CCV Methods Study - Pre-Analysis Plan

By: Nicholas Owsley and Winnie Mughogho
Date; 1 September, 2021

Overview

In recent years, remote and self-administered survey modes have increasingly become preferred for conducting research, especially where in-person data collection proves too costly or impossible. This was greatly accelerated by the pandemic, which required social distancing. With less supervision from study staff and environmental distractions, inattentive (or careless) responding could be prevalent in these self-administered surveys, posing a threat to data quality, thereby biasing results.

There is thus a need to test respondent attention on remote surveys in this context, with the aim of contributing towards methods that can improve attention in remote surveys - an area that is understudied in the measurement literature.

This study aims to develop a set of generalizable and tractable measures of respondent inattention, adopted from measures used in the literature (Meade & Craig, 2012; Berinsky et al. 2019). These will be tailored to the Kenyan population but with a generalizable framework that can be used in other contexts without much complexity. These will allow us to detect low survey attention and motivate attention, while expanding the literature on survey inattention in remote surveys.

These measures are embedded in a larger study on behavioral biases in Kenya. The design, implementation and analysis plan for this wider study is available on request.

This document describes the research questions, sample, treatments, outcome variables, and the estimation strategy of the study.

Objective and Research Questions

Objectives

- Establish a set of generalizable and simple to use measures of respondent inattention that researchers can apply to remote surveys
- Derive recommendations for survey design features to reduce inattention in remote surveys
- Test the reliability of measures of survey inattention

- Investigate whether respondent demographic and socioeconomic characteristics are predictive of survey inattention (sampling pool, income, age, gender, education)
- Quantify different levels of inattention measured by varying difficulty of attention measures

Research questions

1. How can we incentivize better attention?
    a. Does providing financial incentives increase attentiveness?
    b. Does providing an obvious attention check (Instructional Manipulation Check) at the beginning promote more attentiveness?
2. What socioeconomic and demographic factors correlate with inattention?
3. Can we distinguish between different levels of inattention by varying difficulty levels of attention measures?
4. Do the attention check measures fulfill reliability criteria?
5. What is the impact of attention measures on other survey outcomes (responses to other modules, response rates) [Secondary analysis]
6. Which parsimonious subset of attention measures performs best at identifying inattention?

Hypotheses

H1: Participants show increased attention as a result of being offered an incentive
H2: Participants show increased attention later in the survey as a result of an early obvious attention check
H3: Participants show greater attention earlier than later in the survey
H4: Different demographic groups demonstrate different levels of attention
H5: Participants demonstrate less attention in response to more difficult to answer ('costly') attention checks
H6: A subset of attention measures correlate with each other i.e they are internally reliable
H7: Higher levels of attention will correlate with higher response rates and differential answers on other survey modules

Sample Characteristics and Size Calculations

The survey will be administered online to respondents in low-income and low-middle-income settlements in Nairobi, specifically the Kibera and Kawangware constituencies, between the ages of 18 and 60 years old. We will also administer the survey to a university student sample in Nairobi. Invites will be sent to a randomly selected subset of a respondent pool constituting over 20% of the adult population within these areas. This represents a highly relevant group for future development-focused research projects planning to use remote surveys.

We will target a final sample of 2000 individuals, which is similar to that in Berinsky et al (2019) in which they robustly identify the reliability and validity of inattention measures using similar approaches. Moreover, this will allow us to detect an effect size of Cohen's d 0.177 for the primary effects of our attention-encouraging interventions with 80% power (conservatively using the 'long model' approach of Muralidharan, Romero & Wüthrich [2019] assuming interaction effects are relevant). We are thus well-powered to detect the effects found by Hauser & Schwartz (2015) and Meade & Craig (2012) applying similar interventions, where these studies detected effects of between 0.23 and 0.37 sd.

Treatments

This study aims to both detect and deter survey inattention. To this end, participants will receive a link to complete a survey that explores behavioural measures online. We will randomly assign participants to either a) be offered a 'bonus' (small financial incentive) for completing certain questions correctly, or not, and to b) complete an "instructional manipulation check" at the beginning or end of the survey. This will generate a 2x2 design, with 4 treatment cells as in the table below, which allows us to identify whether each intervention, or their interaction, affects survey inattention.

Treatment groups

	Early IMC	Late IMC
Bonus incentive	Group 1	Group 3
No bonus incentive	Group 2	Group 4

The Instructional Manipulation Check is considered an "obvious" attention check as it informs the respondent that the question is an attention check, thereby signalling that attention is being monitored. On realising that their attention is being monitored, it is anticipated that respondents would pay greater attention to the survey. Respondents will be exposed to further not so obvious attention checks, where they are not informed their attention is being monitored - described in Outcomes.

1. **Obvious attention check - Instructional Manipulation Check (IMC)**

We will randomly assign half of the participants to receive an obvious attention check (an IMC) early in the survey - in the first module. The other half will receive the same attention check only later in the survey. Apart from the IMC being a measure of attention, it is designed to signal to participants that their attention is important and is being assessed, thereby encouraging greater attention thereafter (Hauser et al. 2015).. The IMC is contained in a preference module which will be varied across respondents. Varying the preference module which will contain the other attention measures will allow us to gauge where there are differences in attention early or late in the survey.

2. **Incentive bonus**

We will randomly assign participants to receive a small financial incentive of KSH75 for getting the IMC correct. They will be informed in advance that there is a reward for getting a specific question within the survey correct, but not which one. This is also intended to sustain the respondent's attention, as they hope to get the specific question right.

Outcomes

Primary outcomes

We will have two unrelated primary research outcomes. We are interested in whether an obvious early attention measure influences responses to less obvious attention measures later in the survey. An index will therefore be created from the measures that appear later in the survey. Because these outcomes are unrelated, there will not be a need for multiple inference adjustments.

The index will be created from the following:

- Responses to infrequency questions (items on which most attentive respondents will provide nearly identical (highly skewed) responses)
- Response consistency (comparing responses to matched item pairs positioned at different points in the survey)
- Responses to Bogus Questions (questions with obviously correct responses).

Post-hoc measures of attention

Two additional methods of calculating attention have been used in the literature, which we will investigate as an additional set of comparisons to the index described above.

1. Long stringing is the tendency of respondents to choose the same answer in blocks of questions regardless of the question. The long string analysis is therefore able to detect invariant responses typically of inattentive respondents, This measure will be calculated as the longest number of consecutive similar responses to items in a module. Additionally we will calculate the Average Long String which is the average of Long String variables for the module. These variables will allow us to compare deviations from the mean for respondents.
2. Multivariate outlier analysis is used to detect patterns of responses that are statistically unlikely across items, For this we calculate the Mahalanobis distance which is the distance between the respondent's response vector and the vector of sample means (Meade & Craig, 2012)

Other data quality outcomes

The rationale behind using attention measures particularly in remote surveys is to ensure that data quality is not compromised. Besides quality data, data must be complete for the researcher to answer their research questions. We will therefore also analyse response outcomes:

- Early drop-off rates (response indicating 'yes' in consent but not completing the rest of the survey)
- Survey completion rates

Analysis & Econometric Strategy

Outcome Analysis

H1: Participants show increased attention as a result of being offered an incentive
H2: Participants show increased attention later in the survey as a result of an early obvious attention check

For our primary outcomes we analyze the effect of an early IMC and/or an incentive on attention in the survey. As such we will investigate whether these interventions are indeed associated with changes in survey attention using the models 1(a) and 1(b) below. Y is the attention measure of interest (an index of responses to infrequency/inconsistency/bogus questions). *EAC* is the obvious early attention check while *INC* represents the bonus incentive. *X* is a vector of control variables that include socioeconomic group and gender, Randomization is stratified on these controls. β_1 represents the effect of the early attention check or incentive on survey attention.

$$Y_i = \beta_0 + \beta_1 EAC_i + \beta_{2i} X_i + e_{ij} \quad \text{(1a)}$$

$$Y_i = \beta_0 + \beta_1 INC_i + \beta_{2i} X_i + e_{ij} \quad \text{(1b)}$$

A subset of respondents will receive both interventions. For those who are informed of a potential bonus in the survey, receiving the obvious early attention check might serve as a reminder that they were informed by way of incentive that their attention is being monitored. This reminder should therefore have greater effects on later survey attention. To confirm this, we will also analyze the effect of the interaction of the two interventions:

$$Y_i = \beta_0 + \beta_1 EAC_i + \beta_2 INC_i + \beta_3 EAC * INC_i + \beta_{2i} X_i + e_{ij} \quad \text{(1c)}$$

where *EAC* indicates whether the respondent got an early attention check or not. *INC* indicates whether the respondent was offered a bonus incentive. β_1 represents the effect of the early attention check on survey attention, β_2 represents the treatment effect of the bonus incentive on survey attention, β_3 represents the interaction effect of the bonus incentive and early attention check.

H3: Participants show greater attention earlier than later in the survey

To compare early to late attention, we will examine responses to the index of attention questions. The questions contained in the module require likert scale responses which we will use to create binary indicator variables (e.g. disagree/agree, true/not true). From these we shall conduct t-tests to analyse differences in means for these responses, comparing those who received the preference module early to those who received it late in the survey.

H4: Different demographic groups demonstrate different levels of attention

We will conduct the analyses for H1, H2 and H3 with subgroup analyses for the following demographic characteristics: sampling pool (university student vs low income), income, age, gender, education.

We will investigate whether demographic and socioeconomic factors such as age, gender and education correlate with attention. This knowledge would influence survey design for those who are studying different subpopulations. Y is the full attention index, made up of all attention measures, which X is the vector of explanatory variables.

$$Y_i = \beta_0 + \beta_1 X_i + e_{ij} \quad (2a)$$

Though our interventions are not explicitly targeted to increase completion rates, the change in behaviour from the interventions could influence even this. Berinsky et al. (2019) study whether a "screener", in this study referred to as IMC, has an effect on attrition. In this study we will analyse whether both the IMC and bonus incentive has an effect on response outcomes, broken down by completion and drop off rates. If the respondents are not certain whether they have answered the question that will earn them the bonus incentive, they might carry on with the survey until the end. We will run the regression below, where Y is the response outcomes, EAC is the obvious early attention check and INC is the bonus incentive.

$$Y_i = \beta_0 + \beta_1 EAC_i + e_{ij} \quad (2b)$$

$$Y_i = \beta_0 + \beta_1 INC_i + e_{ij} \quad (2c)$$

We will also study the response outcome by socioeconomic and demographic characteristics using the following model:

$$Y_i = \beta_0 + \beta_1 X_i + e_{ij} \quad (2d)$$

Where Y is the response outcome of interest, X is the vector of explanatory variables which will include age, gender, whether a student or not, and other socioeconomic factors.

H5: Participants demonstrate less attention in response to more difficult to answer (‘costly’) attention checks

Survey respondents display different levels of inattention. The difficulty of attention measures tends to determine what would constitute high or low levels of inattention. Some attention measures are quite easy to get correct, for example asking a respondent their age in different parts of a survey. Other measures have a higher cognitive demand, such as IMCs which require respondents to answer in a specific way. The pass rates of the different measures will vary with level of difficulty. We will study the passage rates of the different measures of attention, and compare these to the literature.

H6: A subset of attention measures correlate with each other i.e they are internally reliable

We will calculate Cronbach’s alpha and an inter-item correlation matrix for the index of attention measures. We will further calculate whether responses to the attention index correlate with long stringing, and Mahalanobis distance (multivariate outlier analysis).

H7: Higher levels of attention will correlate with higher response rates and differential answers on other survey modules

We will calculate whether responses to the attention index correlate with early drop-off rates, survey completion rates, long stringing, and Mahalanobis distance (multivariate outlier analysis).

We will further conduct t-tests examining the relationship between attention measures and: self-report of whether they paid attention, whether they experienced any distractions and whether they carried out the survey in a quiet environment.

Finally, we will examine the relationship of attention with predictability of performance on a variation of Tversky and Kahneman’s “Asian Disease Problem”.[1] We will verify whether in our case those flagged as inattentive, do not conform to predicted response to the problem

.

[1] Given the sensitivities of the COVID-19 pandemic, this will be contextualised into a relevant but non-current hypothetical situation about a locust invasion.

References

Berinsky, A. J., Margolis, M. F., Sances, M. W., & Warshaw, C. (2019). Using screeners to measure respondent attention on self-administered surveys: Which items and how many?. Political Science Research and Methods, 9(2), https://doi.org/10.1017/psrm.2019.53

Hauser, D. J., & Schwarz, N. (2015). It's a trap! Instructional manipulation checks prompt systematic thinking on "tricky" tasks. Sage Open, 5(2), https://doi.org/10.1177/2158244015584617

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. Psychological methods, 17(3), 437. https://doi.org/10.1037/a0028085

Muralidharan, K., Romero, M., & Wüthrich, K. (2019). Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments (Working Paper No. 26562; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w26562