**Research Report Series**

THE UNIVERSITY
OF BRITISH COLUMBIA
UBC Paragon Research Initiative
*Statistical Science of Measurement*

# Surveys Using Self-Report Questionnaires Require a Nuanced Interpretation

## The Impact of Tacit Assumptions about the Response Process, Structure of the Target Population, and Reasons for Nonresponse

Bruno D. Zumbo , University of British Columbia

When interpreting surveys of self-report questions, the response rate is at the forefront of the mind of those wishing to make claims from the survey results. However, unlike response rate, which is directly observable (i.e., how many respondents compared to the total possible respondents) hidden from anyone making claims from the survey results, three elements of the survey response process have more impact on the findings than response rate. The three elements involve heterogeneous item response processes, population structure that informs the item response, and nonparticipation (unit nonresponse) processes that are not ignorable. These elements need to be considered in a more nuanced view of the interpretation of findings from self-report surveys.

Keywords: Unit nonresponse, Nonparticipation, Response Process Heterogeneity, Structured Target Population, Response Rate

Version date: September 15, 2021

## TABLE OF CONTENTS

## 1. Surveys Using Self-Report Questionnaires

### A recent report of a survey of self-reported vaccination status

Against the backdrop of debate on campus and in the local and national media about post-secondary institutions' insistence on a return to in-person teaching, the University of British Columbia sent the following media release in late August.

> UBC Broadcast | August 26, 2021 ["Subject: UBC implements vaccine declaration and rapid testing for COVID-19"]
> *A recent survey of all registered UBC Vancouver students (undergraduate and graduate; domestic and international) revealed that, as of August 16, 92% of the responding students (16,093 participated) are partially or fully vaccinated and 83% are fully vaccinated, with the majority of the remainder intending to be vaccinated before the start of classes.*

This announcement of the recent survey results was pointed to by UBC Department Heads/Directors, Deans, and on upward in the leadership structure in communications to faculty, staff, and students that the partial vaccination coverage (having received at least one dose) was 92% and that it will be higher as the response rate (i.e., the proportion of students who participate in, respond to, the survey) increases. This value of 92% was repeatedly relied on as part of the rationale to push ahead with the policy of in-person instruction and to suggest that their's was an empirically-based policy. Other local surveys of this nature continued to be reported weeks later by UBC and other post-secondary educational institutions across Canada and globally.

> In stark contrast to the claim by UBC of 92 percent vaccination coverage in late August, a sensitivity analysis shows that this survey result could have arisen from a population whose vaccination coverage could range anywhere from 58 to 87 percent vaccination coverage, depending on why students chose not to participate in the survey.
>
> It is precarious to set policy with this level of uncertainty about the population vaccination coverage.

The claim of 92% vaccination coverage by UBC based on the local survey stood in contrast to a national vaccination coverage of 72.5% and British Columbia vaccination coverage of 74.63% reported by the national health agency (Public Health Agency of Canada, 2021) for the same time. In addition to providing information about definitions, data sources, and data limitations and caveats, the national agency does not rely exclusively on self-report vaccination status.

### Purpose and structure of this research report

A sense of dis-ease over the naive interpretation of the campus survey results led me to ask how one should interpret the claims from the campus surveys of vaccine status and coverage. Much of the discussion around the interpretation of the campus survey results have been about the response rates. However, the interpretation should be more nuanced, involving heterogeneous item response processes, population structure that informs the item response, and nonparticipation (unit nonresponse) processes that are not ignorable. Unlike the response rate, population structure and these processes are often hidden from anyone interpreting the survey results.

The purpose of this research report is to demonstrate the impact that these other factors have on the claims made from self-report studies, provide readers with a framework of how to think about findings of this nature and a range of plausible vaccination coverage that one could infer from the campus surveys. I aim neither to present a detailed outline of the concepts nor to offer a grand theory that allocates a proper place to each of the concepts of the tactic assumptions when one makes claims from survey results of self-report questionnaires. Instead, the goal is to roughly outline three elements of survey item responding and examine how they may impact the claims one makes from survey results. Furthermore, this research report intends to be pragmatic. It provides a point of departure for non-specialists who want to know more about these concepts and, it can be seen as an invitation for more conceptual and methodological rigour.

I describe a framework that uses elementary probability and Bayesian computation with implications for interpreting self-reported COVID-19 vaccination status from a self-selected sample of respondents under a range of unit nonresponse (nonparticipation) patterns.

It is important to recognize that the findings reported herein build upon a vast literature in the related fields of probability theory, statistics, survey design, and the statistical theories of mental measurement. It reflects the emerging stochastic view of measurement and instrumentation based on my mathematical science of measurement and psychometrics results. Over and above texts on intermediate to advanced probability and statistics, readers interested in this topic would be well served with a close study of beta-Bernoulli and beta-binomial models of item responding, mixture models, and Bayesian data analysis. Sources that I have found particularly useful in the transdisciplinary research developing the concepts and theory include Neubauer, Djuraš, and Friedl (2011); Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin (2014); Gelman, King, and Liu (1998); Stoner, Economou, and Drummond (2019); Moreno and Girón (1998); Schouten, Schlomo, and Skinner (2011); Schouten, Cobben, and Bethlehem (2009); Brick (2013); Vermunt (2005); Banfield and Raftery (1993); Kroc and Zumbo, (2020) and sources cited therein.

The remainder of this research report is organized into three sections. Section 2 provides a list of remarks about self-report surveys and tacit assumptions when making claims from their results. Section 3 describes a stochastic view of the process of item responding in self-report questionnaires. The final section, Section 4, applies the models described in the third section to interpret the UBC self-report survey of COVID-19 vaccination status.

## 2. Remarks About Self-Report Surveys And Tacit Assumptions When Making Claims From Their Results

Survey participants are often asked to report their health status, opinion, attitude, values, or dispositions toward products, policies, or political parties. Examples include sensitive or politically charged questions about vaccine status, income levels, and socially stigmatized behaviours, such as sexual expression and engagement.

- Please note that this report's main message about how tacit assumptions about the response process, structure of the target population, and reasons for nonparticipation/nonresponse can greatly alter the claims one makes from the results of surveys apply to all self-report questions and not only about sensitive or politically charged questions.
- Although the questionnaire and survey design field are replete with references to sensitive or politically charged categories of survey questions, the demarcation between political, nonpolitical, and apolitical survey questions, for example, is not straightforward. In the end, this distinction is not as relevant as it first seems because all survey questions are mental probes of an individual respondent and hence are always political, sensitive, and potentially emotionally charged.

The survey involves soliciting responses from a target group of potential respondents. We can refer to this group as a population and the respondents as the sample. However, it should be noted that in the instances I imagine, this process is closer to an incomplete census than a survey that draws participants from a probabilistic sampling scheme.

When interpreting surveys of self-report questionnaires, the response rate is at the forefront of the thoughts of those wishing to make claims from the survey results. The common understanding is that if one were to have the responses from all potential respondents (i.e., all respondents in the population), then the survey results are purely descriptive. There would be no uncertainty in claims being made. At first blush, this is a reasonable assumption; however, it is often stated to imply a property of the survey statistics as the sample size gets larger that requires a subtle interpretation of the sort seen in the sub-section below titled 'Statistical properties of the proportion (percentage) of survey participants who responded yes.' That is, subtlety comes when considering whether increasing the response rate alone provides a better estimate of the population parameter. However, unlike the response rate, several assumptions are not explicitly voiced nor necessarily understood by those making claims or decisions from the survey results.

Unlike response rate, which is directly observable (how many respondents compared to the total possible respondents), these tacit assumptions are hidden from anyone making claims from the survey results, include heterogeneity of the population, heterogeneity of the response process to the survey question, and the reason for nonparticipation (a more nuanced view of response rate.

## 3. A Stochastic View of Item Responding in Self-Report Questionnaires

This section of the report aims to unpack three elements of the survey response process that may impact the findings of a self-report questionnaire based on a survey soliciting responses from a target group of potential respondents. The respondents from self-select to respond to the survey. Alternatively, stated differently, the potential respondents self-select to opt-out of the survey. The three elements of the response process I will highlight include heterogeneous item response processes, population structure that informs the item response, and nonparticipation (unit nonresponse) processes that are not ignorable.

These three elements are tacit assumptions of the survey response process. They need to be considered in a more nuanced view of the interpretation of findings from self-report surveys. A simple but plausible framework will be described in sufficient detail to inform a sensitivity analysis to investigate the dependence of the reported survey results on the three tacit assumptions.

I will describe a model for the item response followed by a description of the sampling of survey respondents. I will start with the simpler homogeneous case for each item response and survey sampling aspects of a self-report sample survey and then a model that allows for individual differences in the response probabilities.

**A model for the item response process for a yes/no response- survey participants provide a yes/no response based on the beta-Bernoulli**

Although providing this level of detail for a simple process like the Bernoulli may seem unnecessary, it will be useful to move on to more complex models. It is worth clarifying that a Bernoulli trial is an instantiation of a Bernoulli event for which the probability the event of interest occurs is $p$ and the probability the event of interest does not occur is $1-p$.

The Bernoulli event is the (simplest) model for responding to a yes/no survey question. A sequence of Bernoulli trials is called a Bernoulli process. Suppose the probability $p$ that the event of interest occurs remains the same from trial to trial. In that case, the sequence of Bernoulli trials is referred to as a Bernoulli process. We may describe the Bernoulli process with a sequence of indicator random variables $Y=(Y_1,Y_2,...)$ that take only the values 1 and 0, which in the setting of a self-reported yes or no response, respectively, to a survey question such as: *Are you partially or fully vaccinated for COVID-19?* Therefore, the indicator variables are independent and have the same probability density function $\mathbb{P}(Y_i=1)=p$. Bernoulli trials process corresponds to sampling from the Bernoulli distribution.

A simple mathematical characterization of the response process for each survey respondent $i$ of the first $n$ trials $(Y_1,Y_2,...,Y_n)$ forms a random sample of size $n$ respondents from the Bernoulli distribution. Therefore, the homogeneous item response process for a survey question involving two choices, the outcome of the response $Y_i$ for survey respondent $i$ may be characterized as a Bernoulli process,

$$Y_i \sim Bernoulli(p) \tag{1}$$

where the probability of responding ***yes***, denoted $p$ for a sequence of $i$ Bernoulli trials $i \in \mathbb{N}_+$, $\mathrm{i}\forall,\ \mathrm{i} \in [1,...,\mathrm{n}]$.

By randomizing the Bernoulli item response process for each $i$ of the Bernoulli trials, one can vary the probability of responding 'yes' for each survey respondent. Therefore, the indicator variables no longer have the same probability density function $\mathbb{P}(Y_i=1)=p_i$ for some or all of the survey respondents.

To introduce item response heterogeneity by randomizing the Bernoulli item response process, suppose we imagine a two-stage process in which we select a random probability of the event of interest, such as responding yes to a survey question, according to the beta distribution with parameters $a$ and $b$. Next, we apply the Bernoulli process with this probability of a yes response for each respondent $i$. This randomization of the item response process results in a beta-Bernoulli process represented as

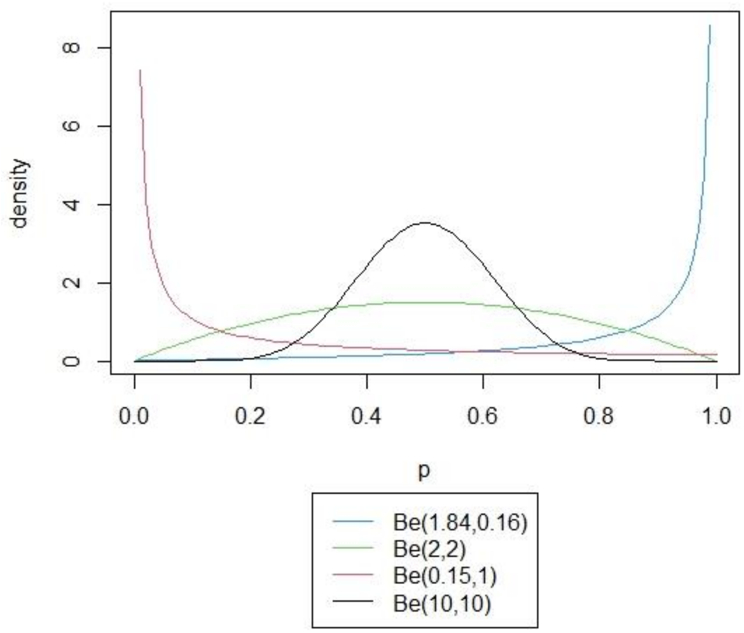
$$\begin{aligned}&Y_i \mid p_i \sim Bernoulli(p_i),\\&p_i \sim beta(a,b),\\&\mathrm{i}\forall,\ \mathrm{i} \in [1,...,\mathrm{n}] \text{ survey respondents in the population.}\end{aligned} \tag{2}$$

**Some remarks on the beta-Bernoulli model**

- If one views the beta-Bernoulli from the lens of Bayesian statistics, the original distribution $P$ is the prior distribution. Furthermore, the conditional distribution of given the data is the posterior distribution. Finally, a useful result is that the posterior distribution is beta whenever the prior distribution is beta, which means that the beta distributions are conjugate to the Bernoulli distribution.

- The beta-Bernoulli can be considered a response process model which adds plausible measurement uncertainty (measurement error or variability) to the response process to an individual's response process.
  - Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin (2014) describe this model as a form of unintentional missing data wherein the phrase *missing data* is used in this context to mean that for some or potentially all survey respondents, there is uncertainty in the observed response due to, if you wish, a latent process characterized by the beta distribution. From this lens, latent variables are, in a sense missing because they are not directly observed. Rather they are variables that govern the observed individual differences in the item response process. It should be noted that the beta-Bernoulli model in equation (2) ***does not represent purposeful disingenuous*** responding but rather natural individual differences among survey respondents.
    - In studies in biology or vision science, the a and b parameters of the beta distribution are interpreted as pseudo counts that correspond to prior beliefs of the data before starting the statistical experiment. In our setting, these would be the prior beliefs of the survey respondent about the question being asked in the survey.
    - Therefore, Gelman et al.'s interpretation of the uncertainty stems from the individual differences induced by the beta distribution for $p_i$, $\text{i}\forall$, $\text{i}\in[1,...,\text{n}]$ survey respondents in the sample. As such, equation (2) has a plausible interpretation as a latent response variable model.
      - Namely, each two-choice response is assumed to be related to a survey respondent's ***underlying (latent) response tendency*** for that survey question. These latent response tendencies are assumed to be independent beta r.v. for $\text{i}\forall$, $\text{i}\in[1,...,\text{n}]$ survey respondents in the sample.
      - For those not well-versed in the mathematics of probability, one can gain some intuition for the difference between the Bernoulli and beta-Bernoulli models if one considers that the expected value of a random variable is the weighted average of all possible values of the variable. For our purposes, the weight here means the probability of the random variable taking a specific value.
      - For a discrete random variable such as the two-choice survey question, the expected value of this random variable, denoted by $E(Y)=\sum y_i \times prob(y_i)$, indexed over the survey respondents $i$ for all survey respondents in the sample, where $y_i$ is the value with 1 (yes) and 0 (no) responses, and $prob(y_i)$ is the probability that $Y$ takes the value of $y_i$. Similarly, one can define variability $\text{var}(Y)$.
      - For the Bernoulli model in equation (1), $E(Y)=p$, $\text{var}(Y)=p(1-p)$, whereas $E\left(Y\right)=\frac{a}{a+b}$ and $\text{var}(Y)=\frac{a}{a+b}\frac{b}{a+b}$ for the beta-Bernoulli model in equation (2).

- One can take advantage of the useful result that the posterior distribution is beta whenever the prior distribution is beta, which means that the beta distributions are conjugate to the Bernoulli distribution. Several different parameterizations of the beta distributions can help interpret a posterior distribution of item responses.
    - The mean of the item response distribution would be $\frac{a}{a+b}$, the variance would be $\frac{ab}{(a+b)^2(a+b+1)}$, and the concentration of the distribution would be $a+b$.
        - The phenomenon that many random quantities are close to their mean with high probability is called concentration of measure quantified by the concentration parameter.
    - If one reparameterizes the beta distribution in terms of its mean $\mu$, concentration $\kappa$, and we solve for a and b of the beta distribution, we get $a=\mu\kappa$ and $b=(1-\mu)\kappa$.
- A graph of the density function of four beta distributions and their corresponding mean, variance, and concentration follows.
    - The difference between the shape of beta(2,2) and beta(10,10) is noteworthy.
    - In the case of beta(1,1), the uniform distribution was not graphed because it would be a flat line. Beta(1,1), therefore, recovers a homogeneous Bernoulli response process. The mean=0.50 and var=0.25 for beta(1,1).



| Beta(a,b) | | mean | var |
|---|---|---|---|
| 1.84 | 0.16 | 0.92 | 0.07 |
| 2 | 2 | 0.50 | 0.25 |
| 0.15 | 1 | 0.13 | 0.11 |
| 10 | 10 | 0.50 | 0.25 |

**Statistical properties of the proportion (percentage) of survey participants who responded yes**

The purpose of the self-report surveys is to quantify the proportion (percentage) of respondents who reported yes to the question of interest. Because these variables are central to the study of response processes, it is worth making some remarks about the expected value and variance of the sum and means as random variables in their own right to reassure us that the indices are interpretable.

- From probability theory and returning to equation (1), we can define a partial sum process associated such that for $n \in N_+$, let $S_n = \sum_{i=1}^{n} Y_i$ denote the number of yes responses for the first $n$ trials (experiments, if you wish). Therefore, $S = (S_0, S_1, ...)$ is the partial sum process associated with $Y = (Y_0, Y_1, ...)$.
- Next for $n \in N_+$, let the sample mean, or equivalently the proportion of yes responses in the first $n$ trials, be defined as $M_n = {S_n}/{n} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, where the properties of the r.v. $M_n$ follow easily from the corresponding properties of $S_n$.
- For a Bernoulli process described in equation (1) with the endorsement (yes or success) parameter $p$ for a sequence, we know that $M_n \to p$ as $n \to \infty$ with probability 1 (in mean and therefore also in distribution). Likewise, if one considers equation (2), the beta-Bernoulli, $M_n \to p$ as $n \to \infty$ with probability 1 (however, in mean square and therefore in distribution).
- As a final remark, the $E(M_n)$ is in a sense constant in $n \in N_+$ suggesting that $M_n$ has, in a sense, some kind of limit as $n \to \infty$. On the other hand, $\text{var}(M_n) = \frac{1}{n}\frac{ab}{(a+b)^2} + \frac{n-1}{n}\frac{ab}{(a+b)^2(a+b+1)}$ and therefore one can describe $\text{var}(M_n) \to ab/(a+b)^2(a+b+1)$ as $n \to \infty$.

**A Structured Population Related to the Phenomenon of Interest**

Having established a model for item responding allows for heterogeneity of the response process; the same task will be considered in terms of the structure of the population of respondents. Should the respondents be considered homogeneous (ignorable in terms of the item response process)? I would suggest that a plausible model of item responding may include a characteristic of a <u>sub-population structure</u> that informs the item response. Therefore, I problematize the response process by considering some characteristics of the item respondents (a structural item response propensity) across the population units that inform the item response.

In the model I describe in this section, the process described in equation (2), or equation (1), for that matter, is independent of the influence of a structured population related to the phenomenon of interest. They are independent but embedded. Although the distinction is somewhat arbitrary, one can get an intuition for this; one can think of equations (1) or (2) as cognitive involving person-level individual differences. In contrast, the model of sub-population strata, or population level propensity, I will introduce in this section is more societal or community-based, hence more sociological or ecological (see Zumbo et al. 2015 for an ecological model of item responding that motivates this approach).

**Latent population strata of survey respondents**

Suppose one were to take the question of self-reported vaccination status. In that case, it is reasonable to assume that the whole population of survey respondents has $k$ latent subpopulations (latent strata) who respond by a beta-Bernoulli process in equation (1) wherein $Y_i \mid p_{ij} \sim Bernoulli(p_{ij})$, $p_{ij} \sim beta(a_{ij}, b_{ij})$, and as described in equation (2), $i$ indexes respondents, and $j$ indexes the subpopulation for $j = 1,2,...,k$ strata, and the strata sample sizes, $n_j$, $\sum_j^k n_j = n$, such that $\pi_j = n_j / n$ are nonnegative and $\sum_{i=1}^k \pi_i$. The resultant probability function of the mixture distribution, which is a mixture of a finite number of Bernoulli models in equation (1) or beta-Bernoulli in equation (2), can be defined as

$$P(Y = y) = \sum_{j=1}^{k} \pi_j P(Y_j = y), \tag{3}$$

where, $E(Y) = \sum_{j=1}^{k} \pi_i E(Y^{(i)})$ and $\mathrm{var}(Y) = \sum_{j=1}^{k} \pi_j^2 \,\mathrm{var}(Y^{(i)})$.

**Latent continua of survey respondents**

It should be noted that although it is beyond the scope of this research report, one could consider reparameterizing equation (2) using what in our framework would be considered latent continua rather than strata. There are many subtleties in doing, so that go beyond the scope of this research report. For example, if one were to use a hyperprior, then the prior distribution (on the parameter of the underlying model) itself is a mixture density. That is, one could add a layer of priors on the $a$ and $b$ parameters of the beta-Bernoulli model in equation (2). Of course, it would be best if the resultant mixture distribution could be set up to take advantage of conjugacy.

**When choosing to participate in, respond to, a survey is not ignorable**

The survey reported by UBC involves soliciting responses from a target group of potential respondents. We refer to a collection of the target participants as a population and the respondents as the sample. However, it should be noted that the UBC survey is widely used and reflects a kind of self-selection of survey participants that this is closer to an incomplete census than a survey that draws participants from a probabilistic sampling scheme. The issue in these cases is not a matter of a sampling scheme but rather of participants opting out of the survey, assuming all members of the target population have been informed and given an opportunity to participate. The statistics research literature describes the outcome of opting out of responding as nonparticipation or unit nonresponse.

Let use the Draper-Lindley-de-Finetti (DLD) framework (see Draper, 1995; Zumbo, 2007 for descriptions of this framework) to get some intuition for when unit nonresponse is or is not ignorable. The main point to be taken from the DLD framework is the necessity to be explicit about the sorts of inferences one makes and that one can make. It is not that some inference is necessarily better than others (because this sort of value judgment needs to take the purpose of the survey into account), but rather that credible and defensible survey science requires one to be explicit about the sorts of inferences that are made and that can be made in a given context.

As Draper (1995) states:

> *Within this approach, the only inferential elements with objective reality are data values X you have already observed and data values Y you have not yet observed. Inference is then the process of quantifying your uncertainty about future observables Y on the basis of things you know, including the X values and the context in which they were obtained. Informally, one might call X the data you have and Y the data you wish you had; with this terminology, a statistical model supporting your inference is a mathematical story that links these two data sets*. (p. 119)

As I highlighted (Zumbo, 2007, p. 54-63), the function of much of contemporary survey statistical methods is to step in when the data are incomplete. In an important sense, we are going from what we have to what we wish we had. If we had available the complete data or information, we would know the true value of the response to the survey question, and no statistics beyond simple summaries would be required. There would be no need for complex models to infer the unobserved score from the observed data and, hence, no need to check the adequacy and appropriateness of such inferences.

In 2007 I stated that the occurrence of complete data or full information, as I describe it, is not commonly encountered, if ever, in the practice of survey research. Naturally, this leads to the common experiences that define what we call statistical modelling. First, the data you have is never really the data you want or need for your attributions, recommendations or decisions. No matter how much data you have, it is never enough. Without complete information, you will always have some error of measurement or fallible indicator variable. We get around data and information limitations by augmenting our data with assumptions. In practice, we are, in essence, using the statistical model to creat-ing new data to replace the inadequate data.

In the DLD framework, a problem of sampling arises, for example, when there is a finite population of participants of scientific or policy interest, and you can only get a subset of the whole population. In this case, the data you have is the information on the sampled individuals and the data you wish you had is the corresponding information from the unsampled people. The problems of sampling may, in turn, be thought of as special cases of a missing data problem. Also, it is important to note that much of the work in the predictive approach to inference described by Draper is founded on the notion of exchangeability. Exchangeability can be thought of as: without prior or additional/supplemental information, the data are similar - and hence, in a mechanical sense, exchangeable.

The conclusions of the DLD framework align with contemporary survey design wherein response rate plays a lesser role when judging the quality of attained survey responses.

Reasons for nonparticipation in a survey are usually categorized as the failure to contact the potential respondent, the inability to persuade the potential respondent to respond, and other reasons such as language problems and health problems that may prevent them from responding (Brick and Montaquila 2009).

Until relatively recently, nonresponse bias and response rates were often treated as equivalent. An implication is that surveys with low response rates were likely to have the potential for high nonresponse bias in the estimates. Great efforts were made to develop and test methods to incentivize participation and increase response rates. Moreover, the quality of sample surveys was determined by the response rate. As was noted earlier in this research report, this is not an entirely unreasonable approach and is likely driven by the fact that response rates have a kind of face validity, are easy to compute, provide a single quality index for an entire survey—several empirical papers demonstrating that the empirical relationship between response rates and nonresponse bias is not strong (e.g., Groves, 2006) spurred researchers to reconsider this presumption.

The exchangeability of the set of respondents and nonrespondents is the key concept. Although the approach based on exchangeability is useful to us to frame the calculations in the sensitivity analysis, it is conceptually related to developments in contemporary statistical survey theory through R-indicators and partial R-indicators indices for evaluating, comparing, monitoring, and improving surveys (Schouten, Bethlehem, Beullens, Kleven, Loosveldt, Luiten, Rutar, Shlomo, & Skinner, 2012). A noteworthy difference is that the class of R-indicators and partial R-indicators have been developed in the context of statistical sampling (rather than the self-selection process I am considering in this research report). These R-indices can be computed from extant survey results to judge the survey quality based on the highly technical definition of representativeness of survey responses. Schouten et al. (2009) describe indicators that decompose the variation in response propensities and are directly linked to R-indicators.

### Some Remarks Setting the Stage for the Next Section of the Research Report- Missing data mechanism and ignorable nonresponse processes

In the next section of the report, the theory described in this report will be used to design the computation for a sensitivity analysis to examine the robustness of the claims made from the UBC study.

The DLD framework's conceptualization of the exchangeability of the respondents and nonrespondents is also closely related to commonly used missing-data mechanisms missing-completely at-random (MCAR), missing at-random (MAR), and not-missing-at-random (NMAR; often described using the more confusing terminology missing not at random, MNAR), see Little and Rubin (2002) for a description of these missing-data mechanisms.

- For the purposes of this research report, one may describe MCAR as the setting wherein the respondents are on average the same as nonrespondents, MAR as the setting wherein within known subpopulations, respondents are on average the same as nonrespondents, and NMAR as even within subpopulations, respondents are different.

The concept of ignorable nonresponse becomes useful. As Matei and Ranalli (2015, p. 145) state:

> *In this paper, we are particularly interested in the case where the missing data mechanism is nonignorable, because nonresponse depends on characteristics of interest that are either observed only on the respondents or are completely unobserved, which leads to data that are Not Missing At Random (NMAR). This is typical of, but not limited to, surveys with sensitive questions (concerning drug abuse, sexual attitudes, politics, income, etc). (*Matei & Ranalli, 2015, p. 145)

By definition, if nonresponse is ignorable for a survey question, it does not contribute to bias in the estimates of those variables. For the purposes of this research report, the exchangeability criterion may be operationalized as a consistency of the ratio of the participation rates within each stratum to the base rate in the population. When they are equal, the nonresponse is ignorable.

In the next section of the report, the results from Section 3 will inform a sensitivity analysis dependence of the reported survey results on the three tacit assumptions.

A noteworthy initial observation from the material presented in Section 3 is that a change in the response rate should not substantially change the average reported COVID vaccination status. That is, as anticipated from the mathematical results in Section 3, the mean proportion (e.g., 92% vaccinated) is, in a sense, constant, so if the resultant self-report vaccination status from self-selected survey respondents creeps up with more respondents (i.e., in the language of Section 3, an increase in the number of trials), it may be a sign that some uncontrolled contaminating variable in the response process that is not reflected in the analysis of the survey responses. That is, a lurking nuisance variable may be confounding the survey findings and leading to nonignorable nonresponse.

## 4. A Sensitivity Analysis to Aid in the Interpretation of the UBC Self-Report Survey of COVID-19 Vaccination

A simple but plausible framework for item responding is described in Section 3 of this research report in sufficient detail to inform a sensitivity analysis to investigate the dependence of the reported survey results on the three tacit assumptions. The framework uses elementary probability and Bayesian computation with implications for interpreting self-reported COVID-19 vaccination status from a self-selected sample of respondents under a range of unit nonresponse (nonparticipation) patterns.

- If the self-report vaccination status from a large number of self-selected survey respondents creeps up with additional respondents (e.g., 92% to 98%), it *may* be a sign that an unaccounted-for lurking nuisance variable may be confounding the survey findings leading to nonignorable nonresponse (nonparticipation).

- In line with contemporary statistical science, the response rate does not play a role in the sensitivity analysis.

The sensitivity analysis aims to mimic the reported survey result on August 16 of 92% of the 16,093 students who responded to the survey question reported being partially or fully vaccinated. From UBC records earlier in the year, the size of the group of possible respondents was 58,562 students.

> UBC Broadcast | August 26, 2021 ["Subject: UBC implements vaccine declaration and rapid testing for COVID-19"]
> *A recent survey of all registered UBC Vancouver students (undergraduate and graduate; domestic and international) revealed that, as of August 16, 92% of the responding students (16,093 participated) are partially or fully vaccinated and 83% are fully vaccinated, with the majority of the remainder intending to be vaccinated before the start of classes.*
>
> *Of the 70,024 students that go to UBC, 58,462 are enrolled at UBC Vancouver, and the remaining 11,562 are at UBC Okanagan. Undergraduates makeup approximately 80 percent of students at UBC and are the majority on both campuses.*
> **https://www.ubyssey.ca/news/2020-21-enrolment-report/**

### Individual Differences in Item Response Processes: A beta-Bernoulli process with three latent strata of respondents

What is important to keep in mind is that the population structure is related to (i.e., informs) the item response. That is, the three strata reflect three tendencies or propensities toward (attitudes toward) COVID-19 vaccination. Although characterized by a discrete singular index of propensity to want to vaccinate, these strata reflect many different experiences, attitudes, and life courses that result in *individual differences and variability within latent strata*- each stratum includes respondents reflecting the entire range of propensities, $p$ of 0 to 1.

- Latent Strata 1: Characterized by a propensity to want to vaccinate with an average of 92% in response to the question of vaccination status
- Latent Strata 2: Characterized by an uncertain propensity toward vaccination with an average of 50% in response to the question of vaccination status
- Latent Strata 3: Characterized by a propensity not to want to vaccinate with an average of 13% in response to the question of vaccination status

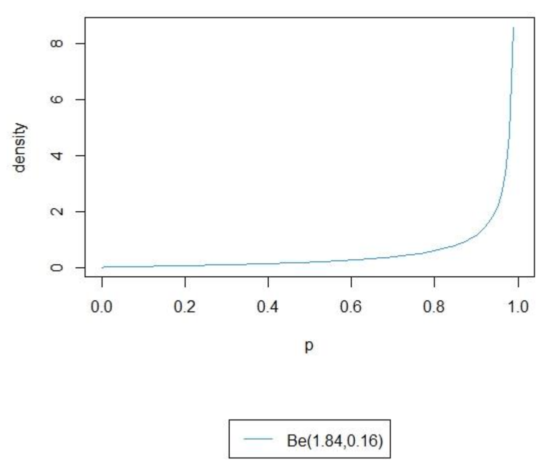
Below you will find a description and density plot for each of the three strata in this example.

Latent sub-population 1, latent stratum 1: propensity to want to vaccinate.

$$Y_i | p_i \sim Bernoulli(p_i),$$
$$p_i \sim beta(a_1, b_1),$$
$$i\forall, i \in [1,...,n_1]$$

For stratum 1, a=1.84 and b=0.16

| $beta(a_1, b_1)$ | | Mean | var |
|---|---|---|---|
| 1.84 | 0.16 | 0.92 | 0.07 |

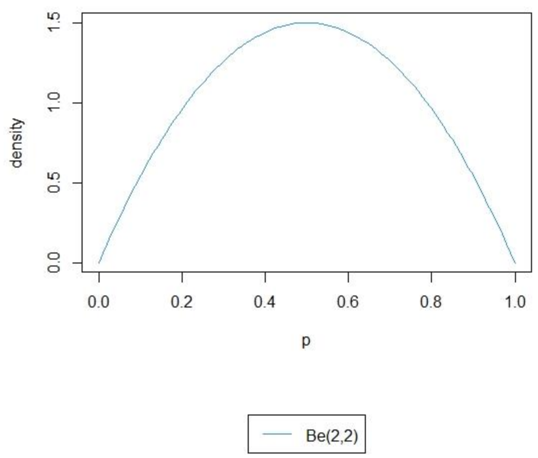


Latent sub-population 2, latent stratum 2: uncertain propensity toward vaccination.

$$Y_i \mid p_i \sim Bernoulli(p_i),$$
$$p_i \sim beta(a_2, b_2),$$
$$i\forall,\ i \in [1,\ldots,n_2]$$

For stratum 1, a=2 and b=2

| $beta(a_2, b_2)$ | | **Mean** | **var** |
|---|---|---|---|
| 2 | 2 | 0.50 | 0.25 |

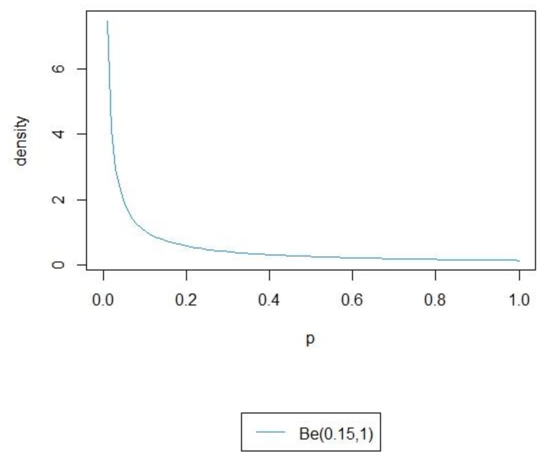


Latent sub-population 3, latent stratum 3: propensity not to want to vaccinate.

$$Y_i \mid p_i \sim Bernoulli(p_i),$$
$$p_i \sim beta(a_3, b_3),$$
$$i\forall,\ i \in [1,\ldots,n_3]$$

For stratum 1, a=0.15 and b=1

| $beta(a_3, b_3)$ | | Mean | var |
|---|---|---|---|
| 0.15 | 1 | 0.13 | 0.11 |




### 5. Conclusion: Base Rate Of Respondents In Each Stratum And Ignorable Nonresponse

Although many possible options could be proposed, I investigated five base rate scenarios with the base rates and corresponding percent of vaccinated students of the whole 58,462 students enrolled at UBC Vancouver.

| | Latent Strata 1: Characterized by a propensity to want to vaccinate | Latent Strata 2: Characterized by an uncertain propensity toward vaccination | Latent Strata 3: Characterized by a propensity not to want to vaccinate | Percent vaccinated of the whole 58,462 students enrolled at UBC Vancouver |
|---|---|---|---|---|
| Scenario A | 40% | 35% | 25% | 57.55% |
| Scenario B | 45% | 35% | 20% | 61.50% |
| Scenario C | 60% | 35% | 5% | 73.35% |
| Scenario D | 70% | 20% | 10% | 75.70% |
| Scenario E | 90% | 7% | 3% | 86.69% |

- Suppose we focus on scenario A. From the exchangeability definition provided in Section 3, the missing data mechanism is ignorable if the proportion of respondents matches each of the three strata.
    - Furthermore, suppose the survey of the self-reported vaccination status for the self-selected sample of respondents is uniform so that each stratum has 33.33% of the respondents. This distribution of the proportion of respondents across the strata does not match the base rate; therefore, the reported percent vaccinated would, in this case, be an underestimate of 51.67%.
    - If instead, the response proportions were 0.98, 0.019, and 0.001, then the reported percent vaccinated would, in this case, be a very large overestimate of 91.12%.
    - Therefore, a population vaccination coverage of 57.55% could result in survey findings ranging from 51.67% to 91.12%, depending on why students do not participate in the survey.
- If the UBC student population had the same vaccination coverage of 74.63% reported by the national health agency (Public Health Agency of Canada, 2021) for the same time, the base rates would need to be 69%, 20%, and 11% for each of stratum A, B, and C, respectively, which lends credence to the plausibility of the framework described in Section 3. This result is a reasonable validity check for the item response framework.
- It is worth noting that for the parameters I have mimicked in the sensitivity analysis, any of Scenarios A to E could result in survey findings ranging from 51.67% to 91.12%, depending on why students do not participate in the survey when the population vaccination coverage is 57.55% to 86.69%.
    - Stated another way, a survey finding of 91.12 percent, similar to what UBC reported in August of 2021, could arise from a population vaccination of anywhere from a population ranging from 57.55 to 86.69 percent vaccinated.
- This research report highlights a point made in contemporary statistical science that the reason why a respondent chooses not to participate in a survey (i.e., why they self-select to participate, leading to ignorable or nonignorable nonresponse) is relatively more informative about the quality of survey findings than response rate per se. The quality of survey data is about a lot more than the response rate alone.
    - Again, as seen throughout the contemporary statistical science literature, it is strongly advised that media releases be accompanied with a link where one can find information about response rates and, more importantly, sensitivity analyses of the sort described in this research report that compares characteristics of responders with non-responders, and if possible compare that information with known characteristics of the target population.
- Sensitivity analyses should include consideration of individual differences in the item response process. In this research report, I used an elementary probability model, the beta-Bernoulli, that allows for plausible variation in the response process.

## About the Author and Suggested Citation of this Report

About the author:

Bruno D. Zumbo, Distinguished University Scholar and Professor, is the Canada Research Chair in Professor Psychometrics and Measurement (Tier 1) and the Paragon UBC Professor of Psychometrics & Measurement at the University of British Columbia. He is a mathematical scientist working on measurement, psychometrics, and scientific methodology at the intersection of the behavioural, social and health sciences.

Webpage: **https://faculty.educ.ubc.ca/zumbo/cv.htm**

Suggested citation (APA Style):

Zumbo, B. D. (2021, September). *Surveys Using Self-Report Questionnaires Require a Nuanced Interpretation: The Impact of Tacit Assumptions about the Response Process, Structure of the Population, and Reasons for Nonresponse*. UBC-Paragon Research Initiative Research Report Series, University of British Columbia. URL: **Interpretation of surveys involving self-report-RR-3**

## References

Banfield, J.D., & Raftery, A.E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics, 49*, 803-821.

Brick, J. (2013). Unit Nonresponse and Weighting Adjustments: A Critical Review. *Journal of Official Statistics,29*, 329-353.

Brick, J.M., Montaquila, J., Han, D., & Williams, D. (2012). Improving Response Rates for Spanish-Speakers in Two-Phase Mail Surveys. *Public Opinion Quarterly, 76*, 721–732.

Draper, D. (1995). Inference and hierarchical modeling in the social sciences. *Journal of Educational and Behavioral Statistics 20*, 115-147.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2014), *Bayesian Data Analysis (Chapman and Hall/CRC Texts in Statistical Science) (Third ed.)*, London: Chapman and Hall/CRC.

Gelman, A. J., King, G., & Liu, C. (1998). Not Asked and Not Answered: Multiple Imputation for Multiple Surveys. *Journal of the American Statistical Association, 93*, 846-857.

Groves, R.M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly, 70*, 646–675.

Kroc, E., & Zumbo, B.D. (2020). A Transdisciplinary View of Measurement Error Models and the Variations of X=T+E. *Journal of Mathematical Psychology, 98*, 1-9.

Little, R. J. A. & Rubin, D.B. (2002). *Statistical Analysis with Missing Data*. New York: Wiley Series in Probability and Statistics

Matei, A., & Ranalli, M.G. (2015). Dealing with non-ignorable nonresponse in survey sampling: A latent modeling approach. *Survey Methodology, 41(1)*, 145-164.

Moreno, E., and Girón, J. (1998). Estimating With Incomplete Count Data a Bayesian Approach. *Journal of Statistical Planning and Inference, 66*, 147–159.

Neubauer, G., Djuraš, G., & Friedl, H. (2011). Models for Underreporting: A Bernoulli Sampling Approach for Reported Counts. *Austrian Journal Of Statistics, 40(1 & 2)*, 85-92.

Public Health Agency of Canada (2021). *Canadian COVID-19 vaccination coverage report*. Ottawa: Public Health Agency of Canada; August 27, 2021. **https://health-infobase.canada.ca/covid-19/vaccination-coverage/**

Schouten, B., Bethlehem, J., Beullens, K., Kleven, Ø., Loosveldt, G., Luiten, A., Rutar, K., Shlomo, N., & Skinner, C. (2012). Evaluating, Comparing, Monitoring, and Improving Representativeness of Survey Response Through R-Indicators and Partial R-Indicators. *International Statistical Review / Revue Internationale de Statistique, 80(3)*, 382-399.

Schouten, B., Schlomo, N., and Skinner, C. (2011). Indicators for Monitoring and Improving Representativeness of Response. *Journal of Official Statistics, 27*, 231–253.

Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology, 35*, 101-113.

Stoner, O., Economou, T., Drummond, G. (2019). A Hierarchical Framework for Correcting

Under-Reporting in Count Data. *Journal of the American Statistical Association, 114*, 1481-1492,

Vermunt, J. K. (2005). Mixed-effects logistic regression models for indirectly observed outcome variables. *Multivariate Behavioral Research, 40*, 281-301.

Zumbo, B.D. (2007). Validity: Foundational Issues and Statistical Methodology. In C.R. Rao and S. Sinharay (Eds.) *Handbook of Statistics, Vol. 26: Psychometrics*, (pp. 45-79). Elsevier Science B.V.: The Netherlands.

Zumbo, B.D., Liu, Y., Wu, A.D., Shear, B.R., Astivia, O.L.O. & Ark, T.K. (2015). A Methodology for Zumbo's Third Generation DIF Analyses and the Ecology of Item Responding. *Language Assessment Quarterly, 12*, 136-151.