Journal of Research in Personality 48 (2014) 61–83

Contents lists available at ScienceDirect

Journal of Research in Personality

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

Caring about carelessness: Participant inattention and its effects on research

Michael R. Maniaci, Ronald D. Rogge *

Department of Clinical and Social Sciences in Psychology, University of Rochester, United States

ARTICLE INFO

Article history:
Available online 7 October 2013

Keywords:
Attentiveness
Random responding
Inattentive responding
Careless responding
Validity scales

ABSTRACT

The current studies examined the adverse effects of inattentive responding on compliance with study tasks, data quality, correlational analyses, experimental manipulations, and statistical power. Results suggested that 3–9% of respondents engaged in highly inattentive responding, forming latent classes consistent with prior work that converged across existing indices (e.g., long-string index, multivariate outliers, even–odd consistency, psychometric synonyms and antonyms) and new measures of inattention (the Attentive Responding Scale and the Directed Questions Scale). Inattentive respondents provided self-report data of markedly poorer quality, sufficient to obscure meaningful regression results as well as the effects of experimental manipulations. Screening out inattentive respondents improved statistical power, helping to mitigate the notable drops in power and estimated effect sizes caused by inattention.

1. Introduction

An oft-neglected issue underlying much research is that not all respondents pay sufficient attention when completing self-report measures. Such responding could introduce error into a dataset, potentially decreasing power and obscuring results. Inattention is sometimes blatant and easy to address (e.g., removing participants from analyses if they exhibit suspiciously fast reaction times or below-chance performance on a task). This logic, however, is only rarely applied to inattentive responding on self-report scales as this form of inattention is more subtle and therefore more difficult to measure. The current studies were designed to identify excessive inattention using a multi-method approach, exploring its impact on: (1) compliance with common study tasks, (2) quality of self-report data, (3) correlational and experimental analyses and (4) statistical power, in order to determine the potential scope of this problem and the degree to which addressing it might improve statistical analyses. Estimated rates of inattention in the existing literature have varied widely, from 3% to 46% of respondents (e.g., Berry et al., 1992; Johnson, 2005; Meade & Craig, 2012; Oppenheimer, Meyvis, & Davidenko, 2009). In part, this wide range of estimates is due to a lack of clarity on how best to measure inattentive responding and on what thresholds correspond to unacceptably error-ridden data. The present research used indicators of non-compliance, data quality, and statistical power as criteria for comparing methods of measuring inattention and establishing concrete, practical thresholds for researchers to use to screen their data.

1.1. Forms of non-compliance

It is common for a small portion of participants to exhibit poor attention and effort in research. For example, subjects with excessively short reaction time latencies on implicit measures like the Implicit Association Test (IAT; Greenwald, Nosek, & Banaji, 2003) are routinely excluded from analyses. Although such practices are common in research using reaction time paradigms and experimental manipulations, this logic has not typically been extended to research utilizing self-report methods. This discrepancy is likely not because researchers believe that experimental manipulations or reaction time measures are more prone to non-compliance than self-report measures; rather, it is simply easier to identify non-compliance on such tasks. As researchers typically do not screen for inattention on self-report scales, the prevalence and impact of such problematic responding is largely unknown.

1.2. Inattentive responding as a distinct construct

Although a number of constructs are occasionally lumped together under the heading of *validity scales* (e.g., socially desirable responding, faking good, faking bad, random responding), the present research focuses on a specific form of invalidity: inattention

* Corresponding author. Address: Department of Clinical and Social Sciences in Psychology, University of Rochester, Meliora Hall, RC Box 270266, Rochester, NY 14627-0266, United States.
E-mail address: rogge@psych.rochester.edu (R.D. Rogge).

0092-6566/$ - see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.jrp.2013.09.008

when completing self-report measures. This form of inattention is distinct from other types of invalidity. For instance, the response sets of faking good, faking bad and social desirability imply a motivation to present oneself in a particular manner. Ironically, these forms of invalidity may be *negatively* related to inattentive responding because presenting oneself in a particular manner requires carefully attending to questions (Meade & Craig, 2012). In contrast, inattentive responding corresponds to a lack of motivation to present oneself in a certain manner, and should therefore contribute little more than error variance to analyses. Extreme levels of inattention conceptualized in this manner are consistent with the extremely inattentive latent class identified by Meade and Craig (2012) as comprising approximately 9% of an undergraduate sample, and with what Nichols, Greene, and Schmolck (1989) called "content nonresponsivity." Although inattention could be correlated with individual differences, in the current study we view inattentive responding as a proximal behavior enacted during the completion of research studies. We therefore conceptualize it as more of a transitory (state) phenomenon, allowing for the possibility that the same individual might provide high levels of attention in one study (e.g., a short and particularly interesting study) but insufficient levels of attention in other studies.

1.3. Approaches to measuring inattention

1.3.1. Infrequency and inconsistency scales

Much of the work examining inattentive responding on self-report measures has been conducted in the development of clinical assessment batteries like the Personality Assessment Inventory (PAI; Morey, 1991) and the Minnesota Multiphasic Personality Inventory (MMPI; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). The length of these inventories (typically containing several hundred items) demands high levels of sustained attention from respondents, necessitating the development of scales designed to identify problematic responding. Two types of validity scales within these clinical batteries (infrequency and inconsistency scales) assess inattentive responding, with a focus on identifying extreme (and therefore problematic) levels. Infrequency scales (e.g., the infrequency scale of the PAI, the "bogus" items of Meade & Craig, 2012) are made up of items that elicit nearly identical (highly skewed) responses from most respondents (e.g., "I have been to every country in the world"). Respondents receive higher scores on infrequency scales for each increasingly unlikely response across the set of items, and cut-scores are used to identify inattentive responding excessive enough to yield an invalid protocol. Inconsistency scales (e.g., the VRIN scale of the MMPI, the inconsistency scale of the PAI) are made up of pairs of items with nearly identical content that are presented in opposing halves of a survey (e.g., "I am an active person" paired with "I have an active lifestyle"). Absolute differences in responses are summed across the item pairs so that higher scores reflect more inconsistent responding. Thus, although these scales make use of self-report items, they do not ask subjects to report on their own levels of attention but instead use their responding behavior on a set of heterogeneous items to indirectly assess their attentiveness to item content. Such scales are effective at distinguishing randomly generated data from actual data (e.g., Bruehl, Lofland, Sherman, & Carlson, 1998; Pinsoneault, 2005), but have only rarely been implemented outside of the clinical instruments for which they were designed (e.g., Saavedra, Chapman, & Rogge, 2010).

1.3.2. Instructional manipulation check

More recently, Oppenheimer and colleagues (2009) developed the instructional manipulation check (IMC), a single item measure presented as a separate page in an online survey that uses critical instructions embedded at the end of a lengthy paragraph to assess participants' attentiveness to instructions. Although the IMC moderated the effectiveness of text-based manipulations, its usefulness is minimized by the fact that it only measures one form of inattention (skipping instructions) which is relatively common and therefore identifies a high proportion of participants (35–45%) as inattentive. As Oppenheimer and colleagues note, eliminating that many participants could potentially reduce power and bias results, leading them to suggest using the IMC as an intervention to encourage attentiveness rather than as a measure of inattention. Despite that suggestion, published studies using the IMC have simply excluded high proportions of inattentive respondents (e.g., Simmons & Nelson, 2006, Study 12). This highlights the need for a measure of attention with greater variability and specificity, potentially allowing for a smaller proportion of participants to be identified as excessively inattentive.

1.3.3. Post hoc indices

In addition to adding measures in order to identify highly inattentive respondents, researchers can also calculate post hoc indices of inattention using virtually any body of self-report items after data have been collected. Meade and Craig (2012) examined the convergence of several such indices, including highly correlated item-pairs (psychometric synonyms and antonyms assessing the consistency of responding), even–odd consistency (split-half reliabilities measured within respondents across scales), multivariate outlier distances (assessing statistically unlikely response patterns), long string analyses (measuring the tendency to choose identical answers in blocks of items), and time spent on the survey. Latent profile analyses of these indices along with an infrequency scale identified two main types of inattention: one reflecting more general inattentive responding and another marked by subjects frequently selecting the same answer for entire blocks of questions and consequently completing the survey in suspiciously short periods of time. Their analyses suggested an overall frequency of 10–12% of inattentive respondents in a sample of undergraduates.

1.4. Effects of inattention

Careless or inattentive responding might act as a source of measurement error that could obscure meaningful results. Identifying and removing inattentive respondents before data analysis could therefore offer a relatively easy method of decreasing error variance and increasing statistical power in research using self-report measures. Consistent with this assertion, Oppenheimer et al. (2009) found that previously reported experimental results with two different manipulations did not replicate in a group of participants identified as inattentive, although they did replicate among attentive participants. Similarly, inattentive responding can adversely affect correlational and factor analyses (Johnson, 2005; Meade & Craig, 2012; Woods, 2006), moderating findings even to the point of generating spurious results.

1.5. Addressing challenges to validity measures

In contrast to these findings, Piedmont, McCrae, Riemann, and Angleitner (2000) challenged the utility of validity scales as a set, demonstrating that scores on a diverse array of 13 different validity measures (including measures of inattention as well as distinct constructs, such as social desirability, faking good, faking bad) failed to moderate substantive results when formed into a heterogeneous composite. The methodological decisions underlying their non-significant results highlight four fundamental elements of the current approach. First, the current work focused exclusively on inattentive responding, as it is not clear that all forms of invalid responding would have equivalent effects on data quality and statistical power. Consequently, forming composites of disparate