Journal of Operations Management xxx (2017) 1–8

Contents lists available at ScienceDirect

Journal of Operations Management

journal homepage: www.elsevier.com/locate/jom

Technical note

Attention by design: Using attention checks to detect inattentive respondents and improve data quality

James D. Abbey [a, *], Margaret G. Meloy [b]

[a] Department of Information & Operations Management, Mays Business School, Texas A&M University, 320 Wehner Building, 4217 TAMU, College Station, TX 77843-4217, United States

[b] Calvin E. and Pamela T. Zimmerman University Endowed Fellow, Department of Marketing, Smeal College of Business, The Pennsylvania State University, 444 Business Building, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 3 August 2016
Received in revised form
2 November 2016
Accepted 7 June 2017
Available online xxx
Handling Editor: Mikko Ketokivi

Keywords:
Data validation
Attention checks
Manipulation checks
Response validation

ABSTRACT

This paper examines attention checks and manipulation validations to detect inattentive respondents in primary empirical data collection. These prima facie attention checks range from the simple such as reverse scaling first proposed a century ago to more recent and involved methods such as evaluating response patterns and timed responses via online data capture tools. The attention check validations also range from easily implemented mechanisms such as automatic detection through directed queries to highly intensive investigation of responses by the researcher. The latter has the potential to introduce inadvertent researcher bias as the researcher's judgment may impact the interpretation of the data. The empirical findings of the present work reveal that construct and scale validations show consistently significant improvement in the fit statistics—a finding of great use for researchers working predominantly with scales and constructs for their empirical models. However, based on the rudimentary experimental models employed in the analysis, attention checks generally do not show a consistent, systematic improvement in the significance of test statistics for experimental manipulations. This latter result indicates that, by their very nature, attention checks may trigger an inherent trade-off between loss of sample subjects—lowered power and increased Type II error—and the potential of capitalizing on chance alone—the possibility that the previously significant results were in fact the result of Type I error. The analysis also shows that the attrition rates due to attention checks—upwards of 70% in some observed samples—are far larger than typically assumed. Such loss rates raise the specter that studies not validating attention may inadvertently increase their Type I error rate. The manuscript provides general guidelines for various attention checks, discusses the psychological nuances of the methods, and highlights the delicate balance among incentive alignment, monetary compensation, and the subsequently triggered mood of respondents.

> *"To avoid any space error or any tendency to a stereotyped response, it seems desirable to have the different statements so worded that about one-half of them have one end of the attitude continuum corresponding to the left or upper part of the reaction alternatives … These two kinds of statements ought to be distributed throughout the attitude test in a chance or haphazard manner."* —Rensis Likert (1932)

1. Introduction

Generations of researchers have struggled with the "stereotyped response"—a response that does not accurately represent subjects' attitudes (Likert, 1932). The challenge for researchers is in distinguishing between a true attitude, belief, or behavioral response versus a stereotyped response without introducing bias from the researchers themselves. With the ready availability of online resources that facilitate primary data collection, the issue of response accuracy is particularly relevant. To assess such data quality issues, this paper addresses the following: how far have we come in identifying stereotyped responses and what methods can be effectively used to address response validity at a fundamental level without introducing bias? The answers are not as simple as they

* Corresponding author.
E-mail addresses: jabbey@mays.tamu.edu (J.D. Abbey), mmeloy@psu.edu (M.G. Meloy).

http://dx.doi.org/10.1016/j.jom.2017.06.001
0272-6963/

might seem. While much progress has been made and many papers written on the topic across empirical disciplines in the social sciences, the need remains for continued research into evolving methods for establishing the validity and quality of primary empirical data (Meade and Craig, 2012).

Though theoretical expansions of mechanisms to detect and isolate inattentive respondents have seen a resurgence in the literature over the past decade, the quantifiable effects of attention checks—questions designed explicitly to detect inattentive responses through direct queries of attention or through questions designed to catch inattentive respondents—have rarely been reported (Van Dam, Earleywine, and Borders, 2010; Meade and Craig, 2012). In behavioral operations, a common alternative to attention checks has been to prompt attention through the use of incentive alignment mechanisms in keeping with the rich tradition of behavioral economics literature (Bendoly et al., 2006). Such incentive alignment mechanisms are not a catch-all for inattentiveness, however, and may have unexpected repercussions as discussed below (Meloy et al., 2006; Angner and Loewenstein, 2012). Further, though lab experiments using student subjects remain a dominant method in behavioral operations (Bendoly and Schultz, 2006; Katok, 2011; Siemsen, 2011), the behavioral operations field has recently seen an expansion of the research methods used as highlighted by the recent state of the literature noted by Croson et al. (2013). Such validation of results through methodological mechanisms, sub-sample analysis, and replication-based triangulation is a distinctly welcome and necessary step in the scientific process (McGrath, 1981; Flynn et al., 1990; Malhotra and Grover, 1998). Yet, even with this much needed surge in the variety of methods employed to address the significant complexity of behavioral issues, some recent advancements in the use of relatively simple and easy to implement attention checks have yet to appear in the majority of the operations management literature.

Fortunately, the use of replicate robustness studies—replication of some or all of the original experiment within a single manuscript—has been a trend in recent behavioral operations research (for some recent examples, see Kremer et al., 2011; Kremer et al., 2015). Of course, replication of empirical studies to alleviate Type I error concerns is quite common in the sister disciplines of psychology, marketing, management, and other related fields of behavioral study (McGrath, 1981; Rosenthal, 1990; Peterson, 2001). Further, some recent works do intensive post-hoc examination of manipulation effectiveness—a more involved concept related to the manipulation checks discussed below (for excellent examples, see Bendoly, 2013; Bendoly et al., 2014). That said, robustness validation through experimental replication does not preclude the use of attention checks. Both the original experiment and robustness validations benefit from the prima facie data validation offered by attention checks. This inherent validation mechanism strengthens the case that both the original study and robustness replicate—through a new sample or sub-sampling—do indeed isolate the measures and effects of interest. Additionally, the use of attention checks has the potential to mitigate Type II error in a robustness replicate. In other words, the replicate may have failed to detect the original effect due not to a lack of effect but the often smaller sample size (i.e., lower inherent power) or random subject pool quality issues that could be detected by incorporating attention checks into the original design.

In addition, attention checks can be extraordinarily valuable for study replicates both in terms of budgetary requirements and inherent data quality. In terms of budgetary savings, studies can be designed to withhold payment or incentives for respondents who fail a sufficient number of attention checks (i.e., failure of all or some subset of attention checks). With loss rates such as those seen in the discussion below, the savings could be considerable, particularly when payouts in some recently published studies average over $80 per subject (Kremer et al., 2015). In terms of data quality, the results below demonstrate that attention checks can yield significant benefits, particularly for measurement models related to the focal manipulations.

As will become clear, there is no *single* solution to the complex issue of identifying inattentive respondents. Recent evolutions in data collection, particularly the advent of easy, fast, and efficient electronic data collection, lead to the increased need for carefully designed studies that do exactly as Likert suggested—find means to establish data validation through careful measure design. These prima facie data validations to capture respondent inattentiveness range from the simple, such reverse scaling as noted nearly a century ago (Hambleton et al., 1991; Nardi, 2003; Hughes, 2009), to the complex, such as evaluating response patterns and timed responses via online data capture tools (Huang et al., 2012). Moreover, the choice of method can range from easily implemented mechanisms, such as automatic detection (Oppenheimer et al., 2009), to highly intensive investigation of responses by the researcher. The concern with the latter is that the researcher's judgment will impact the interpretation of the data (i.e., potential researcher bias). The methodological investigation that follows evaluates such mechanisms and examines recent innovations to assess and detect inattentive respondents.[1]

In light of the discussion above, this manuscript focuses on the effects of attention checks in terms of sample loss rates, statistical impacts on measurement models, and the influence of attention checks on observed statistical significance in experimental models. In particular, the manuscript highlights the far higher than typically reported loss rates among respondents. Further, the manuscript demonstrates the critical need for use of attention checks for measurement models, though the effects on experimental models are less clear in their criticality. Specifically, this manuscript provides a snapshot of the evolving methods surrounding the complex and difficult topic of validating primary data collection with a particular focus on data collection from undergraduate student populations. To this end, the following discussion provides general guidelines, a methodological overview that includes examples of various forms of attention checks, discusses the advantages and disadvantages of different methods, highlights the objectivity of implementation, and underscores the potential outcomes from their use. The work also provides an initial exploration of attention checks empirically by examining the frequency of data loss from failed attention checks and the general impact of such losses on statistical results. In the end, as the analysis demonstrates, the greatest benefits appear to accrue for researchers involved predominantly in scale and construct-based research (e.g., survey research).

2. Historical attempts to detect and prevent inattentive responses

As noted in the previous section, ensuring high quality data in experimental and field studies has been an issue that has plagued researchers for centuries. In particular, researchers would like to ensure that participants in their research studies are responding in ways that accurately reflect their true beliefs, attitudes, and behaviors. Further, such researchers would like to identify those individuals who are failing to follow instructions or who are not devoting attention to the instructed task (Barnette, 1999). These individuals often add unnecessary noise to the research findings,

[1] We extend our deepest thanks to numerous researchers who generously donated their data from more than 30 studies for the examination of these ever-evolving attention check mechanisms.

Please cite this article in press as: Abbey, J.D., Meloy, M.G., Attention by design: Using attention checks to detect inattentive respondents and improve data quality, Journal of Operations Management (2017), http://dx.doi.org/10.1016/j.jom.2017.06.001