

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 relatively 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 sample representativeness in these scenarios may be 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]), sample weighting alters the proportional contributions of *individual*
30 *respondents* within a data set. Some respondents' responses are assigned greater relative
31 contribution and others are assigned less. This practice is commonplace in the summary of
32 general population polling data reflecting, for example, elections and politics (e.g., Rivers &
33 Bailey, 2009), prevalence rates of psychological disorders (e.g., Kessler et al., 2009), or
34 feelings of physical safety (e.g., Quine & Morrell, 2008). It is also seemingly in the periphery
35 of awareness and interest within the published organizational surveying literature (see, for
36 example, Kulas et al., 2016; 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, higher locally realized response rates have been interpreted
43 as indicators of 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.¹

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).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 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.³ 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).

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).

114 The more commonly encountered form of organizational nonresponse appears to be

115 passive (Rogelberg et al., 2003; Rogelberg & Stanton, 2007), although subgroup rates may

116 evidence variability - men, for example, have a higher proclivity toward active nonresponse

117 than do women (Luong & Rogelberg, 1998; Rogelberg et al., 2000; Spitzmüller et al., 2007).

118 The organizational surveying baseline default expectation is that, *on average*, roughly 15% of

119 nonrespondents should be expected to be accurately characterized as “active” (Rogelberg et

120 al., 2003; Rogelberg & Stanton, 2007; Werner et al., 2007). It is this second, less frequently

121 anticipated form of nonresponse that also carries the greater resulting threat of biased sample

122 estimates (see, for example, Kulas et al., 2017; Rogelberg & Stanton, 2007). It is these

123 biased estimates that are the desired target of remediation when applying sample weights.

124 **Sample Weighting - a Brief Overview**

125 Within public opinion polling contexts, when realized sample constituencies (e.g.,

126 44% male - by tradition from *carefully-constructed* and *randomly sampled* data frames)⁴ are

127 compared against census estimates of population parameters (e.g., 49% male), weights are

128 applied to the sample in an effort to remediate the relative proportional under- or

129 over-sampling. This is because, if the broader populations from which the under- or

130 over-represented groups are sampled differ along surveyed dimensions (e.g., males, within the

131 population, are *less likely to vote for Candidate X* than are women), then unweighted

132 aggregate statistics (of, for example, projected voting results) will misrepresent the true

133 population parameter. This remedial application of sample weights should also be considered

134 an option for researchers pursuing answers to analogous organizational pollings such as:

135 “What is the mood of the employees?” This is because focused queries such as this are of

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.

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 biased⁵ 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
205 in Department A represented 60% of a population, then 40% were in Department B.
206 Marginal constituencies were therefore realized at all combinations (across the two variables)

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.

207 of 20% and 80%, 40% and 60%, 60% and 40%, and 80% and 20%. This resulted in
208 population *cell* constituencies (e.g., Male.A, Female.A, Male.B, Female.B) as low as 400 and
209 as high as 6,400 - see Figure 1 for further clarification of our “cell” and “margin” terminology
210 and relative constituency 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 retained in an attempt to model similarities and discrepancies in
214 construct standing (e.g., commitment, satisfaction, or engagement) across localized
215 respondent groupings. The normal distribution exhibited, on average, a mean of 3.0 whereas
216 the skewed distributions were characterized by average means of 2.0 and 4.0, respectively. In
217 total, eight crossings of distributional type across employee categorization were specified
218 (Table 1 presents the combinations of these distributions). Note that these eight conditions
219 are not exhaustive of all possible combinations of constituent groups and attitudinal
220 distribution - we limited the simulations to combinations projected to collectively be most
221 informative within the manipulated simulation parameters.

222 Individual attitudes were randomly sampled from population distributions at the cell
223 level (e.g., Male.A) without replacement. These response rates (methodologically within the
224 simulation these could equally be conceptualized as *sampling* rates) were specified at 10%
225 increments ranging from 60% to 90%, and these were fully iterated across each of our four
226 marginal groups (Males, Females, Departments A and B). Our cell-level response rates
227 therefore ranged from 36% to 81% - a range of rates that encompass reasonable real-world
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 our operationalization of active versus passive utilizes
235 *consistency* of response rate as a baseline indicator of passive nonresponse. There are several
236 patterns of response that are therefore intended to represent sampling scenarios reflecting
237 passive nonresponse across groups, *regardless of response rate*. These are the scenarios in
238 which all subgroups exhibit the same response rate (e.g., 36%, 36%, 36%, and 36%). All
239 other combinations of response rate are intended operationalizations of active forms of
240 nonresponse (e.g., not *as reasonably* characterized as missing at random).

241 In an attempt to capture the “degree of active nonresponse”, we calculated a simple
242 index of response rate discrepancy (SD; presented in Table 2). The “least” active
243 nonresponse scenarios are characterized by two subgroups with identical response rates and
244 two having a slightly different response rate (e.g., male.a = 36%, female.a = 36%, male.b =
245 42%, and female.b⁶ = 42%; see the second row of Table 2, the SD index = .034)⁷. Also here
246 note that three of our eight Table 1 conditions represent scenarios where the presence of
247 active nonrespondents is not expected to result in bias (e.g., regardless of patterns of
248 nonresponse, the unweighted sample mean is expected to yield an unbiased estimate of the
249 population mean). These are Table 1 conditions one through three, where attitudinal
250 distributions are of *the same form* across groups, regardless of any individual group response
251 rate discrepancy from others’.

⁶ Throughout the Method and Results, “lowercase” specification of simulation strata indicates sample constituencies (e.g., male.b) whereas uppercase implicates population (e.g., Male.B).

⁷ 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.

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

Results

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

We were most interested in comparing the extent to which unweighted (aggregated responses without raking) and weighted (aggregated weighted) sample means approximated the known population means across our controlled specifications of response rate, nonresponse form, and attitudinal distribution. Population means were extracted from each iteration, as the simulations specified a new population at each iteration. “Misrepresentation” between sample and population was operationalized as: 1) the discrepancies between the

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

283 **Unweighted effects**

284 **Role of response rate**

285 Research question 1 asked what singular effect response rate has on population
 286 misrepresentation. This is presented most concisely in Figure 2, with *moderate* response
 287 rates exhibiting the greatest degrees of misrepresentation across our simulated conditions.
 288 Note here again that conditions 1 through 3, which represent populations with similar
 289 distributions of attitude, do not exhibit misrepresentation regardless of response rate (\bar{d}_{Cond1}
 290 = 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
 291 can be contrasted most particularly with conditions 6 ($\bar{d}_{Cond6} = 0.05$, $sd_{Cond6} = 0.04$), 7
 292 ($\bar{d}_{Cond7} = 0.07$, $sd_{Cond7} = 0.05$), and 8 ($\bar{d}_{Cond8} = 0.07$, $sd_{Cond8} = 0.05$), which evidence
 293 considerable misrepresentation, particularly so at moderate response rates (the greatest
 294 degree of misrepresentation occurs with aggregate response rates ranging from roughly 40%
 295 to 70%)⁹. Note also that all conditions exhibit circumstances where low and moderate
 296 response rates result in no misrepresentation.

297 Discrepancies in unweighted means between samples and populations – regardless of
 298 response rate – did exhibit differences across the 8 conditions ($F_{(7,32,760)} = 2,938.50$, $p <$

⁸ 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

⁹ Note that 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.

.001). Tukey's HSD implicated all contrasts other than between Conditions 1, 2, and 3 and also between Conditions 7 and 8. Retaining only Conditions 4 through 8, the relationship between response rate and sample/population discrepancy was estimated via hierarchical regression (step 1 = condition, step 2 = response rate, step 3 = squared response rate). In these analyses there was a significant but trivial linear response rate influence *beyond* the effect of condition ($\Delta R^2 = 0.00; F = 7,862.44$), with the polynomial response rate term further adding slightly to the discrepancy prediction ($\Delta R^2 = 0.02; F = 2,503.61$). Collectively these results reflect inconsistent direct relationships between response rate and population misrepresentation – a range of representative/error-filled estimates were encountered all along the response rate continuum. The next sections explore potential explanatory mechanisms for these ranges of misrepresentation at identical rates of response.

310 **Role of nonresponse *form***

311 Research question 2 asked what role the *form* of nonresponse (passive versus active) 312 plays in population misrepresentation. In terms of explaining the error that did emerge 313 within unweighted means sampled from conditions 4 though 8, this error was largely 314 attributable to form of nonresponse as operationalized by our SD index (See Figure 3). 315 Figure 3 also adds context to the previously noted Figure 2 variabilities in ranges of 316 misrepresentation across response rates, with the most extreme Figure 3 cases of 317 misrepresentation fully echoing circumstances of active nonresponse (e.g., the greatest cases 318 of misrepresentation are always associated with the highest SD index regardless of simulation 319 condition).

320 The Figure 3 scatterplots also reveal patterned heteroskedasticity across the active 321 nonresponse continuum. Similar to the response rate – misrepresentation associations, there 322 are *active nonresponse* scenarios in which no misrepresentation occurs (see, for example, the 323 lower right-hand portions of conditions 4 through 8 where discrepancy estimates of “0” 324 persist at multiple points along the passive-active x-axis). These circumstances are

325 simulated conditions within which the response rate patterns “do not mirror” the *population*
326 *distributional form*. For example, in Condition Eight, the distributional forms across
327 populations were: *PositiveSkew_{Male(A)}*, *PositiveSkew_{Male(B)}*, *NegativeSkew_{Female(A)}*,
328 *NegativeSkew_{Female(B)}*. Response rates that “track along with” distributional patterns,
329 when also characterized by extreme cases of active nonresponse (e.g., SD = .156; 54%_{Male(A)},
330 54%_{Male(B)}, 81%_{Female(A)}, 81%_{Female(B)}), result in *substantial error* in the population mean
331 approximation (average discrepancy = 0.16, SD = 0.03). Alternatively, when the response
332 rates are “not aligned” with distributional patterns for the SD=.156 cases, (e.g., 54%_{Male_A},
333 81%_{Male_B}, 54%_{Female_A}, 81%_{Female_B}), there is very little error in approximation (average
334 discrepancy = 0.00, SD = 0.00; See Figure 6 for visual reference). Here, it is not simply
335 response rate or form that is associated with biased sample estimates, but rather the nature
336 of response rate relative to existing attitudinal differences.

337 ***Attitudinal distribution and patterns of nonresponse***

338 To further expand upon this *attitudinal form/pattern of nonresponse* interplay, the
339 discrepancies between population constituency and sampling proportions were additionally
340 evaluated through the lens of Cattell’s profile similarity index (r_p , Cattell, 1949; Cattell et
341 al., 1966). r_p is sensitive to discrepancies in profile shape (pattern across profile components),
342 elevation (average component score), and scatter (sum of individual components’ deviation
343 from the elevation estimate. Here, the profile similarity index references the relationship
344 between the response rates (NEED YANG TO VERIFY - THINK THIS IS
345 SSmale;SSfemale;SSdepta;SSdeptb from `combo` object) and sample sizes
346 (cellrate.ma;cellrate.mb;cellrate.fa;cellrate.gb) across experimental *cells*. For example,
347 VERIFY BEFORE CLARIFYING HERE. Figure 4 demonstrates the pattern of unweighted
348 sample mean deviation (from the population parameter) when this index is taken into
349 consideration. Specifically, Figure 4 demonstrates a more pronounced *form of* nonresponse
350 association when underlying attitudinal distributions evidence group differences (e.g.,
351 incrementally across the 8 specified conditions), and in these scenarios, active nonresponse is

352 shown to have a fairly large effect on error within the sample estimate (as well as
353 systematically increasing degrees of heteroskedasticity paralleling the Cattell index; omnibus
354 Breusch-Pagan [across conditions] = 3177.2, $p < .001$). The curvilinear nature of these
355 functions was estimated via hierarchical polynomial regression (excluding conditions 1, 2,
356 and 3), with misrepresentation exhibiting a linear association across condition ($R^2 = 0.15$, p
357 $< .001$) as well as incrementally across the Cattell index ($\Delta R^2 = 0.24$, $p < .001$), and also
358 exhibiting an incremental polynomial effect ($\Delta R^2 = 0.07$, $p < .001$).

359 To further elaborate this point, consider, for example, Condition 4 as presented in
360 Table 1. Here, three groups are characterized by similar distributions of attitudes (normally
361 distributed) and one, Female.B, is characterized by negatively skewed attitudes. The
362 greatest unweighted error here arises from sampling scenarios in which there are many
363 Female.B (e.g., in our specifications, 6,400) and fewer males and Department A females¹⁰,
364 but the female.b exhibit a much lower response rate (e.g., 20%) than do other groups, who
365 respond at a high rate (e.g., 80%). That is, it is not merely response rate, but response rate
366 within these identifiable groups, and whether or not those response rate differences parallel
367 underlying attitudinal differences that drives sample misrepresentation.

368 Impact of weighting

369 Research question 3 was focused on the impact of weights on both biased (e.g.,
370 misrepresentative) and unbiased sample estimates¹¹. Figure 5 provides a broad summary of
371 the results across the eight different attitudinal distribution conditions, presenting the
372 average absolute discrepancy from the population mean for the weighted and unweighted

¹⁰ 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.

¹¹ Come back to this phrasing after decision is made on RQ 3 wording (whether to avoid using the term bias or not).

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 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 a second important 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

400 one sampling event in the absence of bias, the likelihood that the sample mean approximates
401 the *mean* of sample means is (slightly) greater for the unweighted estimate. When bias is
402 present (in the unweighted estimate), there is obviously no advantage to “being closer” to
403 this biased mean of means. That is, under some circumstances, the mean of unweighted
404 sample means does not center on the population mean. The implications of this seem quite
405 obvious: Weighting should only be applied if bias is anticipated in the sample estimate. This
406 may seem to be a picayune recommendation, but we note here that this advocacy is not
407 heeded in public opinion polling applications, where the computation and application of
408 weights are default practice (**CITES?** - perhaps AAPOR standards or personal
409 communication with polling 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 and 3
414 (**put SD on Figure 3 - color code**)

415

Discussion

416 We view nonresponse as a serious problem that should be addressed via repeated
417 attempts to survey particularly reluctant or hard-to-reach respondents because nonresponse
418 may be reasonably expected to be greatest in groups that are most unsatisfied (e.g., Taris &
419 Schreurs, 2007). However, several researchers have noted potentially misplaced relative
420 emphasis on response rates, with Cook et al. (2000), Krosnick (1999), and Visser et al.
421 (1996) articulating the point that representativeness of the sample is more important than
422 response rate. We also believe that the goal in organizational surveying should be
423 representativeness as opposed to exhaustiveness. Krosnick (1999) specifically comments that,
424 even when probability sampling is employed, response rate does not necessarily implicate
425 either good or poor sample representativeness. One aim of this paper is to stress this

426 primary ‘representativeness’ orientation to those who may be otherwise inclined to focus on
427 response rate as a sufficient index of quality (while also stressing sample weighting as a
428 practice that can potentially remediate *misrepresentativeness*).

429 With the above in mind, we set out to answer three fairly straightforward questions:
430 What roles do 1) response rate and 2) form of nonresponse have on population
431 misrepresentation, and 3) what impact does the application of weights have on the quality of
432 sample estimates? The simulations demonstrate that the impact of (mere) response rate is
433 contingent on the underlying distributions of population attitude. The effect is moderated –
434 there is not a simple relationship between response rate and misrepresentation. Rather, a
435 wide range of representative/ error-filled estimates can be expected all along the response
436 rate continuum. See Figure 2. Conditions 1 through 3 are fully immune and all other
437 conditions are occasionally immune to response rate influence.

438 Regarding question 2, Figure 3 shows that the largest misrepresentation sampling
439 scenarios are associated with *greater degrees of* active nonresponse. However, there also exist
440 active nonresponse scenarios within which little or no misrepresentation occurs. Active forms
441 of nonresponse can harm the unweighted sample estimate, but only when the pattern of
442 active nonresponse is accompanied by differing distributions of attitudes within the active
443 nonrespondent “populations” (this would appear to be a reasonable expectation based on the
444 literature, e.g., Rogelberg et al., 2000, 2003; Spitzmüller et al., 2007). Weighting “always”
445 helps, as long as you capture the proper strata (which of course we were able to do via
446 controlled simulation).

447 The results are presented with at least three primary takeaways: 1) our simulations
448 are comprehensive, iterating through all possible combinations of response rates – those
449 paralleling population distributions, those inversely mirroring population distributions, and
450 those “orthogonal to” population distributions, 2) the “SD” operationalization of passive to
451 active forms of nonresponse is a bit crude and insensitive to specific combinations of response

452 rates expected to manifest or not manifest in bias, and 3) substantial bias may be present in
453 the unweighted estimate even with only small proportions of active non-response (e.g., only
454 one or two groups exhibiting slightly different response rates, with the resulting discrepancy
455 [population versus sample mean] being quite large).

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

465 It has been stated that active nonresponse is relatively harmless unless the actively
466 nonrespondent group is relatively large [the proportion of active nonrespondents is higher
467 than 15%; Dooley and Lindner (2003); Rogelberg and Stanton (2007); Rogelberg et al.
468 (2003); Werner et al. (2007)]. Note here however, the possible disconnect between the reports
469 of 15% active nonresponse and declining response rates (trending toward 50%). Certainly
470 with decreasing overall response rates, the likely reasons would appear to be more active
471 than passive (e.g., it is difficult to entertain the idea that potential respondents are more
472 likely to forget to respond today than they were 40 years ago). Although the weighted mean
473 proved an unbiased estimate of the population mean across all simulations, in circumstances
474 where no bias existed in the unweighted estimate, the trade-off between bias-correction and
475 random error of precision (e.g., standard error) also needs to be acknowledged.

476 The current findings are of course qualified by the uniqueness of our simulations,
477 most notably our ability to fully capture the correct population parameters (e.g., because

478 these were “created” by us, we were also able to identify these strata as the nonresponse
479 contributors). Even in the extreme conditions (e.g., a small “population” with a
480 correspondingly low response rate), the weighting algorithm was able to provide a bias
481 correction. This is undoubtedly attributable to our random sampling procedure (instead of,
482 for example, sampling conditionally from the population distributions), but here we do note
483 that the raking procedure is applied at the “margins” (e.g., variable level, not intersectional
484 level), although our introduction of a biasing element is at the cell (intersection) level.

485 Future Directions

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

491 Our operationalization of passive nonresponse was based on realized subsample
492 differences in response rate. Of course it is plausible that consistent response rates (e.g., 36%,
493 36%, 36%, 36%) could have corresponding *non-sampled* elements who represent active
494 non-response. Our methodology did not model these scenarios, but future like-minded
495 investigations may wish to do so.

496 There may also be some important implications here regarding sample (and
497 population) size. Because organizational surveyors likely interface with organizations of
498 varying sizes (perhaps some of which are small- or medium-sized), the implications of our
499 simulations particularly in the small population conditions, were highlighted. Findings
500 specific to these conditions were: XXX, XXX, XXX.

501 *Research question placeholder:* What are the important interrelationships between
502 nonresponse form, response rate, and underlying distributional attributes that impact

503 population misrepresentation?

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

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

525 We note the above “advantage” held by organizational surveyors because extensions
526 of the current protocol include investigating how inaccurate census estimates (and/or
527 grabbing the “wrong” group) affects the quality of sample estimates. That is, in our
528 controlled simulations, we were able to know population constituencies, because they were

529 set by us! In real-world applications, there is likely more error between the population
530 estimate and actual population constituency. Similarly, if the association between attitude
531 and group membership were to be controlled, there may be conditions identified whereby
532 weighting loses its efficacy (e.g., low “correlations” between attitude and group membership).
533 Future simulations should test boundary conditions for this type of error, identifying at what
534 point inaccuracy in the population constituency estimate appreciably degrades the weighting
535 procedure. Furthermore, it was demonstrated here that, when bias exists, weighting corrects
536 it. Weighting also, however, results in a larger mean square error (MSE; expected spread of
537 sample estimates around the population parameter). Feasibly then, there is a point at which
538 the decreased bias is accompanied by an unacceptably inflated MSE. At which point does
539 this occur? This is another fertile area for future exploration.

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

550 For Condition 5 (for example, low/high response rates with minority/majority
551 population constituencies). The lower-right to upper-left diagonal reflects response rates that
552 parallel population constituencies. The patterns across these stressors were consistent, with
553 the weighted sample means (red dots) providing unbiased estimates of the population
554 parameter, whereas the unweighted sample means (grey dots) tended to yield unbiased

555 estimates when the population constituencies were roughly equal (e.g., close to 50%/50%).

556 Figure 3 drills down this information further by extracting unweighted and weighted
557 estimates in one specific marginal population parameter combination (here, 60% males and
558 40% females; 40% in department A and 60% in department B). In doing so, the population
559 parameters were in control and sample parameters were set free (see dotted red rectangle in
560 Figure 2). Therefore, Figure 3 was then arranged in a fashion that allowed further
561 investigation into the interactive effect of marginal sample parameters (gender on the x-axis
562 and department on the y-axis) on the effectiveness of post-stratification weighting reflected
563 by the pattern of grey and red dots. **Huh? - find old version or delete**

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

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

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

574 Our operationalization of active nonresponse as subgroup differences in response rates
575 of course merits validation. The literature suggests that individuals with... this Whether or
576 not subgroup differences in response rate can (or should) be investigated as potential
577 indication of active nonresponse is an empirical question and future investigations would
578 benefit from exploring the extent to which such variability in simple response rate across
579 constituent groups *should* be interpreted as indicative of active nonresponse. This would be

580 an extension of Taris and Schreurs (2007), who noted that selection of an individual
581 population element into a realized sample may in fact be predictable (because of, for
582 example, an increased likelihood of not responding when dissatisfied or disgruntled). This
583 operationalization is dependent on subgroup comparison (e.g., is not reflective of an entire
584 organization that collectively exhibits active nonresponse).

585 Most likely nonrespondents are actually those in the middle (not extremely
586 dissatisfied or extremely satisfied).

587 “put differently, a high response rate may not allow for valid inferences and a
588 lower response rate might adequately represent the broader population” [p. 1574;
589 Holtom et al. (2022)].

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

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

605 much more plausible within organizations that do not have updated or integrated

606 HR IT/IS systems (perhaps, ironically, *smaller* organizations).

References

- 607
- 608 Anseel, F., Lievens, F., Schollaert, E., & Choragwicka, B. (2010). Response rates in
609 organizational science, 1995–2008: A meta-analytic review and guidelines for survey
610 researchers. *Journal of Business and Psychology*, 25(3), 335–349.
- 611 Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*.
612 <https://github.com/crsh/papaja>
- 613 Baruch, Y. (1999). Response rate in academic studies—a comparative analysis. *Human*
614 *Relations*, 52(4), 421–438.
- 615 Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational
616 research. *Human Relations*, 61(8), 1139–1160.
- 617 Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality* (Vol. 335). John Wiley
618 & Sons.
- 619 Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating
620 composite scores: A literature review, application to content validity, and meta-analysis.
621 *Organizational Research Methods*, 10(4), 689–709.
- 622 Cattell, R. B. (1949). R p and other coefficients of pattern similarity. *Psychometrika*, 14(4),
623 279–298.
- 624 Cattell, R. B., Coulter, M. A., & Tsujioka, B. (1966). The taxonometric recognition of types
625 and functional emergents. *Handbook of Multivariate Experimental Psychology*, 288–329.
- 626 Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web-or
627 internet-based surveys. *Educational and Psychological Measurement*, 60(6), 821–836.
- 628 Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the
629 index of consumer sentiment. *Public Opinion Quarterly*, 64(4), 413–428.
- 630 Cycyota, C. S., & Harrison, D. A. (2002). Enhancing survey response rates at the executive
631 level: Are employee-or consumer-level techniques effective? *Journal of Management*,
632 28(2), 151–176.
- 633 Cycyota, C. S., & Harrison, D. A. (2006). What (not) to expect when surveying executives:

- 634 A meta-analysis of top manager response rates and techniques over time. *Organizational*
635 *Research Methods*, 9(2), 133–160.
- 636 Deming, W. E., & Stephan, F. F. (1940). On a least squares adjustment of a sampled
637 frequency table when the expected marginal totals are known. *The Annals of*
638 *Mathematical Statistics*, 11(4), 427–444.
- 639 Dooley, L. M., & Lindner, J. R. (2003). The handling of nonresponse error. *Human Resource*
640 *Development Quarterly*, 14(1), 99–110.
- 641 Enders, C. K. (2011). Missing not at random models for latent growth curve analyses.
642 *Psychological Methods*, 16(1), 1–16.
- 643 Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic
644 review. *Computers in Human Behavior*.
- 645 Frohlich, M. T. (2002). Techniques for improving response rates in OM survey research.
646 *Journal of Operations Management*, 20(1), 53–62.
- 647 Fulton, B. R. (2016). Organizations and survey research: Implementing response enhancing
648 strategies and conducting nonresponse analyses. *Sociological Methods & Research*,
649 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 rates:
653 Trends and a validity assessment framework. *Human Relations*, 75(8), 1560–1584.
- 654 Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the impact of
655 growing nonresponse on estimates from a national RDD telephone survey. *International*
656 *Journal of Public Opinion Quarterly*, 70(5), 759–779.
- 657 Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa, S.,
658 Merikangas, K. R., Pennell, B.-E., Sampson, N. A., & Zaslavsky, A. M. (2009). National
659 comorbidity survey replication adolescent supplement (NCS-a): II. Overview and design.
660 *Journal of the American Academy of Child & Adolescent Psychiatry*, 48(4), 380–385.

- 661 Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50(1), 537–567.
- 662 Kulas, J. T., Robinson, D. H., Kellar, D. Z., & Smith, J. A. (2017). Nonresponse in
663 organizational surveying: Attitudinal distribution form and conditional response
664 probabilities' impact on patterns of bias. *Public Opinion Quarterly*, 81(2), 401–421.
- 665 Kulas, J. T., Robinson, D. H., Smith, J. A., & Kellar, D. Z. (2016). Post-stratification
666 weighting in organizational surveys: A cross-disciplinary tutorial. *Human Resource
667 Management*.
- 668 Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions
669 between organizational, mechanical turk, and other convenience samples. *Industrial and
670 Organizational Psychology*, 8(2), 142–164.
- 671 Luong, A., & Rogelberg, S. G. (1998). How to increase your survey response rate. *The
672 Industrial-Organizational Psychologist*, 36(1), 61–65.
- 673 Mellahi, K., & Harris, L. C. (2016). Response rates in business and management research:
674 An overview of current practice and suggestions for future direction. *British Journal of
675 Management*, 27(2), 426–437.
- 676 Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data
677 that are not missing completely at random. *Psychometrika*, 52(3), 431–462.
- 678 Pasek, J. (2018). *Anesrake: ANES raking implementation*.
679 <https://CRAN.R-project.org/package=anesrake>
- 680 Pedersen, M. J., & Nielsen, C. V. ek. (2016). Improving survey response rates in online
681 panels: Effects of low-cost incentives and cost-free text appeal interventions. *Social
682 Science Computer Review*, 34(2), 229–243.
- 683 Quine, S., & Morrell, S. (2008). Feeling safe in one's neighbourhood: Variation by location
684 among older australians. *The Australian Journal of Rural Health*, 16, 115–116.
- 685 Rivers, D., & Bailey, D. (2009). Inference from matched samples in the 2008 US national
686 elections. *Proceedings of the Joint Statistical Meetings*, 1, 627–639.
- 687 Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S., & Knight, W.

- 688 E. (2003). Profiling active and passive nonrespondents to an organizational survey.
- 689 *Journal of Applied Psychology*, 88(6), 1104.
- 690 Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee attitude
- 691 surveys: Examining the attitudes of noncompliant employees. *Journal of Applied*
- 692 *Psychology*, 85(2), 284.
- 693 Rogelberg, S. G., & Stanton, J. M. (2007). *Introduction: Understanding and dealing with*
- 694 *organizational survey nonresponse*. Sage Publications Sage CA: Los Angeles, CA.
- 695 Spitzmüller, C., Glenn, D. M., Sutton, M. M., Barr, C. D., & Rogelberg, S. G. (2007).
- 696 Survey nonrespondents as bad soldiers: Examining the relationship between
- 697 organizational citizenship and survey response behavior. *International Journal of*
- 698 *Selection and Assessment*, 15(4), 449–459.
- 699 Taris, T. W., & Schreurs, P. J. (2007). How may nonresponse affect findings in
- 700 organizational surveys? The tendency-to-the-positive effect. *International Journal of*
- 701 *Stress Management*, 14(3), 249–259.
- 702 Tett, R., Brown, C., & Walser, B. (2014). The 2011 SIOP graduate program benchmarking
- 703 survey part 7: Theses, dissertations, and performance expectations. *The*
- 704 *Industrial-Organizational Psychologist*, 51(4), 62–73.
- 705 Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for election
- 706 forecasting? An evaluation of the columbus dispatch poll. *Public Opinion Quarterly*,
- 707 60(2), 181–227.
- 708 Wainer, H. (1976). Estimating coefficients in linear models: It don't make no nevermind.
- 709 *Psychological Bulletin*, 83(2), 213.
- 710 Werner, S., Praxedes, M., & Kim, H.-G. (2007). The reporting of nonresponse analyses in
- 711 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		

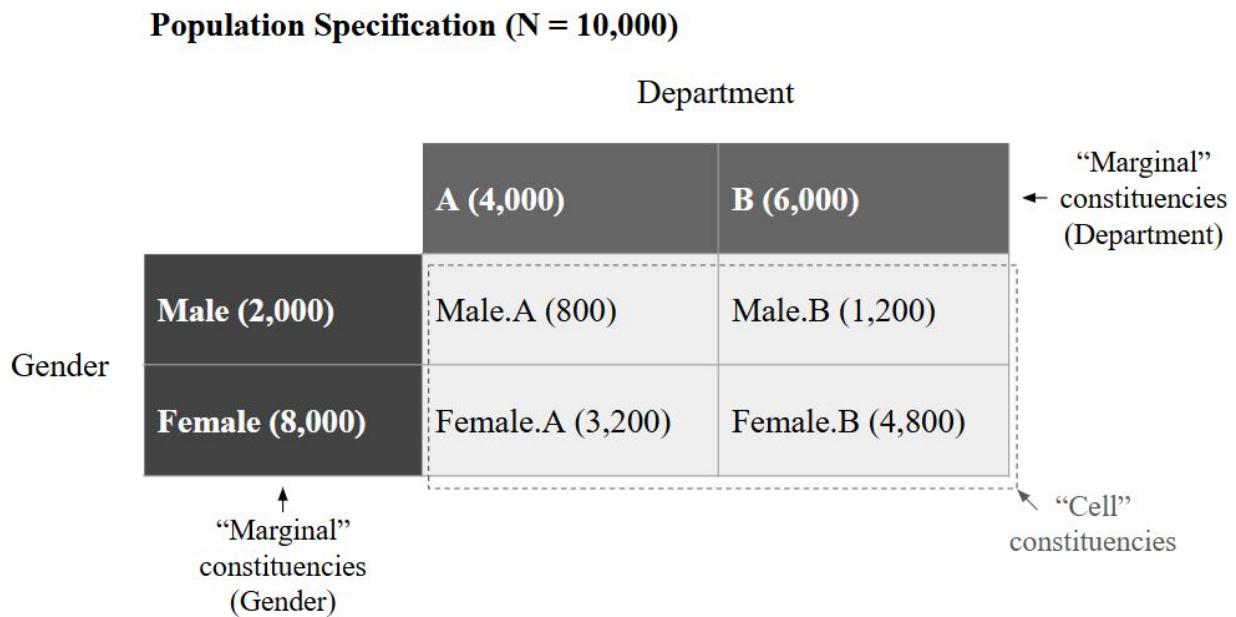
Table 2 continued

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

**Figure 1**

Visual demonstrating terms used to describe population elements.

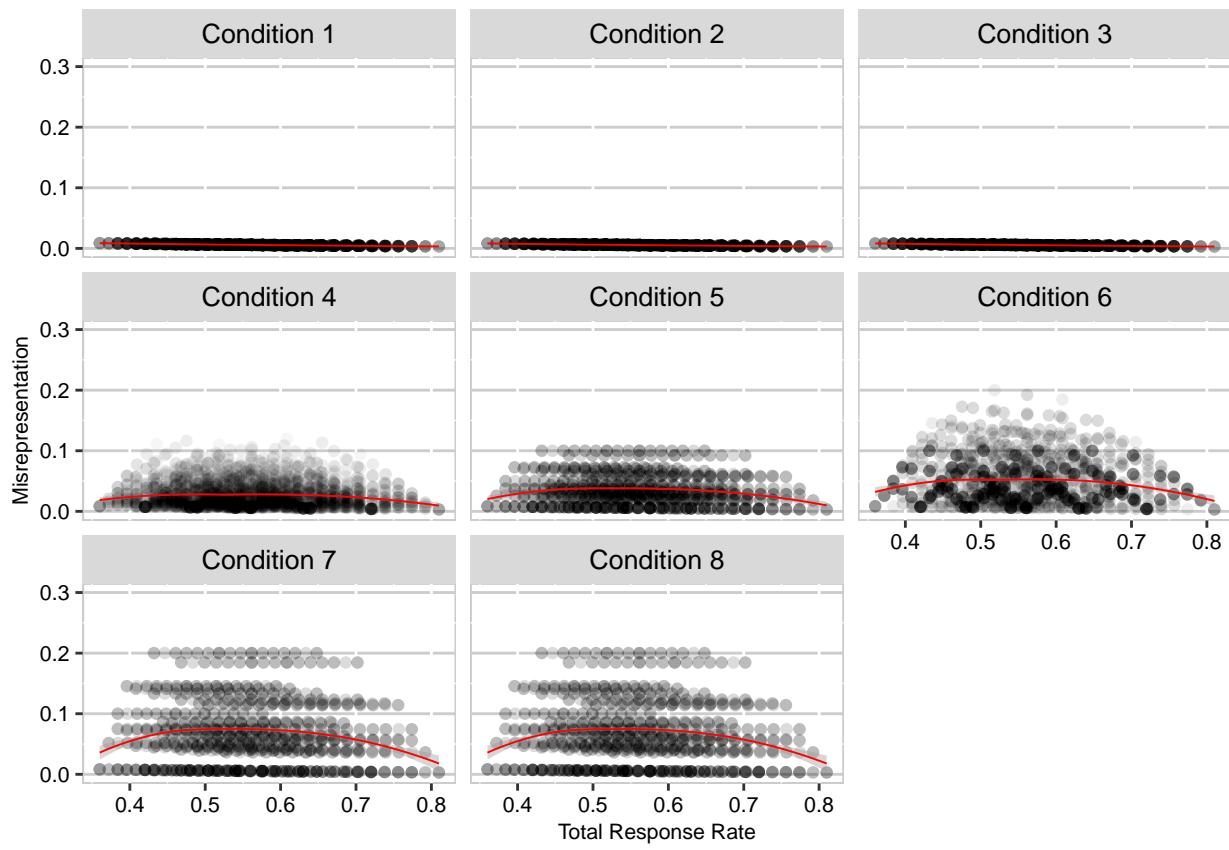


Figure 2

Relationship between total response rate and misrepresentation.

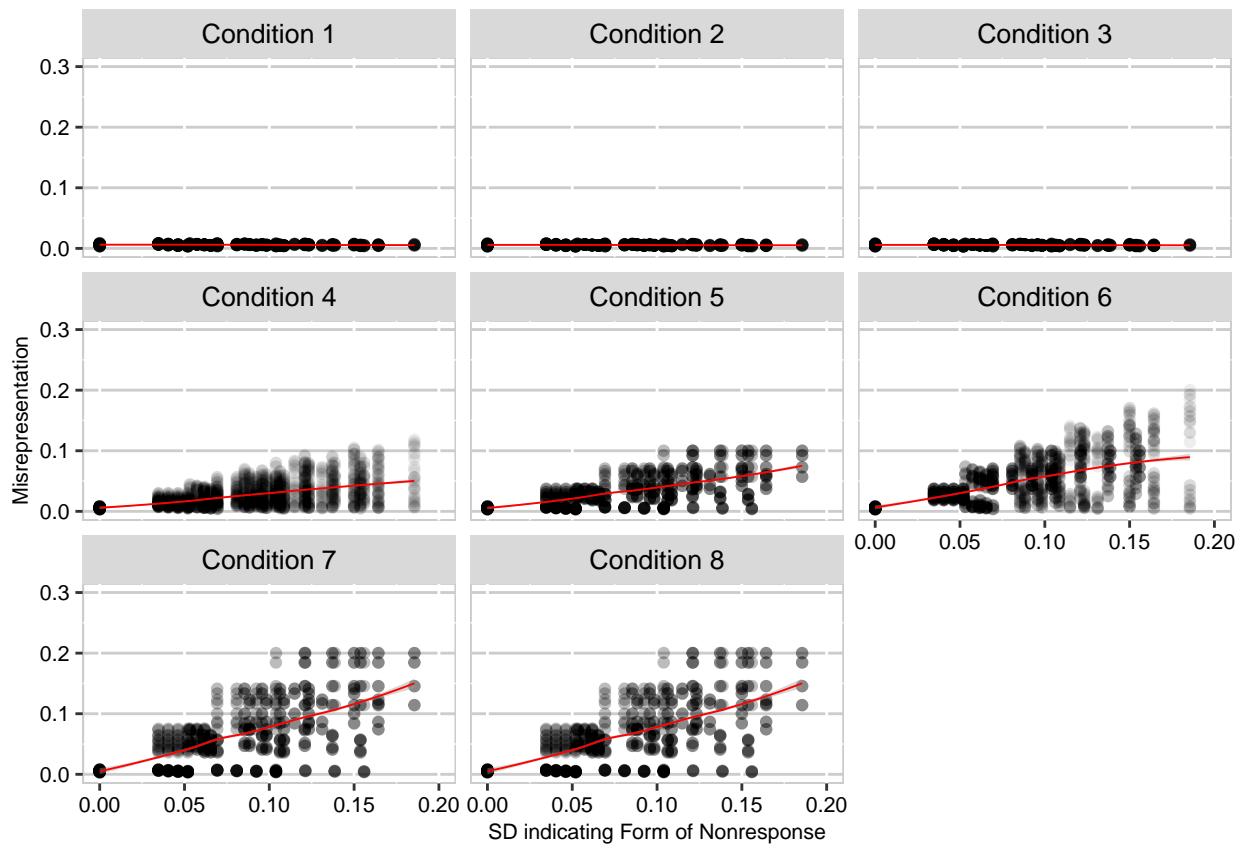


Figure 3

Relationship between nonresponse form and misrepresentation.

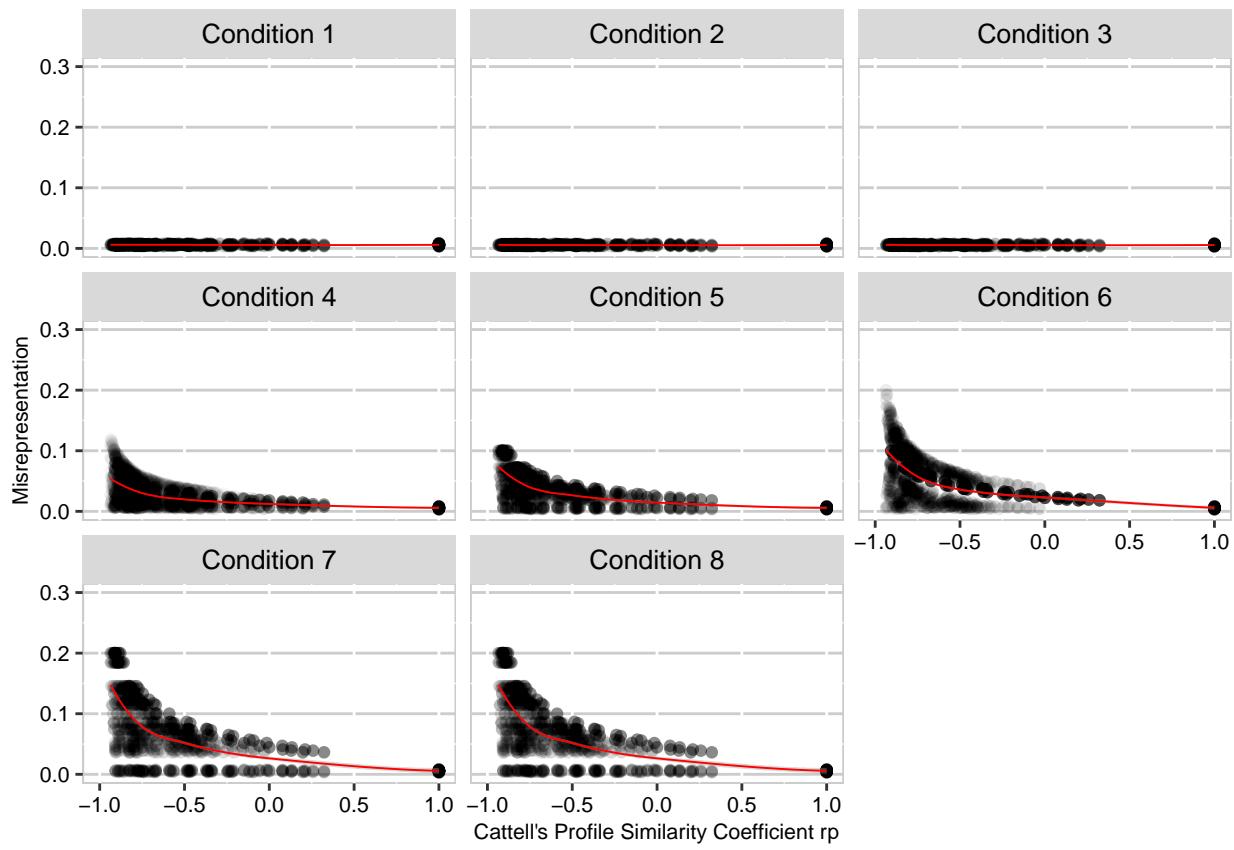


Figure 4

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

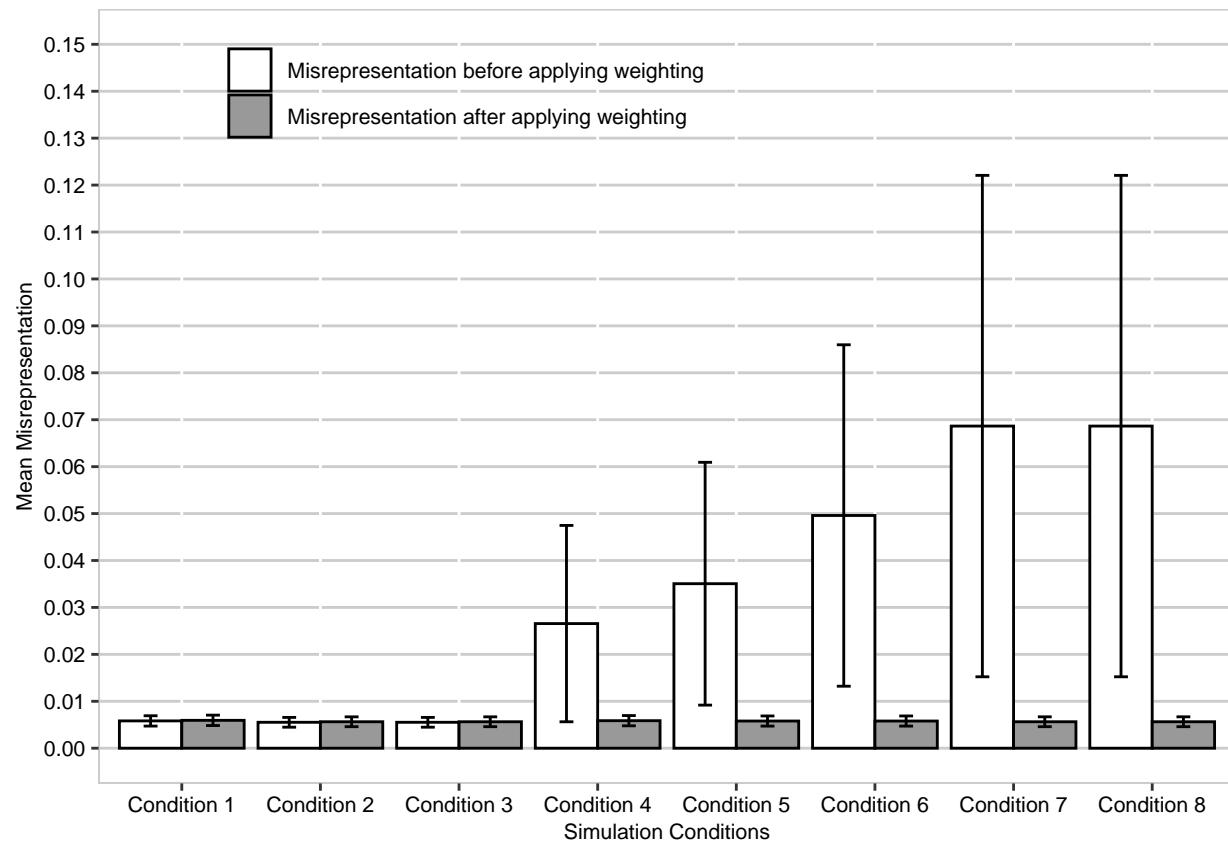


Figure 5

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

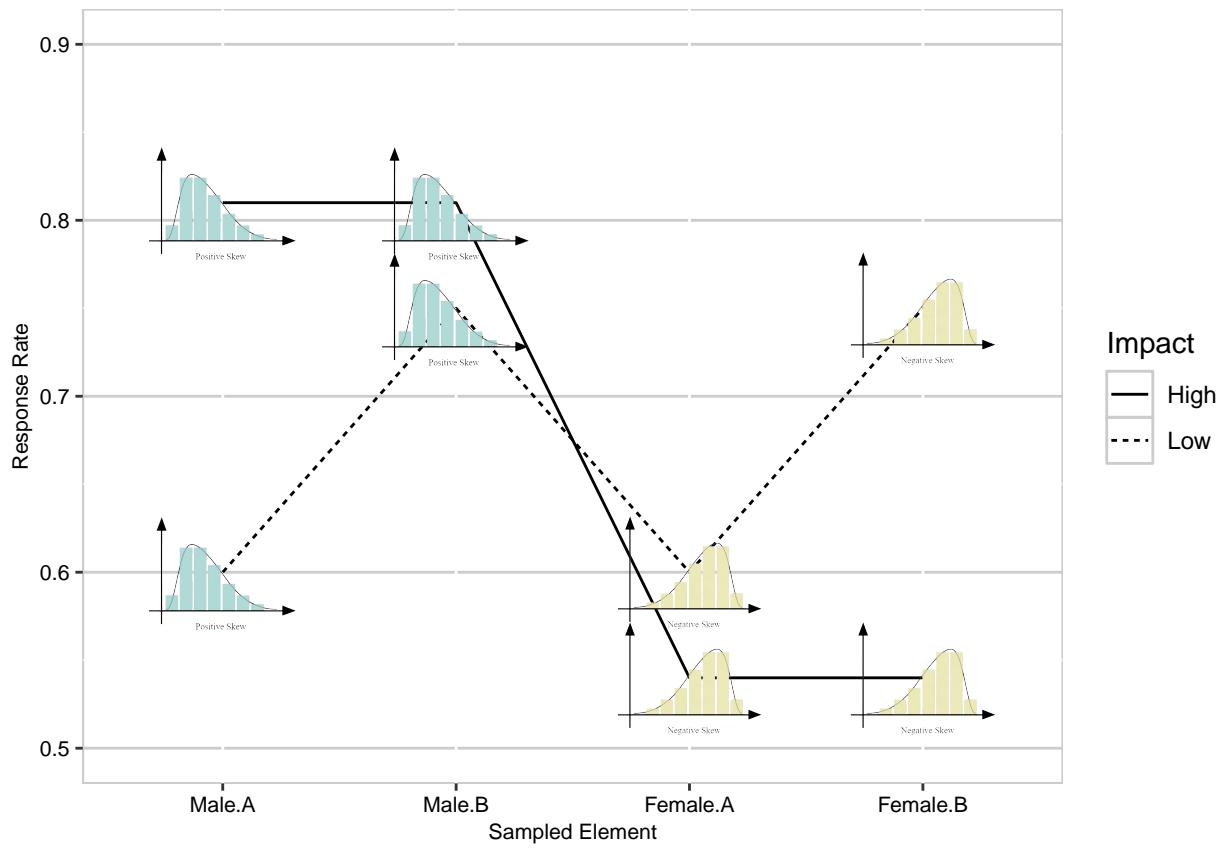


Figure 6

Allocation of response rates relative to underlying distributional form and its impact on population misrepresentation (need to think through hi/lo given Dr Robinsons thoughts)