

Assessing the Impact of Exclusive Extreme Response Style on Estimation of the Treatment Effect in Randomized Control Group Designs

Emily Mo

Abstract

Respondents to Likert-type surveys tend to interpret or respond to the levels of the Likert scale differently depending on factors such as cultural background or different personality traits. One common type of response style is the Exclusive Extreme pattern, in which respondents only select the extreme options on the Likert-type response scale. Data was simulated to mimic a balanced research design in which participants are randomly assigned to a treatment or control condition and respond to pre- and post-intervention surveys. Data was simulated to mimic a moderate treatment effect size. Using structural equation modeling to estimate the treatment effect on the latent variable measured by the survey instrument, an increased proportion of respondents using the Exclusive Extreme response style has a minimal effect on the statistical estimation of the treatment effect. However, as skewness increases in distributions of manifest item responses (on the ordinal response scale), an increased presence of the Exclusive Extreme response style has an increasingly depressive effect on the estimation of the treatment effect. Additionally, an increased number of items on the Likert scale intensifies the negative effect of response style contamination on the treatment effect. Thus, high proportions of the Exclusive Extreme response style in survey respondents may cause treatments to appear less powerful than they actually are, which can lead to the mistaken rejection of effective treatments in domains such as education, psychology, and medicine.

Introduction

The Likert scale is an ordinal response system commonly used on survey measures, and it can have two or more items on the scale, most commonly representing a range of agreement from “strongly disagree” to “strongly agree”. The responses to each question help to quantify an unobservable quality in respondents, which is referred to as a *construct*: for instance, a respondent’s self-reported understanding of an academic concept, or the extent to which a respondent is open to new experiences. This construct is associated, to a varying degree per question, with an individual’s responses to each question. The construct is assumed to be Normally distributed across the population. To generate a response to a particular question, we envision that the construct is multiplied by a constant corresponding to the question’s degree of association with the construct, and then displaced by a random Normally distributed error, producing a numeric value on a continuous scale. Then, responses to Likert-type survey questions can be viewed as a transformation from this value on a continuous scale—which represents the theoretical true extent to which the respondent agrees with the question’s prompt—to an ordinal value on the Likert scale. An individual’s raw continuous response to a question is considered to be an objective Normally distributed score that would be comparable across different individuals if it were observable; however, this response cannot possibly be observed because there are no objective and absolutely comparable means of self-reporting the extent to which one agrees with a survey question. The process of transforming an unobservable continuous response into an observable ordinal response is considered to be modeable by partitioning the continuous response scale into the same number of sections as the number of items on the question’s Likert scale. We refer to the cutoff points of the partitions as *Tau cuts*.

Typically, we assume Tau cuts divide the Normal distribution symmetrically such that the lowest partition of the continuous scale translates to the lowest Likert scale item, the second-lowest partition of the continuous scale translates to the second-lowest Likert scale item, and so on. Such a partitioning style lacks a response

style, or can be called a “neutral” response style, because it represents the way in which respondents are expected to answer Likert-type questions. However, the manner in which respondents answer survey questions often deviates from the neutral response style. For instance, some respondents may select the same item on the Likert scale for all questions regardless of their true latent continuous response. Under what is called the Exclusive Extreme response style, which we focus on in this study, a respondent only selects one of the two most extreme items on the Likert scale. Responding in such a manner increases the variation in Likert-type responses compared to a neutral response style, and essentially compresses the information from the respondent’s true unobserved continuous responses, so the presence of the response style may affect the statistical inference of associative measures based on the survey’s responses. Use of the Exclusive Extreme response style was found to be more prevalent among more individualist cultures and dominant personality types (Johnson et al. 2005), or in the presence of economic or climactic instability (He, Van de Vliert, and Van de Vijver 2016). Therefore, if the response style majorly impacts statistical estimates based on survey data, then two groups that are fundamentally identical in terms of the construct that a survey is measuring could come across as different simply because one group has a higher proportion of exclusively extreme-responding respondents than the other.

Structural Equation Modeling (SEM), which defines directional relationships between multiple variables that can be either latent and observed, is a common statistical modeling framework used to analyze survey data in the social sciences. In SEM, relationships between variables are represented by parameters that indicate the strength of the relationships. When survey data is assumed to be generated by a certain system of structural equations, models can be used to estimate the true parameters of that system based on the data. The results returned through SEM inference are affected by the function that transforms a respondent’s true unobserved responses into observed discrete Likert-type responses. Therefore, a large presence of the Exclusive Extreme response style in survey data may blur the apparent relationship between treatment group assignment and post-survey outcomes. Given existing literature about the prominence of the Exclusive Extreme response style and our understanding of how the response style may affect SEM outputs, we specifically hypothesized that an increasing presence of Exclusive Extreme response style introduces an increasing amount of bias into a treatment effect estimated through SEM from pre- and post-intervention survey data in a randomized control study.

Furthermore, previous studies have discussed the commonly asymmetric shape of ordinal distributions of Likert-type questions, particularly respondents’ bias towards the options near the left side of the scale (Friedman, Herskovitz, and Pollack 1993). This bias can be envisioned as the distribution of continuous responses, which is typically Gaussian in shape, getting mapped to an ordinal scale by Tau cuts that are asymmetrically distributed about that continuous distribution’s mean. Because a high-skew ordinal scale fails to carry the symmetry of its underlying continuous distribution, we also hypothesized that an increase in skewness leads to heightened bias in the estimation of the treatment effect in the presence of the Exclusive Extreme response style.

Since the transformation between continuous response and ordinal Likert-type response is unobservable, and since the true treatment effect parameter is impossible to know in real life and can only be estimated with statistical methods, there is no way to detect bias in treatment effect estimation using real survey data. In order to evaluate the accuracy of statistical inference in estimating an unobservable parameter, it is necessary to simulate many randomized control experiments with the same design, whose treatment effects were gauged through analyzing a pre- and post-intervention survey. This simulation process included the random generation of numeric values that represent the constructs that underlie each respondent’s responses, as well as creation of a basic structural equation model that uses a predetermined treatment effect to transform the constructs into responses for the two surveys. Each simulation represents a survey response body of a large fixed sample size. Certain parameters of choice can be held constant across all simulations (in our case, for instance, the treatment effect parameter is held constant), while other features of either the survey or the respondents can be varied in each simulation. After the generation of a large number of simulations, we use the method of confirmatory factor analysis (CFA) on our structural equation model for each simulation, which returns an estimate for the coefficient corresponding to the intervention’s effect on the post-intervention construct. Path coefficients returned through CFA indicate the estimated magnitude and direction of associations between factors in the structural equation model. Because the true

population-wide treatment effect coefficient is known under this simulated survey, and because it is controlled across all simulations, a linear regression can then be run to evaluate how the estimated treatment effect varies with different survey data features.

Methods

All simulations¹ were based on a structural equation model involving a pre-intervention survey and a post-intervention survey with identical questions, and a treatment administered during the time between the two surveys. Path coefficients were drawn from pre-intervention survey construct to post-intervention survey construct, from treatment assignment to post-intervention survey construct, and from each survey to its respective survey items. The path coefficient from the pre-intervention construct to post-intervention construct was held at 0.5, and represents a moderate treatment effect that we focus on for our study. The path coefficient from the pre-intervention construct to the post-intervention construct was held at 0.3, and the variance of the post-intervention survey was held at 1 to use as a reference point. The number of respondents per simulation was held at 250, and the number of the items on both pre- and post-intervention surveys was held at 15.

We chose to manipulate the following experimental variables across simulations: percentage of respondents using Exclusive Extreme responses, number of items on the Likert scale, skewness of ordinal distribution. We will henceforth refer to the percentage of respondents who used Exclusive Extreme response style as *percent contamination*. The percent contamination in the pool ranged from 0 to 39 in increments of 3. The number of items on the Likert scale varied throughout the values 4, 5, 6, 7, and 11. The amount of skew in the cutoff points ranged from -1 to 1 in increments of 0.2—with no skew, the cutoff points are symmetrically distributed about the mean. For each combination of experimental variables, we generated 1000 simulation iterations. This resulted in a total of 770,000 randomized control group simulations that were perfectly balanced across our experimental variables.

In each simulation iteration, using probability distribution functions in R, we first randomly generated values from a standard Normal distribution to represent a construct for the 250 respondents. For each of the 15 items on the survey, we randomly generated moderate factor loadings ranging from 0.45 to 0.85 from a Uniform distribution, as well as Normally distributed random errors, and used these values to generate each respondent’s true continuous response values based on our structural equation model.

Next, we assigned Tau cuts to partition the continuous distribution of survey responses into an ordinal scale representing the items on a Likert scale. The Tau cuts were determined by the number of items on the surveys’ Likert scales, the response style that an individual respondent used (either neutral or Exclusive Extreme), and the amount of skewness assigned to the ordinal distribution. Using these Tau cuts, we created Likert-type survey data for each of the respondents.

Results

After generating the data set, we first ran a regression of estimated treatment effect based on percent contamination and skewness using the following model structure, where Y represents estimated treatment effect, p represents percent contamination in increments of 10, s represents Tau cut skew in increments of 0.2, ϵ represents random error, and β s represent regression coefficients:

$$Y = \beta_0 + \beta_1 p + \beta_2 s + \beta_3 s^2 + \epsilon$$

The linear term of percent contamination and the quadratic term of skewness both had a moderate negative effect on the estimated treatment effect, while the linear term of skewness did not appear to have a contribution

¹Code to generate these simulations can be found at <https://github.com/emilymo008/survresponsestyles>.

to the estimated treatment effect. This output suggests that both presence of the Exclusive Extreme response style and skewness of the ordinal response distribution individually can lead to underestimation of the estimated treatment effect. After adding an interaction term, $(s^2) * p$, to the model between the quadratic term of skewness and percent contamination (see Table 1 in appendix), this term also had a moderate negative effect on the treatment effect estimate, suggesting that as skewness increases, the negative influence of percent contamination on estimated treatment effect is compounded at an increasingly higher rate. After observing an upward-facing parabolic pattern in the plot of fitted values vs. residuals, we decided that adding a quadratic term of percent contamination would be useful, since it would capture this remaining pattern in the residuals. This parabolic pattern suggested that the estimated treatment effect decreases at a progressively slower rate as percent contamination increases constantly and skewness is held constant.

Next, we fit individual models (see Table 2 in appendix) on the data for each level of skewness (from -1 to 1 in increments of 0.2, resulting in 11 total models), with terms for percent contamination up to the third degree. This way, we can compare the coefficients of each model in order to more easily assess percent contamination’s changing impact on estimated treatment effect as skewness increases. These models are represented by the following structure, in which Y represents estimated treatment effect, p represents percent contamination in increments of 10, j represents one of the skewness levels in the simulation, ϵ represents error, and β s represent regression coefficients:

$$Y_j = \beta_{0j} + \beta_{1j}p + \beta_{2j}p^2 + \beta_{3j}p^3 + \epsilon$$

The coefficients of these models reinforce the earlier finding that percent contamination has a negative impact on estimated treatment effect even when Likert-type survey data is not skewed. Even for the data in which the ordinal distribution of responses was completely symmetric, the corresponding model suggests that the estimated treatment effect still decreases as percent contamination increases—however, the estimated treatment effect begins to plateau around a percent contamination level of 0.4, by then already having fallen to 0.49. Such a bias is displayed in Figure 1: when the original tau cuts are skewed by a magnitude of 1, a percent contamination of 0.4 lowers the estimated treatment effect by more than 10%, whereas in the lack of a response style, a percent contamination of 0.4 only lowers the estimated treatment effect by a bit more than 2%. Skewness in the ordinal response distribution further impacts estimated treatment effect—models corresponding to greater magnitudes of skewness tend to have lower estimated treatment effects for uncontaminated data. Altogether, skewness of the ordinal response distribution combined with a high presence of the Exclusive Extreme response style can greatly bias a treatment effect approximated through SEM.

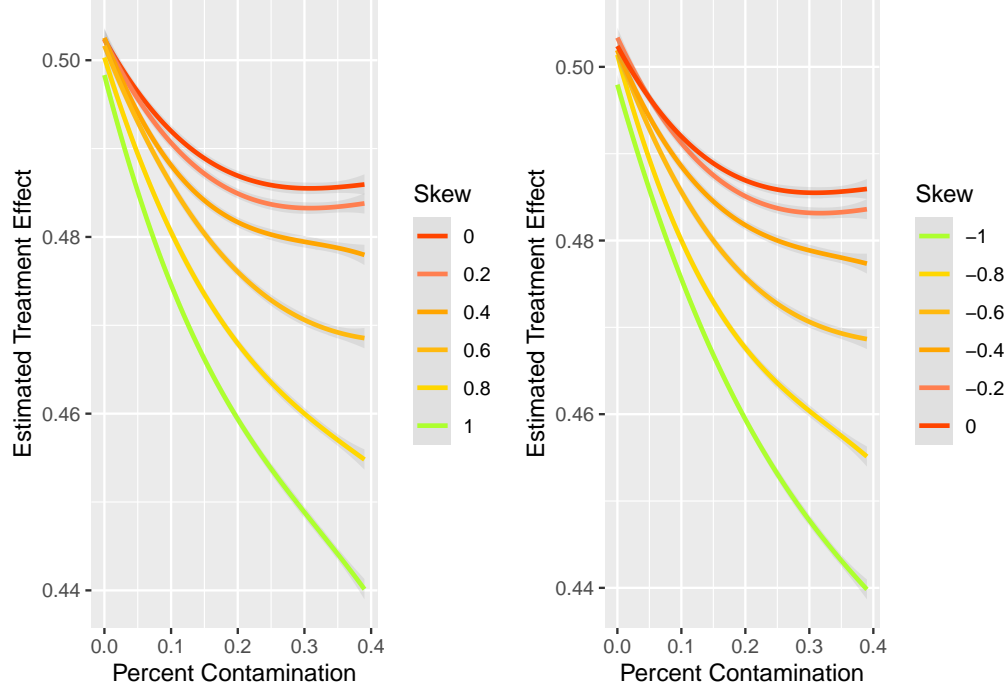


Figure 1: Smooth Curve of Effect of Percent Contamination on Estimated Treatment Effect, Grouped by Skewness

We next sought to understand whether Likert scale granularity had an effect on the estimated treatment effect, or whether there existed an interaction with the observed effects of percent contamination and skewness. We fit a preliminary model (see Table 3 in appendix) of the following structure, where Y represents treatment effect estimate, p represents percent contamination in increments of 10, s represents skewness in increments of 0.2, g represents Likert scale granularity as a factor, ϵ represents error, and β s represent regression coefficients:

$$Y = \beta_0 + \beta_1 s + \beta_2 p + \beta_3 g + \epsilon$$

The outputs of these models show no clear monotonic pattern from increasing the granularity of the Likert scale: the granularity coefficients suggest that when the values of percent contamination and squared skewness are held constant, we can expect the estimated treatment effect of a 5-granularity survey to be lower than that of a 4-granularity survey, that of a 6-granularity survey to be lower than that of a 5-granularity survey, that of a 7-granularity survey to be higher than that of a 6-granularity survey, and that of an 11-granularity survey to be lower than even that of a 6-granularity survey. Although the p-values of these coefficients are all very low, it is not difficult to obtain near-zero p-values given such a high sample size, and there was no obvious explanation for the coefficients' lack of monotonicity as granularity increases, so it was inconclusive from this model whether Likert granularity has a clear effect on the estimated treatment effect.

To better examine how the influence of granularity varies at different levels of skewness, we again fit individual models (see Table 4 in appendix) on the data for each skewness level, but this time predicted estimated treatment effect with percent contamination and Likert scale granularity, including an interaction term between granularity and percent contamination. This model structure is represented below, where Y represents estimated treatment effect, p represents percent contamination, g represents Likert scale granularity, j represents a skewness level, ϵ represents error, and β s represent regression coefficients:

$$Y_j = \beta_0 + \beta_{1j}p + \beta_{2j}p^2 + \beta_{3j}p^3 + \beta_{4j}g + \beta_{5j}p_j * g_j + \epsilon$$

Figure 2 is a visualization of the coefficients in these models which decomposes the complex interactions

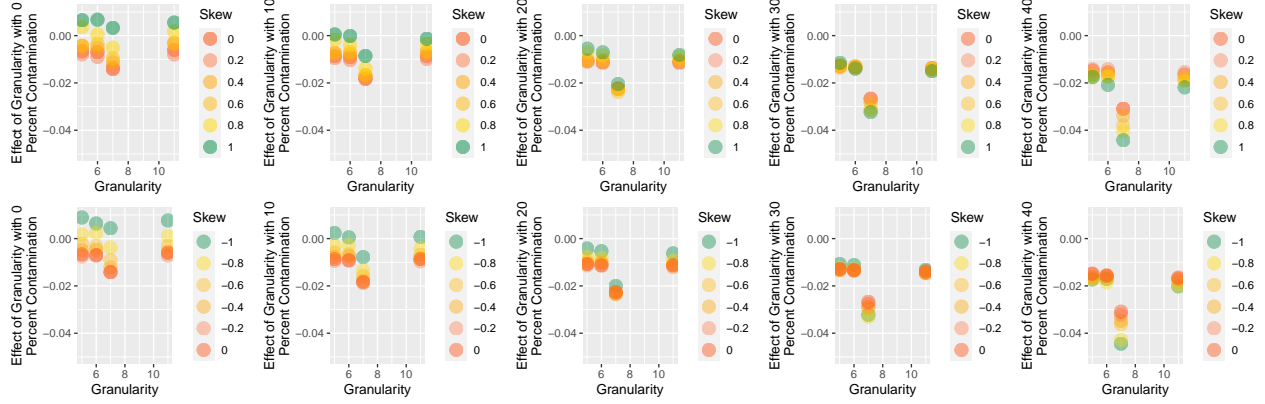


Figure 2: Effect Size of Various Likert Granularity Levels at Various Skewness Levels, Grouped by Percent Contamination

between percent contamination, Likert granularity, and skewness represented in this set of models. Each subplot is conditioned on a different value of percent contamination—for simplicity, there is a subplot 0, 10, 20, 30, and 40 percent contamination. Within each subplot, the horizontal axis represents a Likert granularity level of either 5, 6, 7, or 11, and the vertical axis represents the term for the corresponding granularity level added to the interaction term under the specified percent contamination level, which is equal to the expected change in estimated treatment effect from a granularity level of 4 to the specified level. This effect size is plotted as a separate point for each combination of skewness level and granularity level greater than 4, and colored based on skewness.

Based on Figure 2, the models' estimated difference in treatment effect compared to a granularity level of 4 generally becomes more negative as percent contamination increases, under a constant granularity level. This can be seen by the dots generally moving down as percent contamination increases. This observation may suggest that with an increasingly high presence of the Exclusive Extreme response style in the survey results, a Likert granularity of 4 is expected to have an increasingly smaller negative bias in estimated treatment effect compared to higher granularity levels. Furthermore, the coefficients for each of the granularity levels seem more grouped together around zero when there is zero percent contamination, suggesting that when there is no response style present, there may be a negligible difference between different granularity levels in terms of bias on the treatment effect estimate. However, despite the evidence that different granularity levels have different effects from one another when the response style is present, there is still no consistent pattern in the effect size among the different granularity levels, so it is unclear whether the bias in estimated treatment effect becomes steadily more severe as percent contamination increases.

From Figure 2, we can also see that within each granularity level, the influence of skewness on the effect of granularity appears to reverse as percent contamination increases: at low contamination levels, models fitted on higher-skewed data tend to identify a fixed granularity level of 5, 6, 7, or 11 as being associated with a less negative expected difference in estimated treatment effect from a granularity level of 4, compared to the models fitted on lower-skewed data. However, once percent contamination increases past a certain point (in this case, around 30% contamination), that relationship seems to get inverted due to an apparent interaction between skewness and percent contamination, and so models fitted on higher-skewed data now begin to identify a fixed granularity level of 5, 6, 7, or 11 as being associated with a more negative expected difference in estimated treatment effect from a granularity level of 4, compared to the models fitted on lower-skewed data. Therefore, at lower skewness magnitudes, there appears to be less of a difference in the effects of different granularity levels.

The standard error of the estimated treatment effect is also computed within the confirmatory factor analysis. For our final investigation using this simulation data, we examine how this standard error is impacted by the presence of the Exclusive Extreme response style. We fit a model (see Table 5 in appendix) of the following form, where S represents standard error of the estimated treatment effect, p represents percent contamination,

and ϵ represents error:

$$S = \beta_0 + \beta_1 p + \beta_2 p^2 + \beta_3 p^3 + \epsilon$$

The results of this model suggest that without controlling for other factors such as skewness and granularity, the standard error of the estimated treatment effect generally widens as percent contamination increases—the model’s expected standard error grew from about 0.045 under no contamination to about 0.048 under 40% contamination. Although the model fits a slight curve to the data, the trajectory of the expected standard error as percent contamination increases is roughly linear. The residuals of this model appeared trendless, and adding terms for skewness and Likert granularity to the model yielded relatively small effect sizes despite low p-values, so we did not further consider models with these additional terms. A general increase in standard error is unsurprising as the response style becomes more prominent, because the response style brings an increased variation in the Likert-type responses resulting from the response style, which is expected to negatively impact the certainty of the treatment’s effect on the post-intervention survey. Combining this result with our analysis of the Exclusive Extreme response style’s impact on treatment effect estimate itself, we find that as the response style becomes more prominent, we become increasingly uncertain of the estimated treatment effect, which itself becomes increasingly biased.

Discussion

Once a Likert-type survey response has been codified through the Exclusive Extreme response style, it is impossible to recover the response that would have resulted if the neutral response style was used. Because the Exclusive Extreme response style is so commonly used and we cannot necessarily prevent respondents from using this response style, it is useful to be aware of its biasing properties that this study highlights. Our simulations demonstrate that a heavy presence of the Exclusive Extreme response style in a group of respondents can make treatment coefficient estimates from structural equation models less trustworthy, because the response style binarizes the Likert scale and greatly reduces its precision in measuring continuous constructs. Our simulations also demonstrate that a more skewed underlying distribution of the target population’s ordinal responses can cause structural equation modeling methods to underestimate a treatment effect based on pre- and post- treatment surveys in a randomized control trial, and even more so in the presence of the Exclusive Extreme response style. Additionally, a higher presence of the Exclusive Extreme response style may lead to greater differences between the effect sizes of different levels of Likert scale granularity, and a higher presence of the response style can also modulate the direction of the effect that skewness has on each granularity level’s effect size. Finally, we posit that the presence of the Exclusive Extreme Response Style in survey data steadily decreases the level of confidence that an estimated treatment effect lies within an arbitrary interval. While an increase in percent contamination augmented the standard error at a moderate rate under the chosen initial circumstances that we used to generate our data set, we avoid extrapolating effects beyond the domain of our data, and thus we cannot rule out the possibility of this impact on the standard error being more drastic under varying circumstances—for instance, under different survey lengths or sample sizes.

The most concerning limitation to this analysis is that we were unable to clearly define or justify the estimated change in the differences between different Likert scale granularity levels as percent contamination increased. Although the p-values for the coefficients of each granularity level were near zero across all of the models fitted on data of different skewness levels, the changes in these coefficients were not monotonic, and in fact very erratic, in relation to an increasing granularity level. It is possible that there is a complex tradeoff between compression and clarification of response information as the number of items on the Likert scale increases, which may result in a decrease and subsequent increase in bias of the estimated treatment effect. There is also the less likely possibility that the patterns observed in the differences between the granularity resulted from random chance. In order to better understand this relationship, further simulation studies can be conducted in which more levels of Likert granularity are explored, perhaps under a different random seed. If many simulations are run using different specified treatment effects, it may also be possible to formulate a

model for the “tipping point” in percent contamination level, i.e. the point at which the influence of skewness on the effect of granularity is directionally reversed.

Another limitation of our analysis is that we cannot necessarily expect our models’ projections to be accurate upon extrapolating them beyond the percent contamination, skewness, and granularity levels present in the data on which the models were based. For instance, when other variables are held constant, the fitted values of estimated treatment effect decrease and then eventually increase again as percent contamination increases from 0, as is characteristic of its polynomial structure—however, in real life, we would expect the values to monotonically decrease, for it would be unlikely for the estimated treatment effect to recover its unbiasedness as percent contamination increases. Such monotonicity is expected because the Exclusive Extreme response style is a multiple-to-two mapping; thus, as the response style becomes more prevalent, more variation is introduced into the ordinal data, making it an increasingly imprecise representation of the respondents’ unobserved continuous responses, which consequently makes SEM estimation less able to detect a treatment effect.

Although this study brings forward the impact of the Exclusive Extreme response style on the structural equation modeling estimation of a moderate treatment effect, it does not consider scenarios with much stronger treatment effect sizes, in which a strong presence of the response style could possibly have a penalizing effect that is greater in magnitude, or scenarios with much weaker treatment effect sizes, in which a strong presence of the response style could possibly render the estimated treatment effect negligibly small. Additionally, the down-biasing effect of the response style, as well as its interactions with skewness and Likert scale granularity, may be exacerbated when the underlying treatment effect is larger.

Future research would ideally devise a correction for the bias in estimated treatment effect based on the detected presence of the Exclusive Extreme response style, Likert scale granularity, and estimated skewness of the ordinal response distribution. It may also be beneficial to devise a means of widening the confidence band about the estimate in order to reflect our confidence in the presence of the response style and the skewness of the ordinal response distribution. The task of creating a compensatory method for the bias is complicated by the lack of reliable means of detecting the Exclusive Extreme response style and ordinal response skewness. Detection of the Exclusive Extreme response style can vary in difficulty depending on length of survey and number of extreme responses, and is a strictly subjective judgment. There also remains the possibility that all of a respondent’s unobserved continuous responses on a survey are genuinely extreme. Additionally, even if such a robust compensatory method existed, the amount of skewness of the tau cuts that transform the unobserved continuous response into the ordinal Likert-type response must be assumed in the survey’s target population based on empirical data, and a poorly estimated skew may result in over- or under-compensation. Discarding Exclusive Extreme responses altogether is not recommended because it may reduce the generalizability of survey results, especially since response styles can be correlated with cultural, social, or psychological traits. Future research would ideally devise a correction for the bias in estimated treatment effect based on the detected presence of the Exclusive Extreme response style, Likert scale granularity, and estimated skewness of the ordinal response distribution. It may also be beneficial to devise a means of widening the confidence band about the estimate in order to reflect our confidence in the presence of the response style and the skewness of the ordinal response distribution. The task of creating a compensatory method for the bias is complicated by the lack of reliable means of detecting the Exclusive Extreme response style and ordinal response skewness. Detection of the Exclusive Extreme response style can vary in difficulty depending on length of survey and number of extreme responses, and is a strictly subjective judgment. There also remains the possibility that all of a respondent’s unobserved continuous responses on a survey are genuinely extreme. Additionally, even if such a robust compensatory method existed, the amount of skewness of the tau cuts that transform the unobserved continuous response into the ordinal Likert-type response must be assumed in the survey’s target population based on empirical data, and a poorly estimated skew may result in over- or under-compensation. Discarding Exclusive Extreme responses altogether is not recommended because it may reduce the generalizability of survey results, especially since response styles can be correlated with cultural, social, or psychological traits. Overall, this simulated study is a great oversimplification of real-life survey scenarios because it creates an unrealistically controlled setting with very few non-constant variables, whereas in real life, for instance, there can be multiple response styles within a survey body, or we may not be able to assume uniform Tau cuts across the entire survey population.

Such intricacies can make it even more difficult to detect the Exclusive Extreme response style or begin to understand how it may have affected the results of statistical analyses.

We assert that in a randomized control study setting that utilizes pre- and post-intervention surveys, a greater prominence of the Exclusive Extreme response style in a survey's responses leads to underestimation, and more uncertain estimation, of an intervention effect that the survey seeks to statistically estimate through confirmatory factor analysis. This underestimation is exacerbated when the distribution of the respondents' underlying construct is asymmetric. The amount of underestimation that occurs cannot be retrieved through examination of the survey responses, due to the unobservability of certain latent characteristics of the respondent body. Further research may explore attempts to harness the observable characteristics of the survey data to correct this bias. Survey researchers should pay close attention to the potential use of the Exclusive Extreme response style in their survey answers, and recognize that the true treatment effect may be higher than estimated if there is a large proportion of respondents who used this response style.

Appendix

Table 1: Output of preliminary model of estimated treatment effect using percent contamination, skewness

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.5055	1e-04	4016.3453	<2e-16
10% Contamination	-0.0083	<2e-16	-177.1172	<2e-16
Skewness (by 0.2)	<2e-16	<2e-16	0.2648	0.7911
Skewness (by 0.2) ²	-0.001	<2e-16	-162.237	<2e-16

Table 2: Coefficients of individual models of estimated treatment effect using percent contamination, by skewness

Standard errors are parenthesized.

Skewness	Intercept	10% Contamination	10% Contamination ²	10% Contamination ³
-1	0.498 (0.000587)	-0.026 (0.00135)	0.00392 (0.000825)	-0.000275 (0.000139)
-0.8	0.501 (0.000592)	-0.0271 (0.00137)	0.00639 (0.000833)	-0.000634 (0.00014)
-0.6	0.502 (0.000589)	-0.0201 (0.00136)	0.00408 (0.000828)	-0.000286 (0.000139)
-0.4	0.502 (0.000588)	-0.0176 (0.00136)	0.00464 (0.000827)	-0.000449 (0.000139)
-0.2	0.503 (0.000588)	-0.0157 (0.00136)	0.00389 (0.000826)	-0.000296 (0.000139)
0	0.502 (0.000586)	-0.0137 (0.00135)	0.00352 (0.000824)	-0.000284 (0.000139)
0.2	0.502 (0.000585)	-0.0152 (0.00135)	0.00388 (0.000823)	-0.000304 (0.000139)
0.4	0.503 (0.000587)	-0.0197 (0.00135)	0.00587 (0.000826)	-0.000626 (0.000139)
0.6	0.502 (0.000588)	-0.0189 (0.00136)	0.00348 (0.000827)	-0.000205 (0.000139)
0.8	0.5 (0.000588)	-0.0241 (0.00136)	0.0048 (0.000826)	-0.000411 (0.000139)
1	0.498 (0.000588)	-0.0293 (0.00136)	0.00619 (0.000827)	-0.000644 (0.000139)

Table 3: Output of preliminary model of estimated treatment effect using percent contamination, skewness, Likert granularity

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.5051	2e-04	3193.3119	<2e-16
Skewness (by 0.2)	<2e-16	<2e-16	0.263	0.7926
10% Contamination	-0.0083	<2e-16	-175.8616	<2e-16
Likert Granularity = 5	-0.0088	2e-04	-48.2735	<2e-16
Likert Granularity = 6	-0.0093	2e-04	-51.2995	<2e-16
Likert Granularity = 7	-0.0221	2e-04	-121.7274	<2e-16
Likert Granularity = 11	-0.0099	2e-04	-54.4478	<2e-16

Table 4: Coefficients of individual models of estimated treatment effect using percent contamination and Likert granularity, by skewness

Standard errors are parenthesized.

Skewness	Intercept	10% Contamination	10% Contamination ²	10% Contamination ³
-1	0.49248 (0.00091)	-0.0197 (0.00137)	0.00392 (0.00081)	-0.00027 (0.00014)
-0.8	0.50111 (0.00092)	-0.02214 (0.00138)	0.00639 (0.00082)	-0.00063 (0.00014)
-0.6	0.50548 (0.00091)	-0.0166 (0.00138)	0.00408 (0.00082)	-0.00029 (0.00014)
-0.4	0.50738 (0.00091)	-0.01472 (0.00137)	0.00464 (0.00082)	-0.00045 (0.00014)
-0.2	0.51058 (0.00091)	-0.01354 (0.00137)	0.00389 (0.00082)	-3e-04 (0.00014)
0	0.50916 (0.00091)	-0.01146 (0.00137)	0.00352 (0.00082)	-0.00028 (0.00014)
0.2	0.51 (0.00091)	-0.01342 (0.00137)	0.00388 (0.00081)	-3e-04 (0.00014)
0.4	0.50728 (0.00091)	-0.01662 (0.00137)	0.00587 (0.00082)	-0.00063 (0.00014)
0.6	0.50581 (0.00091)	-0.01549 (0.00137)	0.00348 (0.00082)	-2e-04 (0.00014)
0.8	0.50017 (0.00091)	-0.01938 (0.00137)	0.0048 (0.00082)	-0.00041 (0.00014)
1	0.49391 (0.00091)	-0.02295 (0.00137)	0.00619 (0.00082)	-0.00064 (0.00014)

Skewness	Likert Granularity = 5	Likert Granularity = 6	Likert Granularity = 7	Likert Granularity = 11
-1	0.00891 (0.00111)	0.00638 (0.00111)	0.00437 (0.00111)	0.00768 (0.00111)
-0.8	0.00175 (0.00112)	0.00234 (0.00112)	-0.00377 (0.00112)	0.00113 (0.00112)
-0.6	-0.0022 (0.00112)	-0.00334 (0.00112)	-0.00913 (0.00112)	-0.00306 (0.00112)
-0.4	-0.00484 (0.00111)	-0.00473 (0.00111)	-0.01218 (0.00111)	-0.00552 (0.00111)
-0.2	-0.00752 (0.00111)	-0.00717 (0.00111)	-0.01422 (0.00111)	-0.00717 (0.00111)
0	-0.00664 (0.00111)	-0.00699 (0.00111)	-0.01415 (0.00111)	-0.00603 (0.00111)
0.2	-0.00792 (0.00111)	-0.009 (0.00111)	-0.01348 (0.00111)	-0.00794 (0.00111)
0.4	-0.00426 (0.00111)	-0.00521 (0.00111)	-0.01129 (0.00111)	-0.00295 (0.00111)
0.6	-0.00431 (0.00111)	-0.00346 (0.00111)	-0.00962 (0.00111)	-0.00345 (0.00111)
0.8	0.0036 (0.00111)	0.00043 (0.00111)	-0.00508 (0.00111)	0.00174 (0.00111)
1	0.00649 (0.00111)	0.00674 (0.00111)	0.00328 (0.00111)	0.00552 (0.00111)

Skewness	Likert Granularity = 11	10% Contam. \times Gran. 5	10% Contam. \times Gran. 6	10% Contam. \times Gran. 7
-1	0.00768 (0.00111)	-0.00655 (0.00048)	-0.00586 (0.00048)	-0.0122 (0.00048)
-0.8	0.00113 (0.00112)	-0.00473 (0.00049)	-0.00521 (0.00049)	-0.00975 (0.00049)
-0.6	-0.00306 (0.00112)	-0.00368 (0.00049)	-0.00331 (0.00049)	-0.00676 (0.00049)
-0.4	-0.00552 (0.00111)	-0.00284 (0.00049)	-0.00276 (0.00049)	-0.00563 (0.00049)
-0.2	-0.00717 (0.00111)	-0.00188 (0.00049)	-0.00215 (0.00049)	-0.0045 (0.00049)
0	-0.00603 (0.00111)	-0.00203 (0.00048)	-0.00213 (0.00048)	-0.00419 (0.00048)
0.2	-0.00794 (0.00111)	-0.00157 (0.00048)	-0.00131 (0.00048)	-0.00438 (0.00048)
0.4	-0.00295 (0.00111)	-0.00302 (0.00049)	-0.00267 (0.00049)	-0.00564 (0.00049)
0.6	-0.00345 (0.00111)	-0.00308 (0.00049)	-0.00349 (0.00049)	-0.0071 (0.00049)
0.8	0.00174 (0.00111)	-0.00533 (0.00048)	-0.00442 (0.00048)	-0.00883 (0.00048)
1	0.00552 (0.00111)	-0.006 (0.00049)	-0.00689 (0.00049)	-0.01188 (0.00049)

Skewness	10% Contam. \times Gran. 11
-1	-0.00696 (0.00048)
-0.8	-0.00525 (0.00049)
-0.6	-0.00379 (0.00049)
-0.4	-0.00296 (0.00049)
-0.2	-0.00244 (0.00049)
0	-0.00261 (0.00048)
0.2	-0.00188 (0.00048)
0.4	-0.00393 (0.00049)
0.6	-0.00353 (0.00049)
0.8	-0.00524 (0.00048)
1	-0.00685 (0.00049)

Table 5: Output of model of treatment effect standard error using percent contamination

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.045	<2e-16	3812.0637	<2e-16
10% Contamination	0.0017	<2e-16	61.1026	<2e-16
10% Contamination ²	-3e-04	<2e-16	-19.3071	<2e-16
10% Contamination ³	<2e-16	<2e-16	10.4569	<2e-16

Acknowledgements

Great thanks to Professor Allen G. Harbaugh of the Boston University Mathematics and Statistics department for supervising this project and introducing me to the complicated world of survey response styles.

References

Friedman, Hershey, Paul Herskovitz, and Simcha Pollack. 1993. “Biasing Effects of Scale-Checking Styles on Response to a Likert Scale.” *American Statistical Association 1993 Proceedings of the Section on Survey Research Methods* 2.

He, Jia, Evert Van de Vliert, and Fons Van de Vijver. 2016. “Extreme Response Style as a Cultural Response to Climato-Economic Deprivation.” *International Journal of Psychology* 52: 67–71. <https://doi.org/10.1002/ijop.12287>.

Johnson, Timothy, Patrick Kulesa, Young Ik Cho, and Sharon Shavitt. 2005. “The Relation Between Culture and Response Styles.” *Journal of Cross-Cultural Psychology* 36 (2): 264–77. <https://doi.org/10.1177/0022022104272905>.