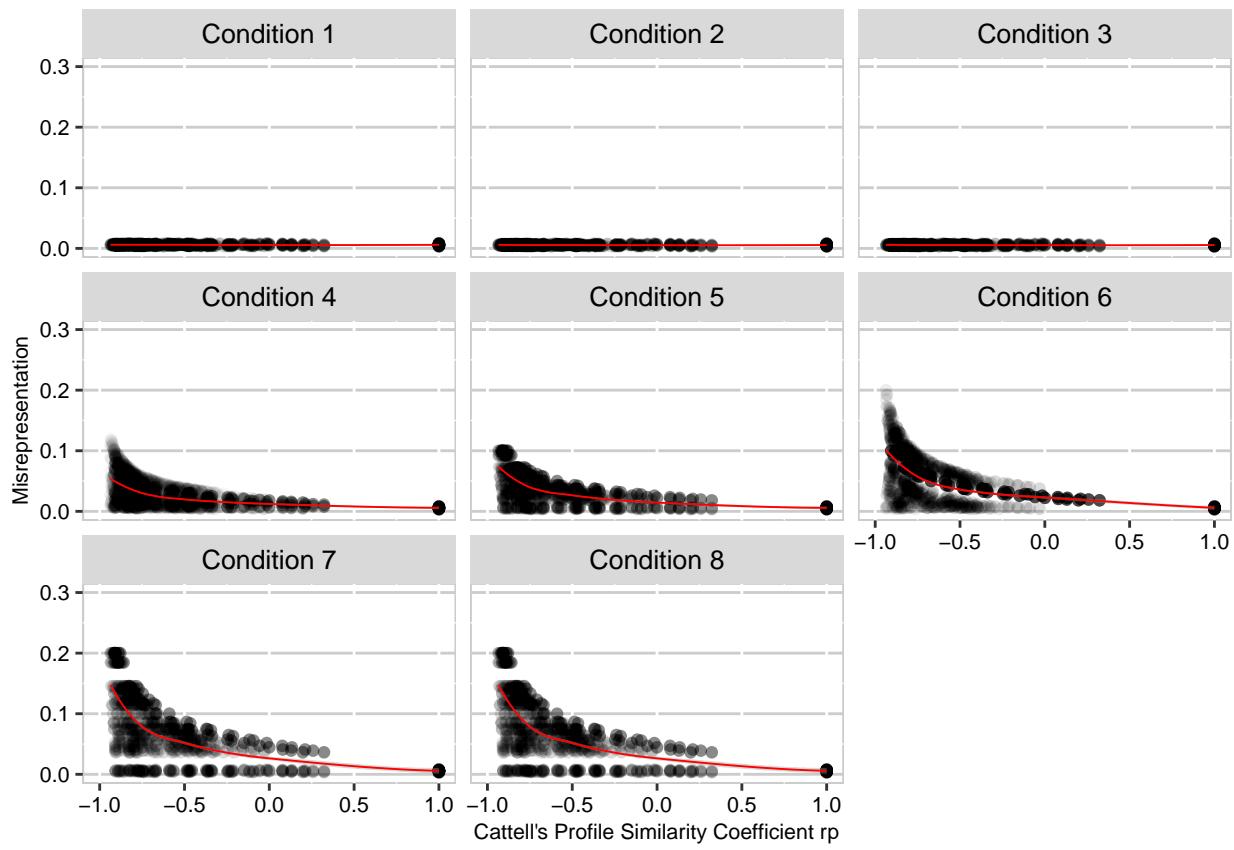
**Figure 2**

*Relationship between total response rate and misrepresentation.*

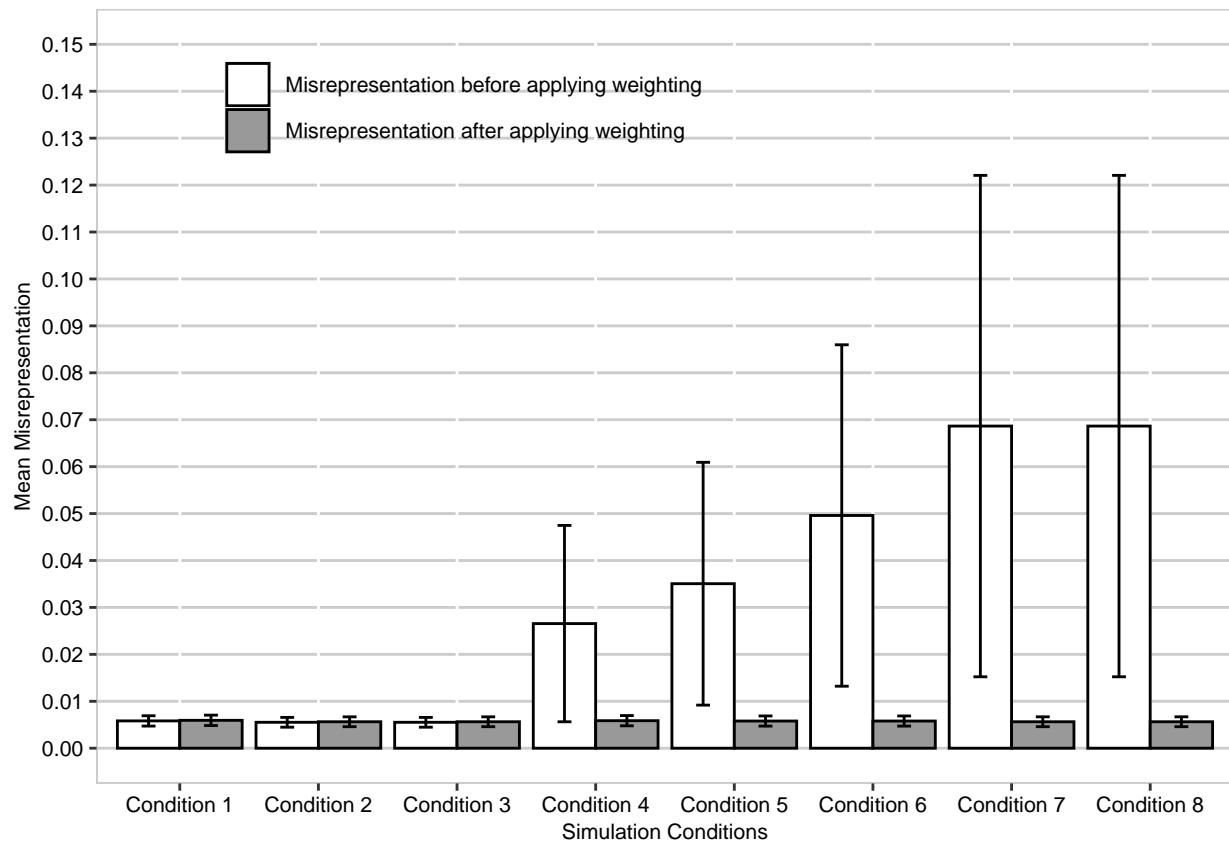


**Figure 3**

*Relationship between nonresponse form and misrepresentation.*

**Figure 4**

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



**Figure 5**

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