Supplement I: Careless Screening Procedure

Careless responses are partially random or inattentive responses in survey data. Previous studies estimate that approximately 10% - 12% of participants may be identified as careless responders (Meade & Craig, 2012). In this study, we used four indicators to help identify participants who may be careless responders on the RPI and C-RPI LongString (LS), Survey Time (ST), Number of Guttman errors (Gpoly), and Mahalanobis Distance (MD). LS is computed as the maximum number of consecutive items with the same response option chosen. In scales with reverse coded items, such as the RPI, larger LS values are indicators of careless responding. ST is computed as the time required for participants to complete the entire survey. If the ST values are too small, which means the participants completed the survey too quickly, this can be an indicator of careless responding. A Guttman error occurs when a less popular item step was taken, when a more popular item step was not taken. Large numbers of Guttman errors indicate the response patterns are inconsistent with the majority of the sample; those responses may be careless responses. MD is a multivariate outlier statistic (Mahalanobis, 1936) that summarizes the distance between observations and the center of the data, while taking the correlational structure into account. Responses with large MD values are likely to be the results of careless responses. All four indicators have demonstrated good sensitivity and specificity with simulated data (Desimone et al., 2015; Meade & Craig, 2012; Niessen et al., 2016).

The screening procedure involved two phases, with phase I prioritizing sensitivity over specificity (to assure that responses that may be careless were sent to phase II validation). In phase I, for each indicator we selected two cut scores to divide the responders into three subsets: individuals at high risk, low risk, and no risk for careless responses. Cut scores were determined based on the distributions of the indicators, outlier statistics, and suggestions from the literature (Desimone et al., 2015; Meade & Craig, 2012; Niessen et al., 2016). The cut scores for each indicator were determined based on data from each sample, hence no fixed cut scores were applied across samples. Additionally, common and unique screening rules were applied across indicators. Three high quantiles of the responses (top 5%, 10%, and 15%) and responses that appear to be outliers were identified for all four indicators across the three samples. When screening careless responses for ST, the researchers took into consideration the minimal time required to carefully complete the entire survey battery (which contained a number of additional scales as well), and participants who competed the survey too fast were marked as at high risk. Participants who used an extremely long time (>12 hours) to complete the survey battery were marked as at low risk. In terms of LS, because C-RPI uses a four-alternative Likert-type scale with 9 items reverse coded, any respondents who answered 1 (lowest response) or 4 (highest response) for 18 or more items in a row were marked as at high risk. Participants who responded 2 or 3 for 18 or more consecutive items may in truth be up and down in the center, so, even though such response patterns may be suspicious, these responders were marked as at low risk. In this first step, no participants’ data were set aside (excluded). Rather, we used these cut scores to assemble two databases: one database (clean) with participants flagged by none of the four indicators, or only flagged as low risk by one indicator; another database (at risk) with the remaining participants, who were flagged by one indicator as high risk, or by two or more indicators.

Using the “at-risk” dataset, the first group determined to be probable careless responders were selected using indicator combination: Responses marked as at (1) low risk by three or four indicators, or (2) high risk by two or more indicators, or (3) at high risk by one indicator and low risk by two or more than two indicators were set aside (excluded from further analysis). Responses that were relatively ambiguous, like respondents with two low-risk indicators, one high risk indicator, as well as one high and one low indicators, were sent to phase II screening.

In phase II screening, we applied several strategies to determine which participants’ data to keep or set aside. Strategies in phase II included investigating consistency in response patterns within subscales (e.g. responses in each subscales should look similar, hence if two or more “1 and 4” response pairs were observed for a responder, that individual’s data was set aside), and score discrepancy across subscales (e.g. after reverse coding, the subscale scores for the two factors should look similar. If substantial between-scale-score discrepancy was observed, that respondent’s data was set aside). By combining both screening procedures, a final list of careless respondents was identified.

Supplement II: Additional Tables

|  |
| --- |
| **Table 1. Corrected Item-Total Correlation** |
|  | China | Taiwan | U.S. |
| *Reward Probability* |  |  |  |
|  RPI1 | .51 | .57 | .57 |
|  RPI2 | .54 | .57 | .52 |
|  RPI4 | .53 | .72 | .64 |
|  RPI5 | .65 | .68 | .63 |
|  RPI6 | .67 | .67 | .63 |
|  RPI8 | .57 | .61 | .58 |
|  RPI10 | .54 | .51 | .56 |
|  RPI11 | .46 | .62 | .57 |
|  RPI15 | .53 | .71 | .61 |
|  RPI18 | .47 | .39 | .53 |
|  RPI20 | .45 | .51 | .53 |
| *Environmental Suppressor* |  |  |  |
|  RPI3 | .44 | .48 | .41 |
|  RPI7 | .24 | .49 | .55 |
|  RPI9 | .42 | .53 | .56 |
|  RPI12 | .35 | .32 | .38 |
|  RPI13 | .48 | .55 | .65 |
|  RPI14 | .52 | .63 | .61 |
|  RPI16 | .62 | .62 | .69 |
|  RPI17 | .47 | .56 | .48 |
|  RPI19 | .41 | .55 | .52 |

|  |
| --- |
| **Table 2. Factor Structure for C-RPI in China, Taiwan, and the U.S. with WLSMV Estimator** |
|  | China | Taiwan | U.S. |
| Items | RP | ES | RP | ES | RP | ES |
| (RPI1) I have many interests that bring me pleasure. | .58 | --- | .64 | --- | .59 | --- |
| (RPI2) I make the most of opportunities that are available to me. | .54 | --- | .58 | --- | .58 | --- |
| (RPI4) I make friends easily. | .52 | --- | .71 | --- | .62 | --- |
| (RPI5) There are many activities that I find satisfying. | .65 | --- | .75 | --- | .61 | --- |
| (RPI6) I consider myself to be a person with many skills. | .73 | --- | .72 | --- | .60 | --- |
| (RPI8) I feel a strong sense of achievement.  | .69 | --- | .67 | --- | .64 | --- |
| (RPI10) It is easy to find good ways to spend my time. | .61 | --- | .56 | --- | .69 | --- |
| (RPI11) I have the abilities to obtain pleasure in my life. | .55 | --- | .67 | --- | .65 | --- |
| (RPI15) I have good social skills. | .54 | --- | .71 | --- | .61 | --- |
| (RPI18) I have been very capable in jobs I have had. | .51 | --- | .40 | --- | .58 | --- |
| (RPI20) I have many opportunities to socialize with people. | .46 | --- | .50 | --- | .56 | --- |
| (RPI3) My behaviors often have negative consequences. | --- | .56 | --- | .66 | --- | .53 |
| (RPI7) Things happen that make me feel helpless or inadequate. | --- | .20 | --- | .61 | --- | .60 |
| (RPI9) Changes have happened in my life that have made it hard to find enjoyment. | --- | .53 | --- | .56 | --- | .66 |
| (RPI12) I have few financial resources, which limits what I can do. | --- | .31 | --- | .27 | --- | .38 |
| (RPI13) I have had many unpleasant experiences. | --- | .47 | --- | .45 | --- | .66 |
| (RPI14) It seems like bad things always happen to me. | --- | .68 | --- | .68 | --- | .69 |
| (RPI16) I often get hurt by others. | --- | .78 | --- | .66 | --- | .69 |
| (RPI17) People have been mean or aggressive toward me. | --- | .55 | --- | .61 | --- | .47 |
| (RPI19) I wish I could find a place to live that brought more satisfaction to my life. | --- | .44 | --- | .66 | --- | .65 |
| Inter-factor correlation | .47 | --- | .55 | --- | .65 | --- |
| Note. RP = reward probability; ES = environmental suppressor  |  |  |  |  |  |  |

Supplement III: DIF Results for Equal Gender Subsamples

|  |
| --- |
| **Table 1. IRT graded model DIF detection results for *Reward Probability* factor – equal gender** |
|  | China (n=152) vs. U.S. (n=212) | Taiwan (n=224) vs. U.S. (n=212) |
|  | Graded G2 (3) | *p* | Graded G2 (3) | *p* |
|  |  |  |  |  |
|  | All other items as anchor | All other items as anchor |
|  RPI1 | 2.5 | .64 | 0.6 | .96 |
|  RPI2 | 5.7 | .22 | **21.0** | **.0003** |
|  RPI4 | 4.9 | .30 | 5.3 | .26 |
|  RPI5 | 7.0 | .14 | 3.7 | .30 |
|  RPI6 | 4.7 | .32 | 2.4 | .66 |
|  RPI8 | 2.0 | .73 | 2.8 | .58 |
|  RPI10 | 0.5 | .91 | 0.9 | .83 |
|  RPI11 | 2.5 | .48 | 0.2 | .97 |
|  RPI15 | 6.5 | .17 | 11.1 | .03 |
|  RPI18 | **15.3** | **.002** | **22.7** | **.0001** |
|  RPI20 | 3.2 | .36 | 2.0 | .57 |
| *Note.* Significant tests with the p-values evaluated using the Benjamini-Hochberg procedure are bold. |

|  |
| --- |
| **Table 5. IRT graded model DIF detection results for *Environmental Suppressor* factor – equal gender** |
|  | China (n=152) vs. U.S. (n=212) | Taiwan (n=224) vs. U.S. (n=212) |
|  | Graded G2 (3) | *p* | Graded G2 (3) | *p* |
|  |  |  |  |  |
| *Step 1* | Anchor item – RPI 14 | Anchor item – RPI 14 |
|  RPI3 | **14.2** | **.00** | **18.9** | **.00** |
|  RPI7 | **28.4** | **.00** | **13.1** | **.01** |
|  RPI9 | **10.9** | **.03** | **20.7** | **.00** |
|  RPI12 | **39.8** | **.00** | **22.0** | **.00** |
|  RPI13 | **21.8** | **.00** | **34.5** | **.00** |
|  RPI16 | 3.9 | .42 | 0.7 | .31 |
|  RPI17 | **27.4** | **.00** | **38.5** | **.00** |
|  RPI19 | **16.8** | **.00** | **11.8** | **.02** |
|  |  |  |  |  |
| *Step 2* | Anchor items – RPI 14, 16 | Anchor items – RPI 14, 16 |
|  RPI3 | **14.5** | **.00** | **18.3** | **.00** |
|  RPI7 | **36.8** | **.00** | **14.4** | **.00** |
|  RPI9 | **12.8** | **.01** | **20.9** | **.00** |
|  RPI12 | **41.3** | **.00** | **22.0** | **.00** |
|  RPI13 | **24.8** | **.00** | **36.5** | **.00** |
|  RPI17 | **22.0** | **.00** | **35.3** | **.00** |
|  RPI19 | **18.9** | **.00** | **11.4** | **.02** |
| *Note.* Significant tests with the p-values evaluated using the Benjamini-Hochberg procedure are bold. |



*Figure 1.* Equal gender subsamples. Graded IRT model expected score curves for *Reward Probability* item 18 with all items as anchor. The vertical axes represent the item expected score, 0 = completely disagree; 1 = disagree; 2 = agree; 3 = completely agree; the horizontal axes is the latent variableReward Probability, -3 = three standard divisions below the population mean, 0 = population mean, +3 = three standard divisions above the population mean. Red dash line = China; Blue solid line = U.S.

*Figure 2.* Equal gender subsamples. Graded IRT model expected score curves for *Reward Probability* items 12 and 18 with all items as anchor. The vertical axes represent the item expected score, 0 = completely disagree; 1 = disagree; 2 = agree; 3 = completely agree; the horizontal axes is the latent variableReward Probability, -3 = three standard divisions below the population mean, 0 = population mean, +3 = three standard divisions above the population mean. Green dash line = Taiwan; Blue solid line = U.S.

*Figure 3.* Equal gender subsamples. Graded IRT model expected score curves for the *Environmental Suppressor* items with items 14 and 16 as anchor. The vertical axes represent the item expected score, 0 = completely agree; 1 = agree; 2 = disagree; 3 = completely disagree; the horizontal axes is the latent variableEnvironmental Suppressor, -3 = three standard divisions above the population mean, 0 = population mean, +3 = three standard divisions below the population mean. Red dash line = China; Blue solid line = U.S.

*Figure 4.* Equal gender subsamples. Graded IRT model expected score curves for the *Environmental Suppressor* items with items 14 and 16 as anchor. The vertical axes represent the item expected score, 0 = completely agree; 1 = agree; 2 = disagree; 3 = completely disagree; the horizontal axes is the latent variableEnvironmental Suppressor, -3 = three standard divisions above the population mean, 0 = population mean, +3 = three standard divisions below the population mean. Green dash line = Taiwan; Blue solid line = U.S.

**References**

Mahalanobis, P. C. (1936). Mahalanobis distance. *Proceedings National Institute of Science of India, 49*(2), 234–256.