

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 sparsely 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 (e.g., 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* and *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 (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

Nonresponse and Sample Weighting in Organizational Surveying

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

We speculate that this form of statistical remediation is gaining research interest in the organizational surveying research domain, at least in part, because industrial psychologists are keenly aware that response rates within organizational surveying applications have been trending downward (see, for example, Anseel et al., 2010; Rogelberg & Stanton, 2007). With lower response rates, surveyors are confronted with heightened levels of scrutiny because, historically, a locally realized high response rate has been widely interpreted as a positive indicator of data quality - if not from the survey specialists themselves, at least from client stakeholders (e.g., Anseel et al., 2010; Cycyota & Harrison, 2002, 2006; Frohlich, 2002).

The orientation of this presentation, however, is that although response rate is a commonly referenced proxy of survey quality, it is not response rate but rather sample *representativeness* that should be the primary focus of concern for survey specialists (see, for example, Cook et al., 2000; Krosnick, 1999). Representativeness can of course be “hurt” by low response rates, but the relationship between these two survey concepts is by no

means exact (e.g., Curtin et al., 2000; Keeter et al., 2006; Kulas et al., 2017). Stated differently, a high response rate is neither a sufficient nor even necessary condition for accurate population sampling.¹

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

¹ There are indisputable benefits *associated* with higher response rates, such as greater statistical *power*. This benefit, however, should not be *attributed to* 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 introduce a *false sense* of methodological superiority when the sample misrepresents the population. Primarily for this reason, we stress that the methodological/statistical concepts of response rate, sample size, and power need to 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 sample/population misrepresentation. The current focus is on deviations from a perfect sampling

70 Nonresponse in Organizational Surveying

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

86 To the likely frustration of those who associate response rate with survey data
87 quality, organizational survey response rates have, on average, been declining for decades.
88 Baruch (1999), for example, summarized response rates of 175 studies published in five
89 leading management and behavioral sciences journals in 1975, 1985, and 1995. His results
90 revealed an average response rate (across time periods) of 55.6% ($SD = 19.7\%$), but also a
91 trend within which response rates declined steadily from 64.4% to 55.7% to 48.4% over the
92 three time points. Nine years later, Baruch and Holtom (2008) conducted a follow-up
93 study of 1,607 studies published in 17 disciplinary-relevant journals in 2000 and 2005 but

methodology as opposed to deviations from an ideal psychometric methodology. We do however note that future advancement of current representations of survey error would benefit from a unified perspective that encompasses error arising from both methodological sources: measurement and sampling strategy.

94 found no substantial differences in response rates compared to those in 1995, suggesting
95 that the declining trend had perhaps reached a lower asymptote. However, a different
96 approach with similar goals (Anseel et al., 2010) analyzed 2,037 survey projects published
97 in 12 journals in Industrial and Organizational Psychology, Management, and Marketing
98 from 1995 to 2008 and did note a slight decline (overall $M = 52.3\%$) when controlling for
99 the use of response enhancing techniques.³

100 **Form of Nonresponse**

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

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

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

116 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 Additionally, it has been noted that selection of an individual population element into a
119 realized sample is often predictable (because of, for example, an increased likelihood of not
120 responding when dissatisfied or disgruntled, Taris & Schreurs, 2007). The organizational
121 surveying expectation is that, *on average*, roughly 15% of nonrespondents should be
122 expected to be accurately characterized as “active” (Rogelberg et al., 2003; Rogelberg &
123 Stanton, 2007; Werner et al., 2007). It is this second, less frequently anticipated form of
124 nonresponse that also carries the greater corresponding threat of biased sample estimates
125 (see, for example, Kulas et al., 2017; Rogelberg & Stanton, 2007).

126 **Sample Weighting - a Brief Overview**

127 Within public opinion polling contexts, when realized sample constituencies (e.g.,
128 44% male - by tradition from *judiciously-constructed* and *randomly sampled* data frames)⁴
129 are compared against census estimates of population parameters (e.g., 49% male), weights
130 are applied to the realized sample in an effort to remediate the relative proportional under-
131 or over-sampling. This is because, if the broader populations from which the under- or
132 over-represented groups are sampled differ along surveyed dimensions (e.g., males, within
133 the population, are *less likely to vote for Candidate X* than are women), then unweighted
134 aggregate statistics (of, for example, projected voting results) will misrepresent the true

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 full methodological consideration of sampling context, but is dependent on accurate “census” population constituency estimates (and, as the results highlight, the presence of an active nonrespondent group). Although beyond the scope of the current project, an acknowledgement of the broader methodological sampling context, and the additional potential sources of error, facilitates a deeper appreciation and understanding of the benefits and potential pitfalls of sample weighting.

135 population parameter. This remedial application of sample weights should also be
 136 considered an option for organizational researchers pursuing answers to similar survey
 137 questions such as: “What is the mood of the employees?” This is because focused queries
 138 such as this are (perhaps somewhat covertly) layered - implicit in the question is a focus
 139 not on survey results, but rather the broader employee population. Acknowledging this
 140 implicit target group is of course important, because the next step (after gauging the mood
 141 of the surveyed respondents) is *doing something* about it. Weighting should be considered
 142 a procedural option for organizational surveyors to potentially transition a bit closer from,
 143 “What do the survey results say”? to “What do the employees feel”?

144 **Procedural application**

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

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

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

- 155 1) Determine proportional weights for all levels within one stratum, and then assign
 156 these weights to cases.
- 157 2) Determine proportional weights for a second group (ratio of population percent to

158 *current* sample percent [the current sample percent will be affected by the step 1
159 weighting procedure]). Multiply previous (step 1) weights by the proportional
160 weights for this second stratum and assign these new weights to cases.

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

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

179 *Research question 1:* What role does overall response *rate* play in
180 sample/population misrepresentation? **[make sure this is reflected in results]**

181 *Research question 2:* What role does nonresponse *form* (passive versus active) play
182 in sample/population misrepresentation? **currently in paper as figures 1-3**

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

184 biased (e.g., misrepresentative) and unbiased sample estimates?

185 *Research question 4:* What is the role of response rate, form, and underlying

186 population attitudes in the *effectiveness* of weighting? [perhaps David can derive/find

187 **a proof to parallel our results?**]

188 Added population attitudes (1/20/23) - not sure if this clutters but more

189 consistent with flow of introduction

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

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

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

193 variable weighting in a selection context (e.g., Wainer, 1976) or simple composite score

194 aggregate (Bobko et al., 2007), we propose that a similar consideration is appropriate with

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

196 weighting via carefully controlled data simulation.

197 Methods

198 We address our research questions within the context of organizational surveying

199 (commonly assessing estimates of, for example, commitment, culture/climate, engagement,

200 or satisfaction). We began the simulations by establishing “populations”, each consisting of

201 10,000 respondents characterized by demographic categorizations across gender (male and

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

203 male-B, female-A, and female-B). For these population respondents, we generated scaled

204 continuous responses (real numbers) ranging from values of 1 to 5, reflecting averaged

205 aggregate scale scores from a fictional multi-item survey with a typical $1 \rightarrow 5$ Likert-type

206 or graphic rating scale response format.

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

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

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

210 80%) with increments and corresponding decrements of 20%. For example, if males

211 accounted for 20% of the simulated population, then females were 80%; also if respondents

212 in Department A represented 60% of a population, then 40% were in Department B.

213 Marginal constituencies were therefore specified at all combinations (across the two

214 variables) of 20% and 80%, 40% and 60%, 60% and 40%, and 80% and 20%. This resulted

215 in population *cell* constituencies (e.g., men in department A) as low as 400 and as high as

216 6,400.

217 Additionally, each of these cell populations was characterized by an attitude

218 distribution in one of three different possible forms: normal, positively skewed, or

219 negatively skewed. These distributional forms were specified in an attempt to model

220 similarities and discrepancies in construct standing (e.g., commitment, satisfaction, or

221 engagement) across respondent groupings. The normal distribution exhibited, on average,

222 a mean of 3.0 whereas the skewed distributions were characterized by average means of 2.0

223 and 4.0, respectively. In total, eight crossings of distributional type across employee

224 categorization were specified (Table 1 presents the combinations of these distributions).

225 Note that these eight conditions are not exhaustive of all possible combinations - we

226 specified scenarios that we expected to be most efficiently informative across our passive to

227 active nonresponse continuum (reflected in Table 1's "anticipated bias" column).

228 Individual attitudes were randomly sampled from population distributions at the

229 cell level (e.g., Department A Males) without replacement. Response rates

230 (methodologically these could also be conceptualized as *sampling* rates) were controlled at

231 the marginal level using 10% increments ranging from 60% to 90%, and these were fully

232 iterated. Our cell-level response rates therefore ranged from 36% to 81% - a range of rates

233 chosen because they are, according to the organizational surveying literature, reasonable

expectations (e.g., Mellahi & Harris, 2016; Werner et al., 2007). We therefore investigated error within the aggregate mean (e.g., grand mean or total sample mean) attributable to different likelihoods of sample inclusion from constituent groups of different relative size and representing populations of different attitudinal distribution, but at response rates reasonably expected to exist in real-world organizational surveying contexts.

It should be noted here that there are several collective patterns of response that are intended to represent sampling scenarios exhibiting *passive* nonresponse, regardless of absolute response rate: all subgroups exhibiting the same response rate (e.g., 36%, 36%, 36%, and 36%). All other combinations of response rate are intended operationalizations of active forms of nonresponse (e.g., *not* reasonably characterized as missing at random, NMAR), although the degree to which a sampling scenario should be reasonably characterized as exhibiting active nonresponse is intended to be incremental across iterations.

In an attempt to capture this “degree of active nonresponse”, we calculated a simple index of response rate discrepancy (SD; presented in Table 2). The “least” active nonresponse scenarios are characterized by two subgroups with identical response rates and two having a slightly different response rate (e.g., Dept A Males = 36%, Dept A Females = 36%, Dept B Males = 42%, and Dept B Females = 42%; see the second row of Table 2, the SD index = .034)⁵. Also here note that three of our eight Table 1 conditions represent scenarios where the presence of active nonrespondents is not expected to result in bias (e.g., regardless of patterns of nonresponse, the unweighted sample mean is expected to

⁵ This method of simplifying the presentation of our response rate conditions is fully orthogonal to 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.

255 yield an unbiased estimate of the population mean). These are Table 1 conditions one
256 through three, where attitudinal distributions are of *the same form* across groups,
257 regardless of any individual group response rate discrepancy from others'.

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

270 **Results**

271 *Research question 1:* What role does overall response *rate* play in sample
272 misrepresentation? **[make sure this is reflected in results]**

273 A couple paragraphs to answer RQ1

274 Have to operationalize "sample misrepresentation" first

275 The following is RQ2:

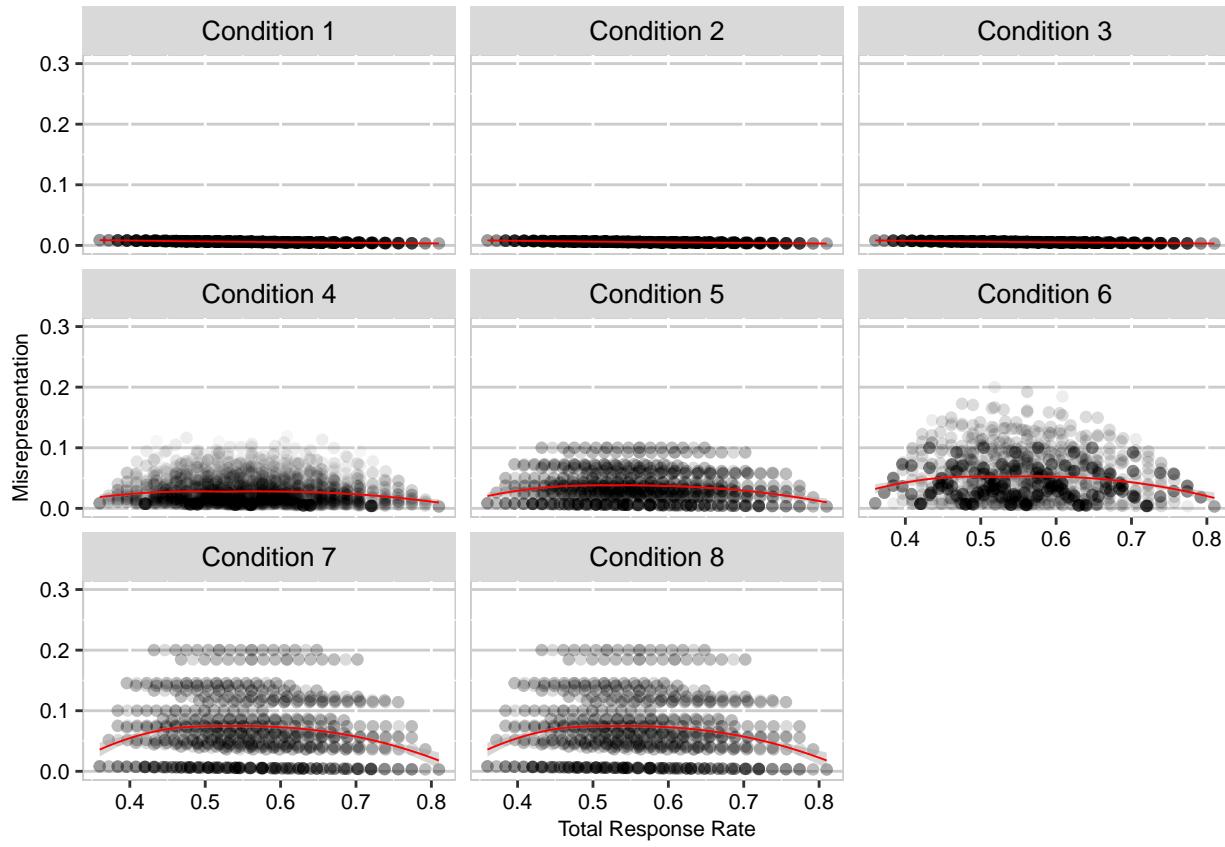
276 In total, we generated 327.68 million samples (4,096 unique combinations of
277 response rate and population constituency across gender and department, simulated 10,000
278 times each across our eight Table 1 conditions). Each of these samples was comprised of,
279 on average, $n = 5,625$, collectively representing an experiment-wide n of 1.8432 trillion.

280 For each individual simulation, weights were applied iteratively to the data at the two
281 marginal (variable) levels via raking, and were estimated via the *anesrake* package (Pasek,
282 2016) in R version 4.2.2 (2022-10-31 ucrt).

283 We were most interested in comparing the extent to which unweighted (aggregated
284 responses without raking) and weighted (aggregated weighted responses) sample means
285 approximated the population means across our controlled specifications of response rate,
286 nonresponse form, and attitudinal distribution (population means were taken from each
287 iteration, as the simulations specified a new population at each iteration). The
288 “effectiveness” of weighting was evaluated by calculating the discrepancies between the
289 population and both weighted and unweighted sample means as well as the averaged
290 deviations of these discrepancies from the population mean (discrepancy in the “mean” of
291 the means is bias, dispersion about the “mean” of the means is error). If the average
292 weighted sample mean was closer to the true population mean, relative to the unweighted
293 one, then the weighting was deemed beneficial.

294 Add a couple of paragraphs here to answer research questions 1(a) and 1(b)

295 Correlation coefficient needed.[Yang to calculate 2/1]



296

297 To partially address the second limitation, discrepancy between population
 298 constituency and sampling proportions was additionally estimated via Cattell's profile
 299 similarity index [r_p ; Cattell et al. (1966)]. r_p is sensitive to discrepancies in profile shape
 300 (pattern across profile components), elevation (average component score), and scatter (sum
 301 of individual components' deviation from the elevation estimate. Figure 3 demonstrates the
 302 pattern of unweighted sample mean deviation (from the population parameter) when this
 303 index is taken into consideration. edits... gain demonstrate these relationships across the
 304 attitudinal form conditions, being grouped by underlying distributions thought to be
 305 susceptible to bias (Conditions 3 through 8) as well as those thought to be relatively
 306 immune to bias (Conditions 1 through 3; aka those sampling situations in which weighting
 307 is unnecessary).

308 ## Warning: The 'size' argument of 'element_rect()' is deprecated as of ggplot2 3.4.0.

309 ## i Please use the ‘linewidth’ argument instead.

310 The plurality of our findings are presented visually, and they focus on the overall
311 mean (e.g., the average rating across all sample members). Figure 1 provides a broad
312 summary of the results across the eight different attitudinal distribution conditions,
313 presenting the average absolute discrepancy from the population mean within each broad
314 condition. Conditions one through three demonstrate that, on average, the unweighted
315 sample mean provides a good (unbiased) estimate of the population mean when the
316 distributional form is held constant across constituent groups (e.g., the distributions of
317 attitudes are of similar functional forms and locations for all constituent groups). This is
318 regardless of form or extent of nonresponse. Additionally, weighting remediates deviations
319 about the true mean in all five attitudinally discrepant conditions, even when considerable
320 error exists in the unweighted estimate (e.g., the rightmost bars in Figure 1).

321 **The Role of Response Rate**

322 In terms of explaining the very little error that did emerge within the passive
323 nonresponse conditions, this error was entirely attributable to response rate (See Figure 2).
324 The nature of the exact relationship was slightly nonlinear, being fit with quadratic
325 functions within each condition (collapsing across conditions did exhibit slight within-array
326 differences [which would affect the statistically perfect relationship]).

327 **Need to Recall Research Questions in appropriate sections**

328 Figure 3 demonstrates how the weighting algorithm operated across conditions one
329 through three taking form of nonresponse into consideration (along the x-axis, with passive
330 nonresponse occupying the left of the figure and active nonresponse scenarios occupying
331 the right). There is a very slight amount of error in the unweighted sample mean with
332 active nonresponse, as well as a systematic pattern of heteroskedasticity across the “passive
333 to active” continuum (studentized Breusch-Pagan = 565.42 [unweighted], 496.67
334 [weighted], p 's < .001). Weighting always corrects this slight amount of error. Figure 3

335 demonstrates a more pronounced *form of* nonresponse association when underlying
336 attitudinal distributions evidence group differences, and in these scenarios, active
337 nonresponse is shown to have a fairly large effect on error within the sample estimate (and,
338 again, predictable heteroskedasticity paralleling the SD index, Breusch-Pagan = 3177.2
339 [unweighted]; 832.91 [weighted], p 's < .001). Weighting again corrects the sample estimate.

340 It should be noted regarding the above-mentioned “heteroskedasticity” that there
341 are active nonresponse scenarios in which no error is found (see, for example, the lower
342 right-hand portion of Figure 3 where values appear all along the passive-active abscissa).
343 These situations are ones within which the response rates “parallel” the distributional
344 form. For example, in Condition Eight, the distributional forms were: Positive Skew_{Male_A},
345 Positive Skew_{Male_B}, Negative Skew_{Female_A}, Negative Skew_{Female_B}. In the most extreme
346 cases of active nonresponse, response rates that fully parallel distributional patterns (e.g.,
347 20%_{Male_A}, 20%_{Male_B}, 80%_{Female_A}, 80%_{Female_B}) result in no error in the population mean
348 approximation (average discrepancy = .0003, $SD = .0002$). Alternatively, when the
349 response rates are inverted, (e.g., 20%_{Male_A}, 80%_{Male_B}, 20%_{Female_A}, 80%_{Female_B}), there
350 is substantial error in approximation (average discrepancy = .51, $SD = .14$). **this is an**
351 **old number - why are our new numbers so low? (see, for example, the y-axis**
352 **on Figure 1) - YANG? (11/17/18)** Again, it is not merely response rate or form that
353 is associated with biased sample estimates, but rather the nature of response rate relative
354 to existing attitudinal differences.

355 To further elaborate this point, consider, for example, Condition 4. Here, three
356 groups are characterized by similar distributions of attitudes (normally distributed) and
357 one, Females from Department B, is characterized by negatively skewed attitudes. The
358 greatest unweighted error here arises from sampling scenarios in which there are many
359 Department B females (e.g., in our specifications, 6,400) and fewer males and Department
360 A females⁶, but the Department B females exhibit a much lower response rate (e.g., 20%)

⁶ Because of the “marginal” versus “cell” specifications of population constituencies, our most extreme

361 than do other groups, who respond at a high rate (e.g., 80%). That is, it is not merely
362 response rate, but response rate within these identifiable groups, and whether or not those
363 response rate differences parallel underlying attitudinal differences.

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

367 To partially address the second limitation, discrepancy between population
368 constituency and sampling proportions was additionally estimated via Cattell's profile
369 similarity index [r_p ; Cattell (1949); Cattell et al. (1966)]. r_p is sensitive to discrepancies in
370 profile shape (pattern across profile components), elevation (average component score), and
371 scatter (sum of individual components' deviation from the elevation estimate). Figure 3
372 demonstrates the pattern of unweighted sample mean deviation (from the population
373 parameter) when this index is taken into consideration. edits...gain demonstrate these
374 relationships across the attitudinal form conditions, being grouped by underlying
375 distributions thought to be susceptible to bias (Conditions 3 through 8) as well as those
376 thought to be relatively immune to bias (Conditions 1 through 3; aka those sampling
377 situations in which weighting is unnecessary).

378 **Summary**

379 Collectively the results highlight three aspects of weighting: 1) our simulations are
380 comprehensive, iterating through all possible combinations of response rates - those
381 paralleling population distributions, those inversely mirroring population distributions, and
382 those "orthogonal to" population distributions, 2) the "SD" operationalization of passive to
383 active forms of nonresponse is a bit crude and insensitive to specific combinations of

example here is necessarily 400 Department A males, 1,600 Department B males, and 1,600 Department A females. This was a decision based on keeping the population N's at 10,000 and certainly more extreme population constituency combinations could be examined in future like-minded explorations.

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

388 Mean square error is our second index for sample quality. It is a well-known
389 mathematical theorem that the application of weights increases (random) errors of
390 precision, which was also empirically true in the current study. For each condition in our
391 simulations, we calculated the standard deviations of 40.96 million unweighted and 40.96
392 million weighted samples means (4,096 possible population-sample combinations by 10,000
393 iterations), which yielded eight empirically-estimated standard errors of unweighted and
394 weighted sample means. Figure XXX <- need to readd this visually presents these
395 standard errors in eight pairs of bars, demonstrating that the standard error of weighted
396 sample means (red bar) tended to be 16% to 18% larger than that of unweighted sample
397 means (grey bar) regardless of condition. These errors highlight the caveat that weighting
398 should only be applied in the active nonresponse case (e.g., although the aggregate effect of
399 weighting with passive nonresponse is error-minimizing, any one sampling condition is
400 *more likely* to result in greater deviation from the population parameter when weighting is
401 applied the passive nonresponse data).

402 In summary, as an aggregate across sampling events, weighting always corrects
403 sample bias, when it is present in the unweighted estimate. However, the standard errors
404 suggest that for any *one* sampling event in the absence of bias, the likelihood that the
405 sample mean approximates the *mean* of sample means is (slightly) greater for the
406 unweighted estimate. When bias is present, however, (in the unweighted estimate) there is
407 obviously no advantage to “being closer” to this biased mean of means. That is, under
408 some circumstances, the mean of unweighted sample means does not center on the
409 population mean. The implications of this seem quite obvious: Weighting should only be
410 applied if bias is anticipated in the sample estimate. This may seem to be a picayune

411 recommendation, but we note here that this advocacy is not heeded in public opinion
412 polling applications, where the computation and application of weights are default
413 procedures (CITES? - perhaps AAPOR standards or personal communication with polling
414 agencies such as Gallop).

415 **Can we look at the “crossing point?” (e.g., when MSE becomes
416 excessive) - David?**

417 **Discussion**

418 We view nonresponse as a serious problem that should be addressed via repeated
419 attempts to survey particularly reluctant or hard-to-reach respondents particularly because
420 nonresponse may be reasonably expected to be greatest in groups that are most unsatisfied
421 [e.g., it may be typical for individuals representing these groups to have their responses
422 diluted; see, for example, Taris and Schreurs (2007)]. However, several researchers have
423 noted potentially misplaced relative emphasis on survey response rates, with Cook et al.
424 (2000), Krosnick (1999), and Visser et al. (1996) articulating the point that
425 representativeness of the sample is more important than response rate. We also believe
426 that the goal in organizational surveying should be representativeness not exhaustiveness.
427 Krosnick (1999) specifically comments that, even when probability sampling is employed,
428 response rate does not necessarily implicate either good or poor sample representativeness.
429 One aim of this paper is to reinforce this primary ‘representativeness’ orientation to those
430 who may be otherwise inclined to focus on response rate as a sufficient index of quality
431 (and propose sample weighting as a practice that can adjust for lack of representativeness).

432 With the above in mind, we set out to answer two fairly simple questions: What
433 impact does the application of weights have on the quality of sample estimates, and what
434 role does nonresponse play? Our answers are that: 1) weighting “always” helps, as long as
435 you capture the proper strata (which of course we were able to do via controlled
436 simulation), but also 2) response rate impact *depends* on relationship between response

437 rate and the underlying distribution of attitudes. conditions 1 through 3 as well as all
438 other conditions are occasionally immune to response rate influence, depending on whether
439 the pattern of nonresponse parallels the pattern of attitudinal distribution differences or
440 not). Active forms of nonresponse can harm the unweighted sample estimate, but only
441 when the pattern of active nonresponse is accompanied by differing distributions of
442 attitudes within the active nonrespondent “populations” [this would appear to be a
443 reasonable expectation based on the literature; e.g., Rogelberg et al. (2000); Rogelberg et
444 al. (2003); Spitzmüller et al. (2007)]. Although the weighted mean proved an unbiased
445 estimate of the population mean across all simulations, in circumstances where no bias
446 existed in the unweighted estimate, the trade-off between bias-correction and random error
447 of precision (e.g., standard error) also needs to be acknowledged.

448 It should be noted that the organizational surveying categorization of passive versus
449 active parallels the broader statistical focus on data that is missing at random or
450 completely at random (MAR or MCAR, see for example, Heitjan & Basu, 1996) versus
451 data not missing at random [non-MCAR, see for example,]. Imputation is the common
452 remediation for data MAR or MCAR whereas non-MCAR solutions may involve strategies
453 such as latent variable estimation procedures (Muthén et al., 1987). In the context of
454 surveying, we are similarly proposing a bifurcation of remediation methods - no
455 remediation with passive nonresponse and post-stratification weighting with active.

456 Previous presentations have noted that bias is sometimes associated with
457 nonresponse and other times it is not - this research has not been explicit in the specific
458 conditions that moderate this association, however. The current paper does make this
459 association explicit. It is not merely the form of nonresponse that determines whether or
460 not bias occurs, but also the underlying distributions that the response probabilities are
461 applied to. Some distributional patterns are immune to the biasing effects of active
462 nonresponse (see, for example, Conditions 1 through 3). Some patterns of active
463 nonresponse also result in no bias even when distributional patterns deviate substantially

464 (see, for example, Condition 8 where a 20%, 20%, 80%, 80% response rate pattern exhibits
465 no error). The target therefore should not be merely form of nonresponse but also
466 underlying attitudes. Regardless, however, weighting always remediates the error when it
467 occurs (and does not add error where it is absent).

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

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

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

490 It should also be pointed out that although the active nonrespondent group seems

491 to be a great concern, it will not seriously bias the results unless the proportion of active

492 nonrespondents is higher than 15% (Rogelberg et al., 2003; Rogelberg & Stanton, 2007;

493 Werner et al., 2007). “In this study we found that the active nonrespondent group was

494 relatively small (approximately 15%), but consistent in size with research conducted by.”

495 (Rogelberg et al., 2003, pp. 1110–1111). “Furthermore, consistent with Roth (1994) who

496 stated that when missingness is not random (as we found for active nonrespondents),

497 meaningful bias will only be introduced if the group is relatively large (which was not the

498 case in this study).” (Rogelberg et al., 2003, p. 1112).

499 “If the results show that the active nonrespondent group comprises a low proportion

500 of the population, fewer concerns for bias arise. If the proportion of active respondents is

501 greater than 15% of the group of individuals included in the interviews or focus groups

502 (this has been the average rate in other studies), generalizability may be compromised.”

503 (Rogelberg & Stanton, 2007, p. 201) * I believe there is an error here. The author want to

504 say that if the proportion of active nonrespondents is greater than 15% of the group .

505 “It has been suggested that it takes a response rate of 85% to conclude that

506 nonresponse error is not a threat (Dooely & Lindner, 2003). We agree that researchers

507 should provide both empirical and theoretical evidence refuting nonresponse bias whenever

508 the response rate is less than 85%.” (Werner et al., 2007, p. 293).

509 Note here however, the seeming disconnect between the reports of 15% active

510 nonresponse and declining response rates (trending toward 50%). Certainly with

511 decreasing overall response rates, the likely reasons would appear to be more active than

512 passive (e.g., it is difficult to entertain the idea that potential respondents are more likely

513 to forget to respond today than they were 40 years ago).

514 Integration of IT/IS systems within HR functions hopefully assists the

515 (un)likelihood that organizatinoal population frames are either deficient or

516 contaminated, although we note that this possibility (frame misspecification) is
517 much more plausible within organizations that do not have updated or
518 integrated HR IT/IS systems (perhaps, ironically, *smaller* organizations).

519 **Future Directions**

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

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

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

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

For Condition 5 (for example, low/high response rates with minority/majority

569 population constituencies). The lower-right to upper-left diagonal reflects response rates
570 that parallel population constituencies. The patterns across these stressors were consistent,
571 with the weighted sample means (red dots) providing unbiased estimates of the population
572 parameter, whereas the unweighted sample means (grey dots) tended to yield unbiased
573 estimates when the population constituencies were roughly equal (e.g., close to 50%/50%).

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

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

References

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

- 615 Management, 28(2), 151–176.
- 616 Cycyota, C. S., & Harrison, D. A. (2006). What (not) to expect when surveying
617 executives: A meta-analysis of top manager response rates and techniques over
618 time. *Organizational Research Methods*, 9(2), 133–160.
- 619 Deming, W. E., & Stephan, F. F. (1940). On a least squares adjustment of a
620 sampled frequency table when the expected marginal totals are known. *The
621 Annals of Mathematical Statistics*, 11(4), 427–444.
- 622 Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A
623 systematic review. *Computers in Human Behavior*.
- 624 Frohlich, M. T. (2002). Techniques for improving response rates in OM survey
625 research. *Journal of Operations Management*, 20(1), 53–62.
- 626 Fulton, B. R. (2016). Organizations and survey research: Implementing response
627 enhancing strategies and conducting nonresponse analyses. *Sociological Methods
628 & Research*, 0049124115626169.
- 629 Heitjan, D. F., & Basu, S. (1996). Distinguishing “missing at random” and “missing
630 completely at random.” *The American Statistician*, 50(3), 207–213.
- 631 Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the
632 impact of growing nonresponse on estimates from a national RDD telephone
633 survey. *International Journal of Public Opinion Quarterly*, 70(5), 759–779.
- 634 Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa,
635 S., Merikangas, K. R., Pennell, B.-E., Sampson, N. A., & Zaslavsky, A. M.
636 (2009). National comorbidity survey replication adolescent supplement (NCS-a):
637 II. Overview and design. *Journal of the American Academy of Child &
638 Adolescent Psychiatry*, 48(4), 380–385.
- 639 Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50(1),
640 537–567.
- 641 Kulas, J. T., Robinson, D. H., Kellar, D. Z., & Smith, J. A. (2017). Nonresponse in

- 642 organizational surveying: Attitudinal distribution form and conditional response
643 probabilities' impact on patterns of bias. *Public Opinion Quarterly*, 81(2),
644 401–421.
- 645 Kulas, J. T., Robinson, D. H., Smith, J. A., & Kellar, D. Z. (2016).
646 Post-stratification weighting in organizational surveys: A cross-disciplinary
647 tutorial. *Human Resource Management*.
- 648 Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary
649 distinctions between organizational, mechanical turk, and other convenience
650 samples. *Industrial and Organizational Psychology*, 8(2), 142–164.
- 651 Luong, A., & Rogelberg, S. G. (1998). How to increase your survey response rate.
652 *The Industrial-Organizational Psychologist*, 36(1), 61–65.
- 653 Mellahi, K., & Harris, L. C. (2016). Response rates in business and management
654 research: An overview of current practice and suggestions for future direction.
655 *British Journal of Management*, 27(2), 426–437.
- 656 Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with
657 data that are not missing completely at random. *Psychometrika*, 52(3), 431–462.
- 658 Pasek, J. (2016). *Anesrake: ANES raking implementation*.
- 659 Pedersen, M. J., & Nielsen, C. V. ek. (2016). Improving survey response rates in
660 online panels: Effects of low-cost incentives and cost-free text appeal
661 interventions. *Social Science Computer Review*, 34(2), 229–243.
- 662 Quine, S., & Morrell, S. (2008). Feeling safe in one's neighbourhood: Variation by
663 location among older australians. *The Australian Journal of Rural Health*, 16,
664 115–116.
- 665 Rivers, D., & Bailey, D. (2009). Inference from matched samples in the 2008 US
666 national elections. *Proceedings of the Joint Statistical Meetings*, 1, 627–639.
- 667 Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S., &
668 Knight, W. E. (2003). Profiling active and passive nonrespondents to an

- 669 organizational survey. *Journal of Applied Psychology*, 88(6), 1104.
- 670 Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee
671 attitude surveys: Examining the attitudes of noncompliant employees. *Journal*
672 *of Applied Psychology*, 85(2), 284.
- 673 Rogelberg, S. G., & Stanton, J. M. (2007). *Introduction: Understanding and dealing*
674 *with organizational survey nonresponse*. Sage Publications Sage CA: Los
675 Angeles, CA.
- 676 Spitzmüller, C., Glenn, D. M., Sutton, M. M., Barr, C. D., & Rogelberg, S. G.
677 (2007). Survey nonrespondents as bad soldiers: Examining the relationship
678 between organizational citizenship and survey response behavior. *International*
679 *Journal of Selection and Assessment*, 15(4), 449–459.
- 680 Taris, T. W., & Schreurs, P. J. (2007). How may nonresponse affect findings in
681 organizational surveys? The tendency-to-the-positive effect. *International*
682 *Journal of Stress Management*, 14(3), 249.
- 683 Tett, R., Brown, C., & Walser, B. (2014). The 2011 SIOP graduate program
684 benchmarking survey part 7: Theses, dissertations, and performance
685 expectations. *The Industrial-Organizational Psychologist*, 51(4), 62–73.
- 686 Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for
687 election forecasting? An evaluation of the columbus dispatch poll. *Public*
688 *Opinion Quarterly*, 60(2), 181–227.
- 689 Wainer, H. (1976). Estimating coefficients in linear models: It don't make no
690 nevermind. *Psychological Bulletin*, 83(2), 213.
- 691 Werner, S., Praxedes, M., & Kim, H.-G. (2007). The reporting of nonresponse
692 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		Anticipated Bias
			Dept A	Dept B	Dept A	Dept B	
1	Normal	3	X	X	X	X	None
	Positive Skew	2					
	Negative Skew	4					
2	Normal	3					None
	Positive Skew	2	X	X	X	X	
	Negative Skew	4					
3	Normal	3					None
	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

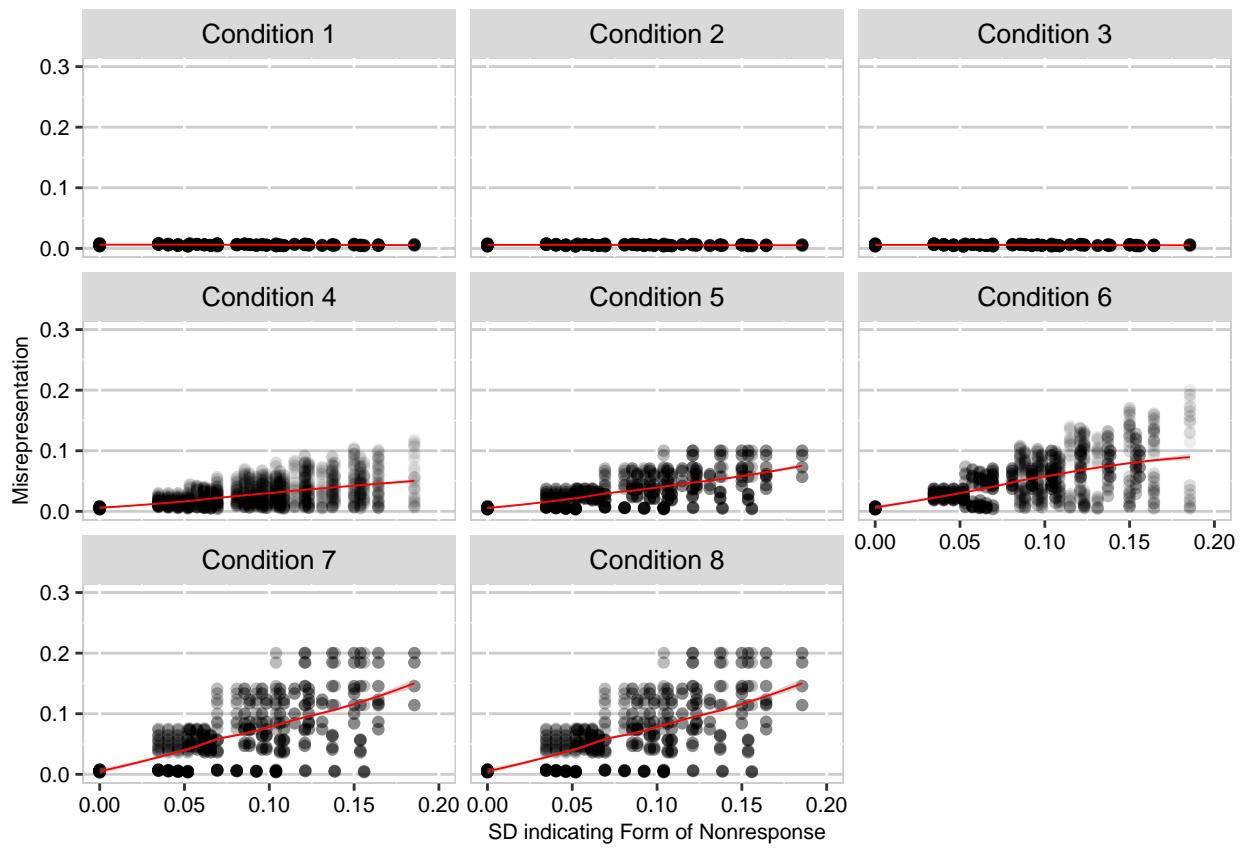


Figure 1

Relationship between nonresponse form and misrepresentation.

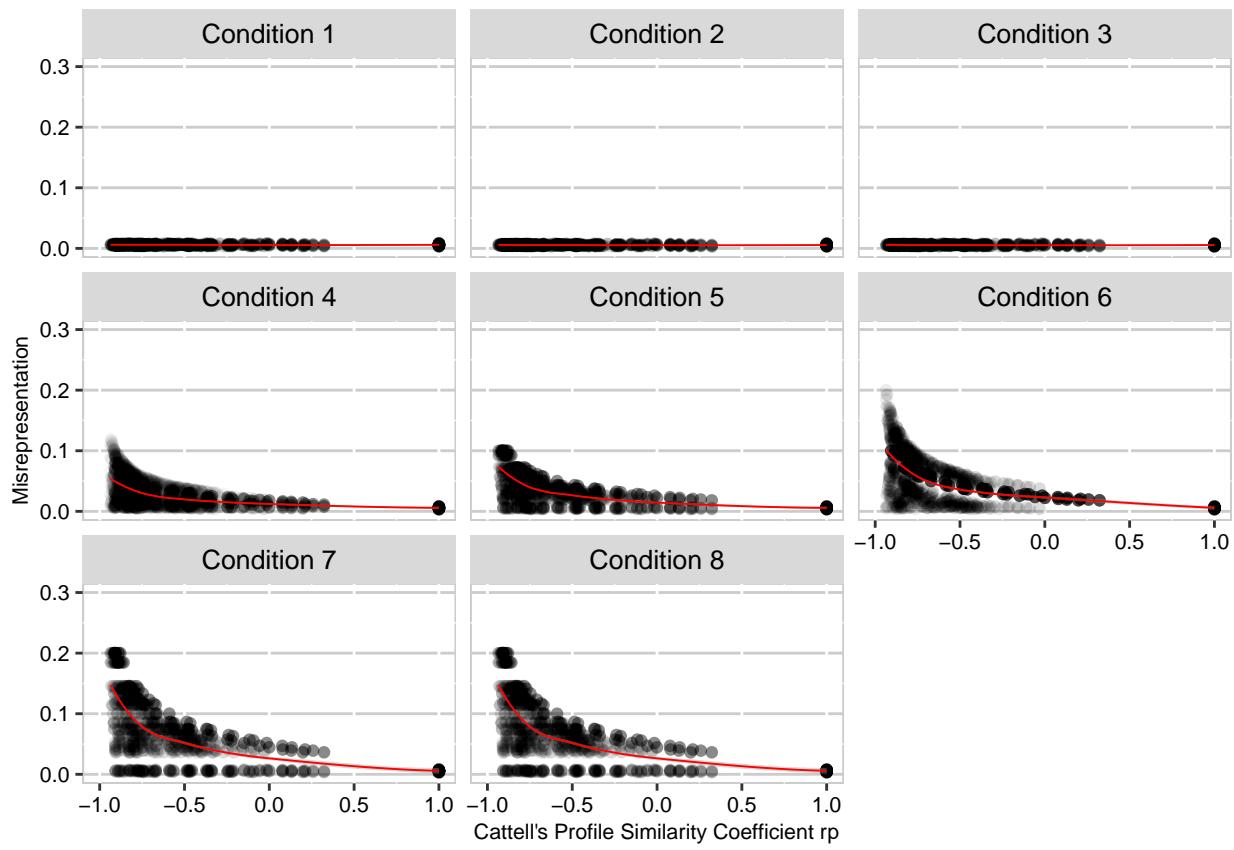


Figure 2

Relationship between sample representativeness and misrepresentation.

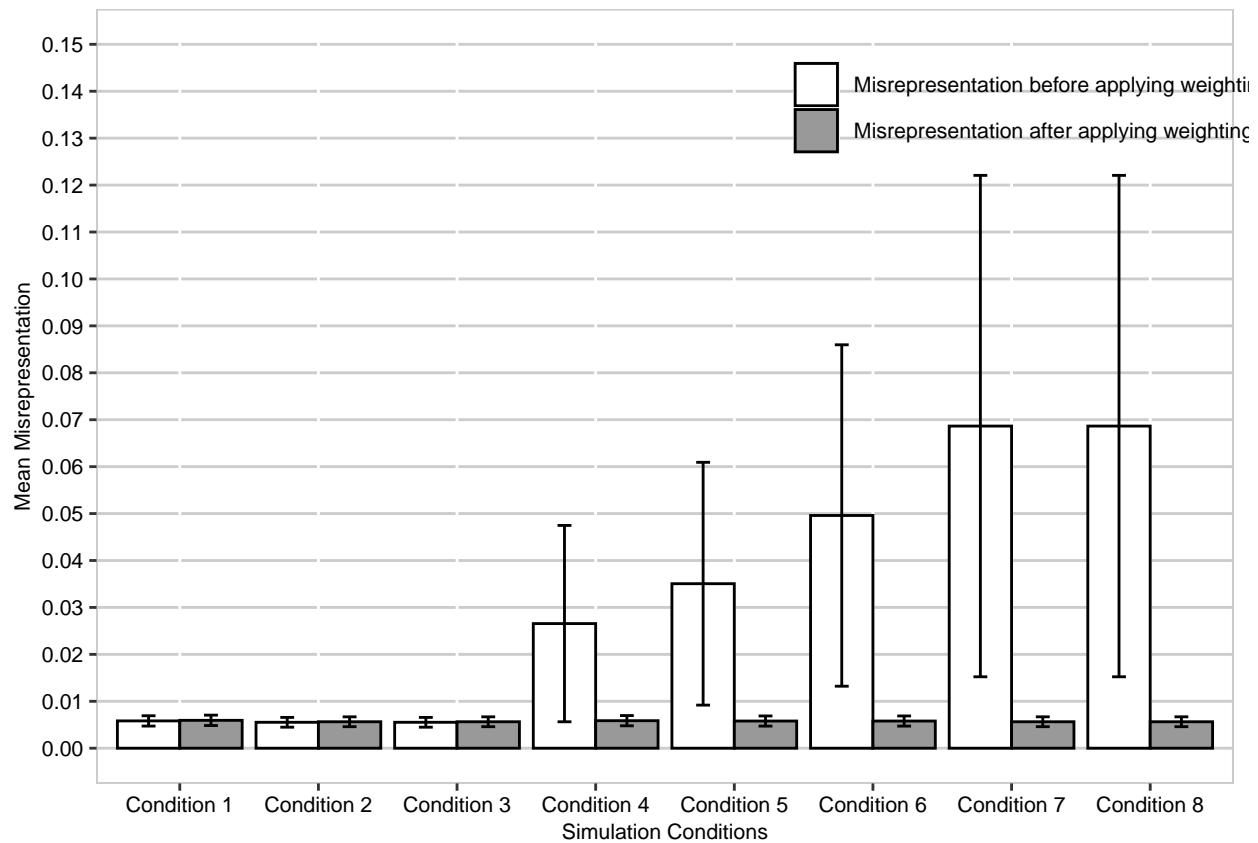
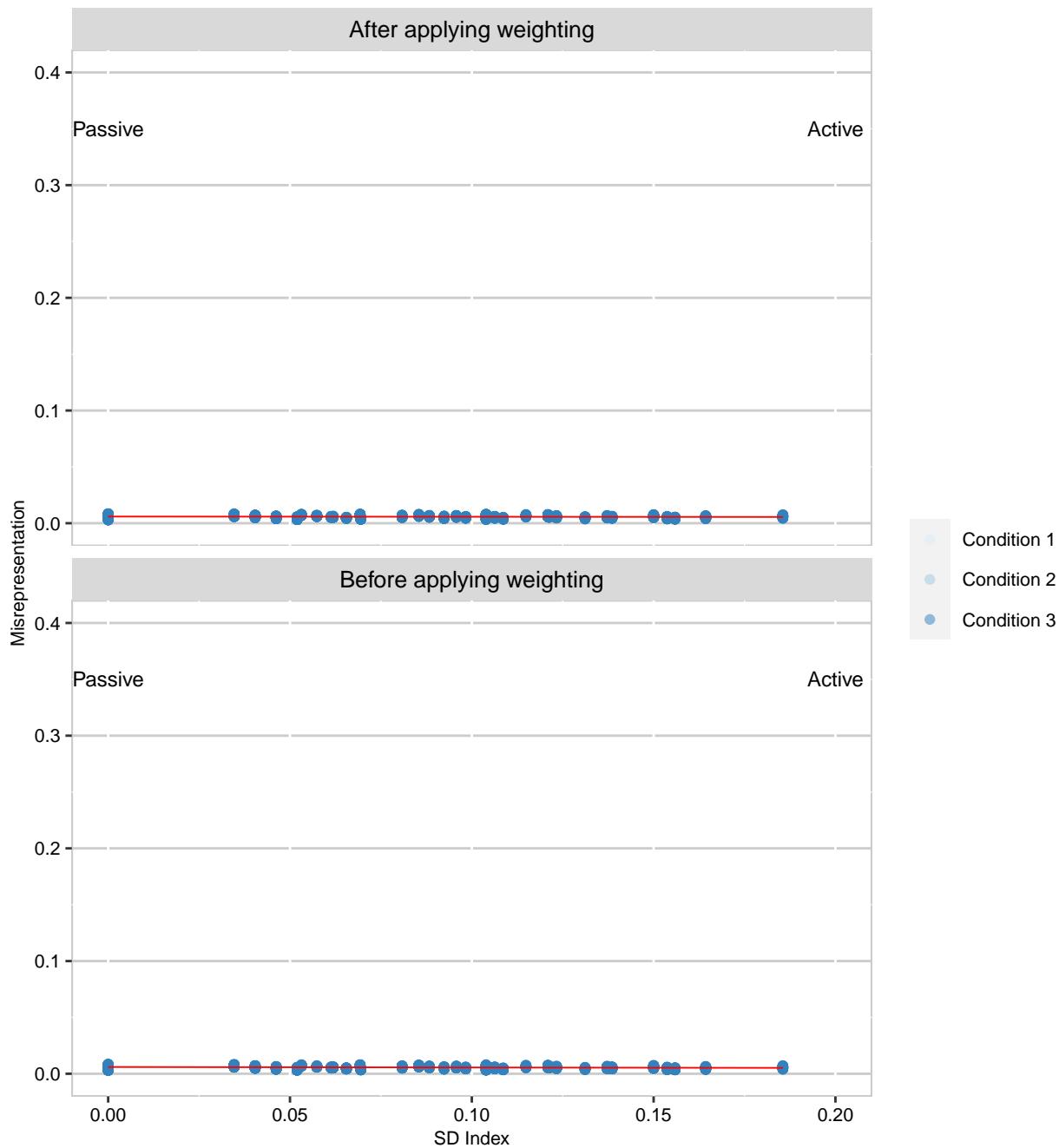


Figure 3

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

**Figure 4**

Presence (or lack) of error in unweighted and weighted sample estimates across passive and active forms of nonresponse (Conditions 1 through 3).

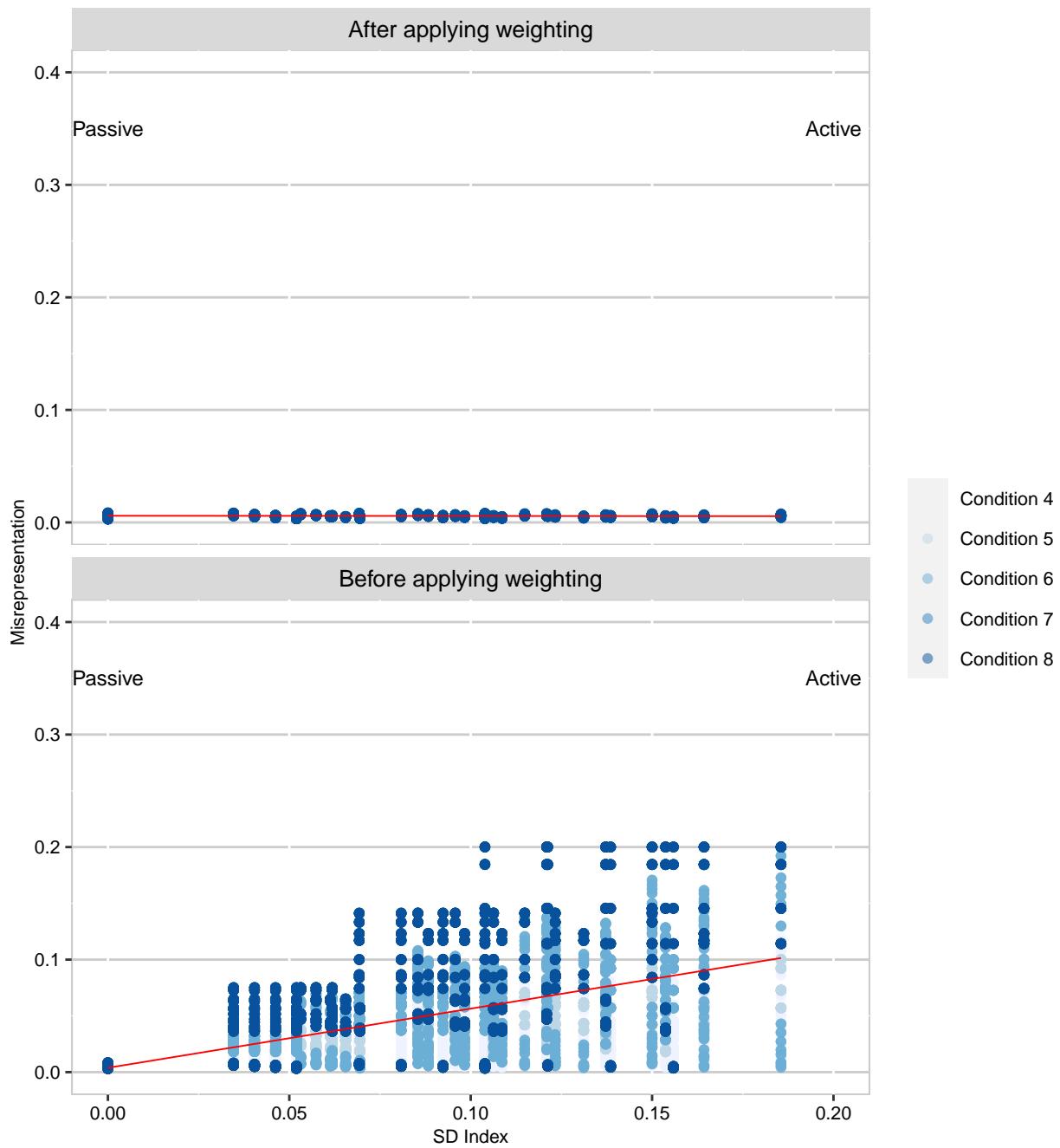


Figure 5

Presence (or lack) of error in unweighted and weighted sample estimates across passive and active forms of nonresponse (Conditions 4 through 8).