

# Survey Quality Metrics Notes

12 September, 2023

## Careless / Insufficient Effort Responding (IER) / Satisficing

### Bogus items / Infrequency Approach

- Definition:
  - Bogus items are questions that have an obvious correct answer, so that an incorrect answer can be regarded as not paying attention to the item.
  - This approach uses items on which most attentive respondents, if not all, will provide the same response.
  - E.g., “I was born on 30<sup>th</sup> FEB”
- Implementation:
- For:
  - Easy to implement
- Against:
  - Can easily confound with IER with impression management and faking.

### Individual consistency / Inconsistent Approach / psychometric antonym / Individual reliability

- Individual consistency:
  - odd-even consistency and resampled individual reliability
  - semantic and psychometric antonyms/synonyms
  - inter-item standard deviation
  - Polytomous Guttman Errors
  - person total correlation
- Definition:
  - Consistency of a response string within an individual
  - This approach typically uses matched item pairs and compares the response on one item to the response on the other item
  - Should be highly positively/negatively correlated, if not, then IER
- Implementation:
  - Test correlation
- For:

- Against:
  - Large size of questions is preferred, Not suitable for shorter questionnaires.
  - Need to choose unique pairs of questions for the survey
- An even-odd consistency index :
  - When the even-odd consistency index cannot be applied due to unavailability of multiple subscales, the psychometric antonyms or psychometric synonyms measures are recommended.
- Psychometric antonyms :
  - The psychometric antonyms and/or the psychometric synonyms indices can serve as alternatives when multiple sub-scales are not available.
  - The psychometric antonyms index is determined by first identifying item pairs with opposite content as indicated by negative inter-item correlations
- (See Curran 2016)

### **Response pattern analysis / Maximum long string / Straight-lining / non-differentiation**

- Definition:
  - Refers to the tendency of respondents to choose the same or a very similar answer option for each item in a grid.
  - The number of times a respondent chooses the same response option consecutively.
- Implementations:
  - Use hypothesis testing (see Cornesse, 2020, page 12)
  - The same response option more than 6/9/10/14 times can be considered as IER.
- For:
- Against:
  - A cutoff score is difficult to establish, particularly for short questionnaires (Can vary from 6 - 14 consecutive responses)

### **Response time analysis (Perhaps most widely used)**

- Definition:
  - Use page time for this approach, page time is the time between the initiation and submission of each survey page online.
  - Page time with extremely small values, which implies the absence of cognitive processing
    - \* Long Time : It can be explained by the survey respondent is taking a break etc. . .
    - \* Short Time : It is very unlikely that respondents read the item content and answer items seriously when the response time is very short.
- Implementation:
  - Need to set a ‘cutoff’ time.
  - Huang et al. (2012) have suggested a cut score for response time at 2 seconds an item

- For:
- Against:
  - Difficult to create concrete rules for response time that differ from normal outlier analysis unless the status of the participant as a C/IE respondent is already known (unlikely)

## Outlier Analysis

- Definition:
  - Unusual data point relative to the remainder of a distribution.
  - Individuals who are responding without sufficient effort are likely to differ from their thoughtful counter-parts.
- Implementation:
  - Examine one value from a distribution relative to other values in that distribution.
    - \* 1.5 standard deviation from the mean
    - \* 1.5 interquartile ranges from the median
  - Use visualization techniques such as box-plot to detect the outliers.
- For:
- Against:
  - Only some subset of survey responses are considered valid.
  - May neglect the fact that there are many reasons why individuals may exist as basic statistical outliers.
- Probability values can be computed by converting Mahalanobis distance to chi-square p-values, and a critical p-value can be selected to flag suspicious respondents.
- This only holds when the assumption of multivariate normality of the item scores is met. Considering the number of response options that are used in survey data, it is unlikely that this assumption will hold.

## Mahalanobis Distance

- Definition:
  - Similar to outlier analysis, but it is a multivariate outlier technique
  - A simple extension of normal outlier analysis into multivariate space.
- Implementation:
  - Calculate the distance of points from the multivariate center.
- For:
  - Able to tell that the observation is on the outskirts of the multivariate distribution formed by responses to all items.
- Against:
  - Computationally intensive procedure
  - Rely on certain degree of normality in the data.

### **Explicit instructed response item**

- Definition:
  - Instruct the respondents to choose a specific response.
- Problem: How many items to include in a questionnaires and what is cutoff score?

### **Item non-response**

- Definition:
  - Skip questions or give a nonsubstantive answer, i.e., answer ‘don’t know’ or ‘don’t want to say’.
- Implementation:
  - if applicable, use hypothesis testing (see Cornesse, 2020, page 12)
  - Calculate the percentage of missing value for each panels. (for skip question)
- For:
- Against:

### **Midpoint selection**

- Definition:
  - A visual design experiment in our questionnaire to investigate whether respondents answer consistently across different answer scales.
- Implementation:
  - Use hypothesis testing on (see Cornesse, 2020, page 13)
  - (proportion of respondents chooses an answer option when it is located at the visual midpoint)
- For:
- Against:

### **Self Report measures**

- (Can be too stringent, complex, and computationally intensive)
- E.g. Asking: “In your opinion, should we use your data?” at the end of a questionnaire.

### **NOTES:**

- The use of any of these techniques should not be applied in a vacuum void of other techniques.
  - The known negative relationship between different metrics (Meade & Craig, 2012) means that the differential application of techniques has the potential to remove certain sets of C/IE responders while retaining others who are equally invalid. For instance, simply using a method that examines consistency of response as a positive characteristic and neglecting a method that looks at excess consistency of response as a negative characteristic is likely to remove those individuals providing highly variable invalid responses while retaining those that are functionally invariant (e.g., all ‘Strongly Agree’ responses)