Contribution ID: **197**   Type: **Symposium**

Understanding, detecting and managing careless responding in survey research.

Friday 25 July 2025 08:30 (15 minutes)

Symposium title

Understanding, detecting and managing careless responding in survey research

Coordinator

Ana Hernández and Inés Tomás

Affiliation

University of Valencia

Abstract

Careless and insufficient effort responding (C/IER) occurs when individuals do not pay sufficient attention to item content. This threatens the validity of measurement and research conclusions. This symposium presents state-of-the-art approaches to understanding, detecting, and managing C/IER in self-report data. Specifically, it examines both simulated and empirical data and focuses on different item formats, including Likert and ipsative (e.g. forced-choice) formats.

The first presentation investigates the stability of C/IER over time, addressing whether it should be considered a stable trait or a transient state. Using longitudinal data from university students, the study examines C/IER patterns identified through Instructed Response Items (IRIs) and explores whether distinct subpopulations display stable or changing response behaviors. The second presentation compares different C/IER detection methods, contrasting attention check items (i.e. IRIs) with a model-based mixture IRT approach that does not require additional items. The effectiveness of these methods and their implications for data quality are discussed. The third presentation uses simulated data to show how different strategies for handling C/IER affect the psychometric properties of scales. It compares doing nothing regarding C/IER, removing careless respondents, treating C/IER as a control variable, and using it as a moderator variable. Finally, the fourth presentation examines two strategies for addressing C/IER in ipsative data. The first strategy identifies and removes careless respondents using "person fit" statistics, while the second controls for C/IER using method factors designed for Thurstonian IRT and factor models. Together, these four studies contribute to advancing best practices in survey data quality.

Keywords

Careless responding, validity, data quality

Number of communicatios

4

Communication 1

Authors

Inés Tomás, Ana Hernández, Clara Cuevas, Vicente González-Romá

Abstract

Background. Careless responding (CR) occurs when individuals do not pay adequate attention to item content. Research has shown that CR introduces bias and compromises data quality (Podsakoff et al., 2012), highlighting the need for effective prevention and management strategies (e.g., Arthur et al., 2021; Edwards, 2019; Ward & Meade, 2022). Different methods have been proposed to detect CR, one of them being Instructed Response Items (IRIs), which direct participants to provide specific answers. Failing these items serves as an indicator of CR. The use of IRIs stands out for its simplicity, transparency, and metric properties (Kam & Chan, 2018). Despite its significance, the nature of CR remains unclear. While some researchers consider CR as a stable trait (Meade & Craig, 2012), others argue it is transient state (Maniaci & Rogge, 2014). However, little empirical evidence has clarified this distinction. A recent study by Tomás et al. (2024), conducted with a sample of adult workers who were paid for their participation in the study, identified subpopulations with distinct CR patterns, some displaying stable CR behaviors, while others exhibited changes over time.

Objectives. This study aims to deepen the understanding of CR's nature and dynamics by analyzing its patterns over time in a sample with different sociodemographic characteristics (university students) and with different contextual factors (individuals were not financially compensated for their participation). Additionally, we examine whether CR operates as a trait or state for the entire population or if distinct subpopulations exist, some for whom CR is a trait and others for whom it is a state. To detect CR, we utilize IRIs.

Methods. A total of 360 Spanish university students (71.7% women; mean age = 25.6 years, SD = 6.3) participated in the study after being offered a free face-to-face training course. We used a within-subject longitudinal design with three data collection points, spaced at 3-month intervals. Participants were first contacted during their final semester (T1), approximately one month before graduation, followed by assessments nine months post-graduation (T2), and four months after T2 (T3). The trajectory of CR over time was modeled using latent growth modeling (LGM), and latent class growth analysis (LCGA) in Mplus.

Results and Conclusions. The results aligned with previous research (e.g., Tomás et al., 2024): while CR exhibited a stable response pattern over time at the population level, distinct subpopulations emerged, each displaying different CR trajectories. Notably, the subgroups identified in this study differed from those found by Tomás et al. (2024). In this study, three distinct subpopulations emerged: a relatively stable group (careful individuals) and two groups whose inattentiveness increased over time (one initially careful but becoming less attentive and another already careless that became even more inattentive). These findings contribute to the understanding of CR's nature and dynamics, highlighting the role of personal factors (e.g., age) and contextual factors (e.g., participation compensation) in shaping CR patterns over time.

Keywords

Communication 2

Detecting careless and insufficient effort responding: A comparison of attention check and model-based approaches

Authors

Esther Ulitzsch, Ana Hernández, Inés Tomás, Clara Cuevas

Affiliation

University of Oslo and University of Valencia

Abstract

Background. Careless and insufficient effort responding (C/IER) on self-report measures produces responses that fail to accurately reflect the trait being measured, posing a major threat to the quality and validity of survey data. While detecting C/IER is vital to ensure validity of conclusions drawn from self-report data, it is a non-trivial endeavor, with each detection method involving distinct assumptions and limitations.

Objectives. This study compares two prominent approaches for C/IER identification and adjustment based on respondent behavior: (1) attention check items, which offer clear interpretability but require careful and parsimonious administration, limiting their ability to monitor C/IER comprehensively, and (2) a model-based mixture IRT approach, which avoids the need for additional items but relies on strong assumptions about respondent behavior.

Methods. Using data from five scales of a job quality survey completed by 707 respondents, we fitted an explanatory mixture IRT approach by means of R.

Results and Conclusions. We observed strong alignment between the two approaches: respondents identified as less attentive by one method were similarly flagged by the other. Overall, both approaches suggested that C/IER remained relatively stable throughout the course of the questionnaire. However, single attention check items consistently indicated substantially lower levels of C/IER at multiple points throughout the questionnaire compared to the scale-level C/IER rates implied by the model-based approach. Both methods had comparable impacts on adjusted psychometric properties. While correlations between latent constructs did not differ markedly from their unadjusted counterparts, adjusted trait estimates were less reliable, especially when obtained using the model-based approach, reflecting greater uncertainty in respondents' trait levels. Implications for C/IER identification and adjustment are discussed, arguing for a triangulation of different approaches.

Keywords

Communication 3

Detecting and managing careless and insufficient effort responding: A simulation approach.

Authors

Clara Cuevas, Inés Tomás, Ana Hernández

Affiliation

University of Valencia

Abstract

Background. Careless and insufficient effort responding (C/IER) occurs when respondents fail to give sufficient attention to item content, which leads to poor-quality data (Podsakoff et al., 2012). There are several methods to detect this phenomenon, one being Instructed Response Items (IRI), valued for its simplicity, robust metric properties, and ability to identify different C/IER patterns (Kam & Chan, 2018). While detecting C/IER is a crucial first step, deciding how to address this phenomenon once identified is equally important, as this choice can determine the extent of its impact on data quality.

Objectives. This study compares four strategies for managing C/IER and their impact on the psychometric properties of questionnaires, specifically reliability and validity evidence based on the internal structure: (1) using the total sample without adjustments, (2) excluding careless respondents to create a "clean" sample, (3) retaining the total sample while treating C/IER as a control variable, and (4) retaining the total sample while treating C/IER as a moderating variable.

Methods. We use simulated data based on the Big Five Questionnaire (Caprara et al., 1993) and the Maslach Burnout Inventory (Maslach & Jackson, 1981). A total of 180 conditions are manipulated, with variations in variables such as Severity of C/IER (25%, 50%, 75%, 100%), Percentage of C/IER (0%, 8%, 24%), or Sample Size (150, 300, 700). For each condition, 100 replications are run.

Expected results and Conclusions. Based on previous studies with empirical data (Tomás et al., 2023), we anticipate that using C/IER as a moderating variable (4) will be the most effective strategy. In contrast, using the total sample without adjustments (1) will likely be the least effective, given that C/IER is ignored. Regarding the exclusion of careless respondents (2), we anticipate a reduction in statistical power and its subsequent impact on the psychometric properties. As for (3) using C/IER as a control variable, based on previous empirical research examining its impact on questionnaire psychometric properties (Tomás et al., 2023) and substantive research models results (Tomás et al., 2025), we expect this strategy to be a less effective approach for addressing CR. We will provide recommendations for managing C/IER, helping to mitigate its impact on data quality in applied research.

Keywords

Careless responding, Montecarlo, Psychometric properties

Communication 4

Detecting careless responding in ipsative data

Authors

Anna Brown

Affiliation

University of Kent

Abstract

Background. To prevent response styles associated with the use of rating scales, test items may be presented in so-called ipsative (or relative to self) formats including popular 'forced choice', and also 'graded preferences' or 'proportions-of-total'. Like any other questionnaires, ipsative questionnaires can be subject to careless responding when respondents are not sufficiently motivated to give their full attention to the questions. However, detecting such responding can be more challenging than when using Likert scales because ipsative response formats usually involve comparisons between items measuring different traits and their modelling is inherently multidimensional. Moreover, the comparative nature of ipsative responses makes challenging the use of a method factor (latent variable) to control careless responding.
Objectives. This presentation will describe and evaluate two alternative strategies for dealing with careless responses in ipsative data: (1) identifying (and ultimately removing from the sample) careless responders using "person fit" indices designed for ipsative formats; and (2) controlling for careless responding using method factors specifically designed for Thurstonian IRT and factor models (Brown & Maydeu-Olivares, 2012).
Methods. The two approaches are illustrated on a sample of N=504 paid Prolific participants in a trial of the Leadership Styles Questionnaire (LSQ) measuring 24 personal styles with 88 multidimensional graded triplets. Under Approach 1, two "person fit" indices were computed for each respondent. The first index summarized the discrepancies between a person's observed responses and responses expected under the fitted Thurstonian measurement model, thus resembling the lco index (Ferrando, 2010). The second index summarized the concordance between a person's observed and expected responses by computing a correlation coefficient between them. Under Approach 2, a random intercept was added to the Thurstonian measurement model to control carelessness expressed as overusing one rating scale category.
Results and Conclusions. The concordance index had a median of 0.572 and a long left tail, identifying at least 8% of aberrant responders. The discrepancy index had a median of 0.820 and a long right tail, again identifying at least 8% of aberrant responders. The Thurstonian model with the random intercept factor fitted better than the baseline model (SRMR were .058 and .075, respectively), and the random intercept explained between 1% and 2% in the variances of observed responses. However, at the individual level the discrepancy, concordance and random intercept agreed only for careful responders. For careless responders, each index provided unique information about the nature of carelessness. We conclude with recommendations for the use of such indices in practice.

Keywords

Careless, ipsative data, person-fit, Thurstonian-models

Primary authors: HERNÁNDEZ, Ana (University of Valencia); BROWN, Anna; CUEVAS UREÑA, Clara; ULITZSCH, Esther (University of Oslo); Dr TOMÁS, Inés (University of Valencia); GONZÁLEZ-ROMÁ, Vicente

Presenters: HERNÁNDEZ, Ana (University of Valencia); BROWN, Anna; CUEVAS UREÑA, Clara; ULITZSCH, Esther (University of Oslo); Dr TOMÁS, Inés (University of Valencia)

Session Classification: Symposium : "Understanding, detecting and managing careless responding in survey research."

Track Classification: Measurement: Measurement