Computers in Human Behavior 48 (2015) 554–568

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Using virtual presence and survey instructions to minimize careless responding on Internet-based surveys

M.K. Ward [a,*], Samuel B. Pond III [b]

[a] Department of Psychology, Campus Box 7801, North Carolina State University, Raleigh, NC 27695, United States
[b] Department of Psychology, North Carolina State University, Raleigh, NC 27695, United States

ARTICLE INFO

Article history:

Keywords:
Online survey
Survey design
Careless responding
Insufficient effort responding
Virtual presence
Instructions

ABSTRACT

Internet-based survey data inform knowledge creation in research and justify work activities in organizations. While there are advantages to online surveys, this mode of administration comes with its own set of challenges. Survey respondents may engage in careless responding (i.e. insufficient effort responding or satisficing) by intentionally or unintentionally responding in a manner that does not accurately reflect their true sentiments. Careless responding can create psychometric problems even after correctly removing careless respondents (i.e. mischievous responders). This study aimed to improve survey methodology by preventing careless responding. Using a 3×3 between-subjects experimental design, we manipulated both virtual presence (none, animated shape, and virtual human) and type of instructions (anonymous, warning, and feedback). Indicators of careless responding were the dependent variables. Results showed that beyond characteristics of survey items, survey design elements can prevent careless responding. The effects of interventions differed by type of careless responding. Instructions, and the interaction of instructions and virtual presence significantly reduced careless responding, but not virtual presence alone. Virtual human presence increased the salience of instructions. Although currently effective, instructions warning of punitive consequences may create difficulty in recruiting participants. Future research should continue investigating non-aversive ways to prevent careless responding on Internet-based surveys.

1. Introduction

Advances in technology have spurred the extensive use of Internet-based surveys. Data from Internet-based surveys support knowledge creation in research and inform applied work in many organizations (e.g., Acquavita, 2009; Anderson, 2010; Berta, 2006; Marx, 2012; Mervis, 2007; Patrick, 2012). While there are decided advantages to Internet-based surveys, this mode of administration comes with its own set of challenges. For example, respondents may ignore important parts of the survey, submit multiple times, and may often exhibit careless responding (CR; Barak & English, 2002; Barge & Gehlbach, 2012; Berry et al., 1992; Curran, Kotrba, & Denison, 2010; Hardré, Crowson, & Xie, 2012; Johnson, 2005; Meade & Craig, 2012; Robinson-Cimpian, 2014). That is, either intentionally or unintentionally respondents may answer survey items in a manner that does not accurately reflects their true sentiments. Understanding and manipulating features of Internet-based surveys that encourage attentiveness may: decrease CR, provide cleaner datasets to support conclusions, and promote better theory development and application. The primary aim of this study is to examine how certain features of survey design can prevent CR by increasing attentiveness among respondents.

2. Psychometric problems associated with Internet-based surveying

CR occurs when a person responds to a survey item in a way that reflects inaccuracy rather than that person's true sentiment. The person may or may not take into account the content of the survey item. Nichols, Greene, and Schmolck (1989) describe CR as manifesting itself in one of two ways. Content-responsive faking occurs when responses relate to the content of items and exhibit some level of inaccuracy. Respondents may intentionally engage in content-responsive faking or unintentionally engage in socially desirable responding (Paulhus, 1984). An alternative form of CR does not relate to item content whatsoever. Content nonresponsivity occurs when one responds to a survey without consideration of item content. Random responding would fall under this

* Corresponding author.
E-mail addresses: mkward@ncsu.edu (M.K. Ward), sbpond@ncsu.ed (S.B. Pond III).

http://dx.doi.org/10.1016/j.chb.2015.01.070
0747-5632/

category, although some response patterns that do not relate to item content are not necessarily random. Some respondents who display content nonresponsivity, for instance, may choose the same response option for multiple items in a row. Such a response pattern is not random, though it is just as reflective of CR as a random response pattern (Johnson, 2005; Meade & Craig, 2012).

2.1. How to detect CR

There are two general ways to identify CR. First, when researchers build a survey they can add special survey items that indicate CR. Some examples of this include self-report items asking the respondents to rate their engagement during the survey, asking if their survey data is of sufficient quality for research use, and asking respondents to select a specific response option. Instructed-response items are particularly effective at identifying CR given that they have one objectively correct response option. Self-report items that ask respondents if their data are adequate for research use are brief and can be effective at screening out CR (Meade & Craig, 2012). The present study used two types of survey items as CR indicators, namely instructed-response items and self-report items asking respondents if their data were adequate for research use.

The second way to identify CR is to calculate values from the survey performance information of each respondent. Survey performance information includes how long it took to complete the survey and what respondents reported in the raw survey data. Effective CR indicators that researchers can derive from survey data include: Mahalanobis distance (Ehlers, Greene-Shortridge, Weekley, & Zaiack, 2009), *Even–Odd consistency* (Jackson, 1977), and response patterns such as the same response option selected consecutively, i.e., *LongString* (Johnson, 2005). Mahalanobis distance is a descriptive statistic that indicates the distance of cases from the means of predictor variables (Field & Miles, 2010). Ehlers et al. (2009) showed that using Mahalanobis distance values to identify extreme values on surveys could indicate CR because those values by their nature are unlikely. The Even–Odd consistency measure splits the odd items from the even items on unidimensional subscales from the survey. According to the Even–Odd consistency measure, small within-person correlations across the subsets of even and odd items would indicate CR. The LongString CR indicator is the maximum number of consecutive items with the same answer choice selected (Johnson, 2005). Other CR indicators include measures of rushing, skipping, early termination, and sensitivity-analysis (Barge & Gehlbach, 2012; Robinson-Cimpian, 2014). Although many more CR indicators exist, prior research supports the efficacy of Mahalanobis distance, Even–Odd consistency, and *Maximum LongString* beyond that of other CR indicators (Meade & Craig, 2012).

Adequately identifying CR necessitates using more than one CR indicator because research suggests CR is a multidimensional construct that can manifest itself in different ways (Meade & Craig, 2012). For example, self-reported CR shows low to moderate correlations with other indicators of CR. This means that self-report items may not sufficiently indicate CR when used alone (Meade & Craig, 2012). Therefore, the current study used several of the previously mentioned CR indicators because they measure CR more effectively and comprehensively than any single CR indicator. Perhaps more importantly, the current study also investigated two methods of obviating CR.

2.2. Importance of finding solutions to CR

Finding a solution to CR should be a priority to researchers and survey administrators for two reasons. First, both research findings and data-driven decision making heavily rely on survey data to justify decisions and actions based on those decisions. Evidence-based conclusions rely on clean datasets assimilated through research. Although CR is a longstanding issue in survey methodology, few studies that employ Internet-based survey methodology examine the quality of responses to surveys beyond typical data cleaning procedures. Recent research is starting to fill this gap by investigating methods beyond typical data-cleaning procedures that identify careless respondents (e.g. Barge & Gehlbach, 2012; Gehlbach & Barge, 2012; Meade & Craig, 2012; Robinson-Cimpian, 2014).

People who use survey data should address CR because estimates of the prevalence of CR range from 3.5% to 60% of the samples (Berry et al., 1992; Curran et al., 2010; Johnson, 2005). In a job application survey, Berry et al. (1992) found that one or more items reflected CR in a *majority* of respondents. Berry et al. (1992) identified the careless responses in only part of the survey. Therefore, respondent inattentiveness may not affect all survey items equally. In a voluntary subject pool, Johnson (2005) found a 3.5% base rate of CR. Depending on the criteria by which they defined inattentive responding to a job satisfaction questionnaire, Curran et al. (2010) estimated the rates of CR to be approximately 5%, 20%, or 50% among a large sample of employee respondents. In sum, the literature warrants two observations: (a) the prevalence of CR depends on the indices used to estimate it, and (b) CR is evident in many datasets derived from Internet-based surveys.

The second reason why people who use survey data should address CR, is that CR can lead to psychometric problems. The frequent presence of CR in data is problematic to scale development (Schmitt & Stults, 1985; Woods, 2006) and factor analysis (Huang, Curran, Keeney, Poposki, & DeShon, 2012; Woods, 2006) that often underlie theoretical development and exploratory studies. CR can distort correlations and internal consistency reliability estimates (Meade & Craig, 2012). For these reasons, prudent researchers in all domains of the social sciences need to address CR in their data. Additional data-cleaning procedures can verify the assumption of sufficient data quality in survey responses. Researchers must identify CR and discover ways to eliminate its effects in order to draw sound conclusions based on survey data.

One approach to addressing issues resulting from CR would be for researchers to omit data from certain respondents. To do this, each respondent would receive values on CR indicators and if the data from any of the respondents returned values beyond cutoff scores, then researchers would exclude their data from further analysis (Tabachnick & Fidell, 2013). The assumption made here is that removing respondent data is preferable to keeping low-quality data. However, correctly extracting data contributed by careless respondents is a limited solution to CR issues.

Removing respondents' data is a reactive approach that, even if perfectly executed, can lead to a host of other problems. For example, it necessarily reduces sample sizes in a non-random way. Such removal can artificially shape the sample distribution. In turn, this limits the external validity of results and narrows implications. Put another way, removing respondents negates random sampling and potentially decreases the generalizability of survey findings. Therefore, it is imperative to find ways of preventing CR in addition to correctly identifying it after it happens.

3. Reasons for CR and how it might be prevented

Preventing CR requires an understanding of why this form of responding occurs. Despite many advantages to online data collection, administrators of Internet-based surveys relinquish much of the control they had when overseeing paper and pencil surveys. Researchers have posited that less direct interaction between the administrator and participant (Johnson, 2005), more