

Identifying the Onset of Careless Responding: A Machine Learning Approach

M. Welz¹ and A. Alfons¹

¹Erasmus University Rotterdam, Department of Econometrics, Burgemeester Oudlaan 50, 3062PA Rotterdam, Netherlands

Questionnaires in the behavioral and organizational sciences tend to be lengthy: survey measures comprising hundreds of items are the norm rather than the exception. However, the literature suggests that the longer a questionnaire takes, the higher the probability that participants lose interest and start responding carelessly. Consequently, in long surveys a large number of participants may engage in careless responding, posing a major threat to internal validity. We propose a novel method to identify the onset of careless responding (or an absence thereof) for each participant. Specifically, our method is based on combined measurements of up to three dimensions in which carelessness may manifest: inconsistency, invariability, and fast responding. We measure the first dimension (inconsistency) in terms of reconstruction errors of the observed responses from auto-associative neural networks. For the second dimension (invariability), we propose a novel measure aimed at capturing identical responses or recurring response patterns. The third dimension (fast responding) can be taken into account via per-item or per-page response times. Since a structural break in either dimension is potentially indicative of carelessness, our method searches for a changepoint along the three dimensions (or a subset of those dimensions). Our method is highly flexible, based on machine learning, and provides statistical guarantees on its performance. In simulation experiments and an empirical example, we find that it achieves high reliability in correctly identifying carelessness onset, discriminates well between careless and attentive respondents, and can capture a wide variety of careless response styles, even in datasets with an overwhelming presence of carelessness. Finally, we provide freely available software in R to enhance accessibility and adoption by empirical researchers.

[1] M. Welz and A. Alfons, “I don’t care anymore: Identifying the onset of careless responding,” 2023. arXiv:2303.07167.