ABSTRACT

WARD, MARY KATHRINE. Using Virtual Presence and Survey Instructions to Minimize Careless Responding on Internet-Based Surveys. (Under the direction of Dr. Samuel B. Pond.)

Internet-based survey data inform knowledge creation in research and justify work activities in many organizations. While there are decided advantages to Internet-based surveys, this mode of administration comes with its own set of challenges. Survey respondents may intentionally or unintentionally respond to the survey in a manner that does not accurately reflect their true beliefs or feelings. The purposes of this study were to examine two approaches that address the problem of careless responding (CR), and increase attentiveness among respondents with better survey design. This study investigated instructional manipulation and virtual human presence as potential buffers against CR. The sample consisted of undergraduate students who voluntarily completed an Internet-based survey. This study used a 3x3 between-subjects experimental design where virtual presence (absent, animated shape, and virtual human) and type of instruction (anonymous, warning, and feedback) were the independent variables. Indicators of CR were the dependent variables. Findings showed that warning instructions significantly reduced some forms of CR, but promising feedback had little effect on CR. The interaction of instructions and virtual presence had a significant effect on CR. Future research will need to tease apart the nuances of the relationship between instructions, virtual presence, and CR. The discussion includes implications for Internet-based survey administration and future directions for addressing the problem of CR.

© Copyright 2013 by Mary Kathrine Ward

All Rights Reserved

Using Virtual Presence and Survey Instructions to Minimize Careless Responding on
Internet-Based Surveys

by
Mary Kathrine Ward

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Psychology

Raleigh, North Carolina

2014

APPROVED BY:

Samuel B. Pond, III
Committee Chair

Adam Meade

Lori Foster Thompson

DEDICATION

To my parents and sister who instilled in me the importance of mental development, individual responsibility, and personal growth.

BIOGRAPHY

M.K. Ward was born in Tacoma, Washington in 1986 to parents John and Denise Ward. She earned a Bachelor of Arts in Psychology and a Bachelor of Business Administration from Loyola Marymount University in December of 2008. During the time between undergraduate and graduate school, M.K. worked full-time as a general store manager of a small soccer retailer. She managed inventory, developed marketing plans, advised the company president about strategic planning, hired, trained and managed employees, and wrote job descriptions and a training manual. In August of 2010, M.K. matriculated to North Carolina State University to begin the doctoral program in Industrial/Organizational psychology. She is a member of the Society for Industrial/Organizational Psychologists (SIOP), North Carolina Industrial and Organizational Psychologists (NCIOP), American Psychological Association (APA), the Association for Psychological Science (APS), and the Academy of Management (AOM).

ACKNOWLEDGMENTS

First, a sincere thank-you to all of my committee members: Bob Pond, Adam Meade, and Lori Foster Thompson for their support and feedback throughout the preparation of this thesis. Special thanks to my advisor and chair, Bob Pond, for having the fastest draft turnaround in the East. His feedback has been both timely and thorough, never leaving a word left unturned.

A deep note of gratitude to my family for their support during times of challenge beyond what I thought I could endure alone. A special thank you to my mother who never failed to lend a sympathetic ear, my father who recalibrated my perspective when necessary, and to my sister who consistently connects with me in a way no one else possibly could. Thank you to my friends from Tacoma, from college, and from Raleigh for their encouragement and inspiration. Thank you to my fellow graduate students for their willingness to share information and provide feedback. Finally, a special thanks to Steven Toaddy for his review of the manuscript and for consistently being a supportive mentor and a tough act to follow! Thank you for being a model of excellence.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
Using Virtual Presence and Survey Instructions to Minimize Careless Responding on Internet-Based Surveys	1
Psychometric Problems Associated with Internet-based Surveying.....	3
How to detect CR.....	3
Importance of finding solutions to CR.....	5
Reasons for CR and How it Might be Prevented.....	8
Manipulating instructions	9
Manipulating virtual human presence.....	12
Method	16
Participants.....	16
Study Design.....	16
Procedure	17
Data preparation.....	17
Survey Instruction Conditions	17
Normal instructions.....	18
Feedback instructions.....	18
Warning instructions.....	19
Virtual Presence Conditions	19
No virtual presence	19
Animated shape.....	19
Virtual human	19
Survey Content.....	20
Survey measures to induce CR.	21
Measures to test hypotheses.....	22
Results.....	26
Analysis of the Fidelity of Manipulations	26
Manipulation checks for instructions.....	27
Manipulation checks for virtual presence.	29
Analysis of the Hypotheses.....	30
Discussion	33
References.....	38
APPENDIX.....	55

LIST OF TABLES

Table 1.	Number of Participants, Mean Age (Standard Deviations), and Sex for each condition.	46
Table 2.	Univariate ANOVAs to Check the Efficacy of Manipulations.	47
Table 3.	Correlations Among CR Indicators with Standard Deviations on the Diagonal.	48
Table 4.	Means and Standard Deviations of Dependent Variables by Level of Instructions.	49
Table 5.	Means and Standard Deviations of Dependent Variables by Level of Virtual Presence.	50
Table 6.	Means and Standard Deviations of Dependent Variables by Condition.	51
Table 7.	Univariate ANOVA follow-up tests to the two-way MANOVA	52

LIST OF FIGURES

Figure 1.	Depiction of the animated shape condition.	53
Figure 2.	Depiction of the virtual human condition.....	54

Using Virtual Presence and Survey Instructions to Minimize Careless Responding on Internet-Based Surveys

Advances in technology have spurred the extensive use of Internet-based surveys, assessments, and measures. In academia, Internet-based surveys collect student feedback as a way to improve pedagogy and assess grant-funding processes (Mervis, 2007). In research, job satisfaction (Acquavita, 2009), assessments of knowledge, skills, abilities and other characteristics (Anderson, 2010), and training (Patrick, 2012) are examples of areas in industrial/organizational psychology that use survey data. In practice, industrial/organizational psychologists use online surveys to screen job applicants (Berta, 2006) and to inform administrative recommendations (Marx, 2012). In sum, Internet-based survey data support knowledge creation in research and inform applied work in many organizations.

Survey administrators are increasingly computerizing paper-and-pencil surveys for several reasons (Barak & English, 2002). Compared to paper-and-pencil surveys, Internet-based surveys are more convenient for both respondents and researchers. Internet-based surveys are easy for researchers to create, to ensure standardization, to expand distribution, and to quicken data collection. Furthermore, respondents can complete surveys when and where it is convenient for them, and they can receive feedback or scoring instantly (Barak & English, 2002). Increased speed and convenience, compared to the paper-and-pencil method, propel the popularity of Internet-based surveys.

While there are decided advantages to Internet-based surveys, this mode of administration comes with its own set of challenges. For example, respondents cannot ask researchers questions when taking Internet-based surveys. Also, the convenience of being able to respond to a survey wherever and whenever Internet connectivity is available introduces a variety of settings into the data collection process that, in some circumstances, may influence responses (Hardré, Crowson, & Xie, 2012). Respondents may ignore important parts of the survey, submit multiple times, and be more apt to respond carelessly (Barak & English, 2002).

The prevalence of survey methodology implies that those using Internet-based surveys to collect data believe that the responses they are collecting are coming from attentive respondents. However, because Internet-based surveys collect data in various places at various times predicated unknown context effects on responses, respondents to this kind of survey may often exhibit what has been described as careless responding (CR; Barak & English, 2002; Berry et al., 1992; Curran, Kotrba, & Denison, 2010; Hardré, Crowson, & Xie, 2012; Johnson, 2005; Meade & Craig, 2012). That is, survey respondents may intentionally or unintentionally respond to the survey in a manner that does not accurately reflect their true scores. Research needs to address this problem because the advantages of Internet-based surveys indicate the longevity of this methodology.

The primary concern of this study, seeks to increase attentiveness among respondents with better survey design. Understanding aspects of Internet-based surveys that tend to encourage attentiveness will increase the quality of data collected. Moreover,

inferring from cleaner datasets leads to sounder conclusions. In an indirect way, deterring and identifying CR in Internet-based surveys improves both research and practice because research conclusions derived from higher quality data supports better theory development and its subsequent application. This research improves Internet-based research methodology by avoiding the problem of CR altogether.

Psychometric Problems Associated with Internet-based Surveying

CR is responding with or without regard to item content and reflects inaccuracy rather than respondents' true scores. 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 refers to when one responds to a survey without consideration of item content. Random responding would fall under this category, although some content nonresponsive response patterns 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).

How to detect CR. There are two general ways to identify CR. First, when researchers build the survey they can add special survey items that would indicate CR.

Some examples include: self-report items asking the respondents to rate their engagement during the survey, self-report items asking if their survey data is of sufficient quality for research use, and instructed response items 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 are 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 survey data itself. 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, 2009). Ehlers and colleagues (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, large within-

person variance on the subsets of even and odd items would indicate CR. The last CR indicator is *LongString*, which is the maximum number of consecutive items with the same answer choice selected (Johnson, 2005). Although more CR indicators exist, 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 the previously mentioned CR indicators because they measure CR more effectively and comprehensively than any single CR indicator. The current study also investigated two methods of obviating CR.

Importance of finding solutions to CR. Finding a solution to CR should be a priority to researchers and practitioners for two reasons. First, both research findings and data-driven decision making heavily rely on survey data for justifying decisions and actions based on those decisions. Turnover intentions (Price & Mueller, 1986), job satisfaction (Stanton et al., 2002), and organizational commitment (Mowday, Steers, & Porter, 1979) are just a few organizational constructs of interest to practitioners that survey data assess. Evidence-based conclusions rely on clean datasets assimilated through research.

Although CR is a longstanding issue in survey methodology, few studies examine the quality of responses to surveys beyond typical data cleaning procedures. Studies that employ Internet-based survey methodology rarely use metrics to estimate CR as a means of screening out inattentive respondents. Recent research is starting to fill this gap by investigating methods, beyond typical data cleaning procedures that identify careless respondents (Meade & Craig, 2012).

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 and colleagues (1992) found that one or more items reflected CR in a *majority* of respondents. Berry and colleagues (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 and colleagues (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 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 a cutoff score, 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. Correctly extracting data contributed by careless respondents is a limited solution to CR issues, however.

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 thus 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.

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 in paper and pencil surveys. Researchers have posited that less direct interaction between the administrator and participant (Johnson, 2005), more environmental distractions (Carrier, Cheever, Rosen, Benitez, & Chang, 2009), lower participant interest (Schwartz, 1999), and survey length (Berry et al., 1992) may elicit CR in Internet-based surveys. Johnson (2005) suggests the distance between respondent and survey administrator and increased ease of survey submission via the Internet increase the likelihood and prevalence of respondents rushing through surveys. Carefully constructing survey instructions can discourage respondents from rushing through surveys at the expense of data quality. The varied environments in which respondents can complete Internet-based surveys can increase distracters and influence responses (Hardré, Crowson, & Xie, 2012). Increasing the perceived presence of a survey administrator could improve and maintain the ability of respondents to focus on an Internet-based survey. The current study changed Internet-based survey design to maintain respondent attention and reduce CR as a result.

Thus far, research in CR and Internet-based surveys has largely focused on identifying and extracting CR. As previously mentioned, the problems of removing careless respondents necessitate the search for preventing CR. The current study investigated ways survey instructions and virtual humans can obviate CR and maintain the integrity of the sampling process.

Manipulating instructions. Regardless of the method of administration, surveys often start with instructions that give respondents information about how to complete the survey. Research suggests that carefully crafted survey instructions may influence the attentiveness of respondents. For example, Huang and colleagues (2012) studied the effects of different instructions on the quality of survey responses. In the control group of their study, respondents read instructions asking for honesty and that clarified that there were no correct or incorrect answers. Survey instructions typically emphasize honesty, accuracy, and anonymity (Huang et al., 2012). Researchers have referred to these instructions as “normal” instructions. In what Huang et al. (2012) termed the warning-instructions condition, respondents read about how statistical control methods would check the validity of their responses, and that if flagged for insufficient effort, the respondents would not receive participation credit. Respondents who read the warning instructions showed markedly less CR than respondents in the control group who read the “normal” instructions. Respondents in the warning group showed higher reliability and consistency in their answers. The significant findings of Huang and colleagues’ (2012) implicate warning instructions as a way to reduce the amount of CR among respondents.

In addition to providing survey instructions that emphasize honest responding or that warn against not responding carefully, Meade and Craig (2012) investigated the impact of instructions that reduced the respondents’ perceived anonymity. Meade and Craig studied the effects of three types of instructions they described as “normal,” “identified,” and “warning” instructions. In their study, “normal” instructions informed respondents that

their responses would be anonymous. “Identified” instructions expressed that responses would be confidential and required respondents to enter their names at the bottom of each web page of the survey. “Warning” instructions emphasized to participants that their responses would be subject to the institution’s academic integrity policy. Respondents who read “warning” instructions had to type their name next to a statement to indicate that they had answered all questions with care and honesty before they could move to the next page of the survey. The statement was similar to an integrity statement one might sign after completing an exam. Thus, “warning” instructions implied that responding carelessly would be in violation of the academic integrity policy (i.e., it would be like cheating on an assignment) and might be subject to university sanctions.

Like Huang and colleagues (2012), Meade and Craig (2012) found that manipulating instructions had a significant main effect on two CR indicators. First, instructions affected the number of times respondents agreed with items when agreement was logically impossible. For example, if a respondent agreed with an item that said leprechauns pay their salaries then a data analyst might reasonably infer CR (Meade & Craig, 2012). Second, the amount of attention respondents reported varied by the type of instructions they read. When instructions required respondents to add their names to each survey page, respondents reported higher levels of attention than did the control group that read instructions confirming anonymity. Instructions warning respondents that their answers would be subject to an academic integrity policy had no significant effect on self-reported levels of attention (Meade & Craig, 2012).

In summary, manipulating the content and length of survey instructions influences CR, but the nuances of the relationship between survey instructions and CR remain unknown (Huang et al., 2012; Meade & Craig, 2012). Instruction manipulation impacted CR indicators differently across studies (Huang et al., 2012; Meade & Craig, 2012). Thus, the effect of warning respondents of potential CR identification on a multivariate composite of CR is not well understood. Although the nuances of the relationship between warning respondents of negative repercussions and CR is unknown, previous findings indicate that warnings may significantly reduce CR.

Hypothesis 1: Respondents who read instructions warning them that statistical methods will be used to evaluate their responses will score significantly lower on a multivariate composite of CR than respondents who read normal instructions.

Although notable effects came from instructions hinting at punitive repercussions, these instructions also diminished respondents' attitudes toward the survey (Meade & Craig, 2012). A more positive approach relative to vague threats of punishment would seem to be preferable and worth finding. There is evidence that feedback can improve task performance (Northcraft, Schmidt, & Ashford, 2011) and job outcomes (Colarelli, Dean, & Konstans, 1987). It is possible that the beneficial effects feedback can have on job performance might extend to performance on surveys. Promising feedback about the quality of respondents' survey data may increase the attentiveness of respondents. According to feedback intervention theory, showing feedback can cue respondents to pay attention to themselves; this shift in attention tends to improve task performance (Kluger & DeNisi

(1996). A search through the literature revealed no protocols about feedback on Internet-based surveys. To date, no studies have investigated effects of survey instructions that include a statement promising respondents that they will receive feedback about their survey responses. The current study investigated the utility of promising feedback for reducing CR among respondents.

Hypothesis 2: Respondents who read instructions that promise that they will be given feedback about the quality of their survey responses will score significantly lower on a multivariate composite of CR than respondents who read normal instructions.

Manipulating virtual human presence. In addition to manipulating instructions, research suggests that a second way to impact CR is to include virtual humans in Internet-based surveys to affect attention and perceived accountability among respondents as they complete surveys (Aiello & Svec, 1993; Behrend & Thompson, 2011, 2012; Park, 2009; Park & Catrambone, 2007; Rickenberg & Reeves, 2000; Zambaka, Ulinski, Goolkasian, & Hodges, 2004). The presence of virtual humans has significantly affected performance on experimental tasks of known difficulty levels, i.e., tasks that could be sorted into easy, medium, and difficult categories based on prior research (Aiello & Svec, 1993; Park & Catrambone, 2007; Park, 2009; Rickenberg & Reeves, 2000; Zambaka et al., 2004). Griffith (1993) studied performance on a data entry task that participants completed using a computer. Participants either entered data while alone, electronically monitored by the computer itself, or in the physical presence of a human supervisor. In the electronic

monitoring condition, the computer screen displayed messages telling participants that the computer was monitoring their progress on the task. Performance improved when participants were electronically monitored compared to participants who worked alone, but the increase did not reach statistical significance.

Virtual humans used in these studies have evolved from visual markers of users, to three-dimensional virtual beings with individual behavioral patterns, personalities, and appearances (Ahn, 2012). As technology has advanced, virtual humans bear a closer resemblance to real human beings. The increasing similarity between humans and virtual humans has enabled virtual humans to elicit the same kinds of social responses from people that one would expect during human-to-human interactions (Gratch, et al., 2007; Haxby, Hoffman, & Gobbini, 2002; Park, 2009). Social interactions involve perceptions and judgments of others as well as relating to and influencing others (Park, 2009). These defining features of social interactions are largely present in exchanges between real humans and virtual humans (Nass, Steuer, & Tauber, 1994; Reeves & Nass, 1996). Respondent reactions to a virtual human during survey completion may be similar to reactions to a real human watching them complete a survey. Thus, adding virtual humans to Internet-based surveys could increase perceived interaction between survey administrators and respondents (Park & Catrambone, 2007) and subsequently influence respondents' behavior.

Unlike in-person surveys, Internet-based surveys severely limit interaction between survey administrators and respondents. As a result, there is limited opportunity, if any, for

the effects of social influence to hold respondents accountable for choosing responses carefully. The lack of social interaction in Internet-based surveys and the evidence suggesting that virtual humans might simulate social interaction suggest an avenue that survey developers could use to improve the design of Internet-based surveys. Adding virtual humans to Internet-based surveys may increase respondent accountability by increasing the respondents' perceptions of social interaction with the survey administrators. Both consequences of virtual human presence may lead to a reduction in CR. Research has neither confirmed nor rejected these possible outcomes.

Researchers differ in their posited mechanisms underlying the effects of social interaction during task completion, (i.e. social facilitation effect). One reason for social facilitation effect is people's perception that they are being evaluated. Park (2009) and Griffith (1993) tested this explanation for social facilitation effect, but these authors did not rule out alternative explanations for performance differences. One alternative explanation for the social facilitation effect is that performance changes because of distraction, not perceived evaluation. It may be that the mere presence of something else distracts people from the task, be it the movement of a virtual human, or the appearance and removal of written messages on screen. Distractions lower performance on a task, e.g. completing a survey, but people performing the task may increase their effort in order to compensate for their drop in performance (Aiello & Douthitt, 2001). In the current study, an animated shape will be included to eliminate distraction as an alternative explanation for changes in performance.

Hypothesis 3: Respondents who see a virtual human displayed on an Internet-based survey will score significantly lower on a multivariate composite of CR as compared to respondents who do not see the virtual human displayed.

Just as the literature suggests that displaying virtual humans during survey administration might decrease CR, instructions may reduce CR as well. Instructions warning respondents that they will be identified for CR indicates a potential negative evaluation insofar as people generally do not want to be seen as careless (Huang et al., 2012; Meade & Craig, 2012). Alternatively, instructions promising feedback about the quality of survey data indicates monitoring in less threatening way. Social facilitation research has found that the expectation of evaluation induces performance changes, especially when people perceive negative outcomes as possible (Feinberg & Aiello, 2006, 2010; Klehe, Anderson & Hoefnagels, 2007). Furthermore, there is some evidence of additive effects in social facilitation effects on performance (Feinberg & Aiello, 2006). It follows that combining the presence of a virtual human and a warning of response evaluations may reduce CR to a greater extent than either change individually.

Hypothesis 4: The interaction between virtual human display and warning instructions will lower scores on a multivariate composite of CR to a greater extent than all other conditions.

Taken together, the prevalence of CR throughout Internet-based surveys and the incomplete understanding of CR indicators call for research exploring ways to obviate CR.

In addition to identifying CR in a survey after it occurs, this study investigates instructional manipulation and virtual human presence as potential buffers against CR.

Method

Participants

A total of 502 participants were recruited from a pool of students enrolled in introductory psychology courses at a large Southeastern university. Deletion of participants who omitted one or more pages of the survey reduced the total *N* of 502 to 424 participants. Cell sizes per condition ranged from 29-64 people due to random assignment and deletion of missing data. Deleted cases ranged from 2-12% of cases across conditions. Table 1 presents sample sizes by condition. Missing data from participants prone to fail to complete commitments, to resist committing to a task, or to procrastinate may be nonrandom. However, it is the aforementioned characteristics that are typical of CR, thereby increasing confidence in significant results. Demographic items assessed the age, sex, ethnicity, and native language of each respondent. The mean age of the sample was 19.11 years old ($SD = 1.30$), the maximum age was 30 years old and 63% of the total sample was female. Table 1 presents descriptive statistics.

Study Design

This study used a 3x3 between-subjects experimental design where virtual presence (no presence, animated shape, and virtual human) and type of instruction (normal, feedback, and warning) were the independent variables. CR indicators were the dependent variables. Random assignment placed participants into the control group (no virtual

presence with normal instructions) or one of the experimental conditions in which respondents saw some combination of virtual presence and instruction content on the survey pages.

Procedure

Instructors of introductory psychology courses recruited student respondents by giving them the option of participating in research to fulfill a course requirement. Participants received a hyperlink to a webpage hosted by the survey administrator after signing up for the study. At the webpage, participants agreed to the conditions specified in an informed consent statement, and a JavaScript routine randomly assigned participants to one of a combination of the three levels of virtual presence and the three levels of instruction for a total of nine experimental conditions (including the control condition). The JavaScript routine automatically directed respondents to the web page corresponding with the appropriate survey. All respondents read a note of gratitude for their participation and a debriefing statement immediately after they completed the survey.

Data preparation. Three of the CR indicators required some modifications before further analysis because MANOVA requires interval level data. Details about the modifications follow the descriptions of each of the CR indicators.

Survey Instruction Conditions

The survey displayed one of three types of instruction sets described in more detail below. Like other surveys, participants read instructions before responding to any survey items. Then participants viewed instructions on separate webpages interleaved with

webpages that contained survey items. For example after submitting the first webpage of the survey, participants again saw the instructions on a webpage that they had to click through to access the remaining webpages of the survey. Thus, participants saw the instructions a total of six times.

Normal instructions. Respondents in the normal instructions condition saw normal survey instructions adapted from Huang and colleagues (2012). The normal instructions stated, “There are no correct or incorrect answers on this survey. Please respond to each statement or question as honestly and accurately as you can. Your answers will be kept strictly confidential.” Respondents in the normal instruction condition served as the control group for this manipulation.

Feedback instructions. Respondents assigned to the feedback instructions condition read a message at the beginning of their survey that confirmed the confidentiality of their responses and informed participants that they would receive feedback about the quality of their survey responses upon completion. Feedback instructions stated, “Your honest and thoughtful responses are important to us and to the study. You will receive feedback about the quality of your survey responses and whether we can use the information that you provided to us upon completion of the survey. Your answers will be kept strictly confidential.” Upon survey completion, participants in this condition read the following message that provided feedback about their data quality: “Thank you for your time spent taking that survey! Your survey responses were of sufficient quality for use in the study.”

Warning instructions. Respondents assigned to the warning instructions condition read that while their confidentiality would remain protected, researchers would employ statistical methods to evaluate the quality of their survey responses. Furthermore, the instructions warned that data would be marked for low quality where applicable. Warning instructions stated, “Your honest and thoughtful responses are important to us and to the study. To ensure the quality of survey data, your responses will be subject to sophisticated statistical control methods. Responding without much effort will be flagged for low-quality data. Your answers will be kept strictly confidential.”

Virtual Presence Conditions

No virtual presence. Participants in this condition served as the control group for this manipulation, as they took the Internet-based survey under typical conditions, i.e., without any virtual human or geometric shape displayed on screen along with the survey.

Animated shape. In the animated shape condition, an animated circle appeared at the beginning of the survey and remained visible until completion. The interior of the circle shape pulsed continuously in order to approximately match the amount of movement of the virtual human condition. Figure 1 presents two screenshots of the animated shape condition.

Virtual human. Participants assigned to this condition saw a virtual human throughout the survey. In this condition, the virtual human exhibited movements such as breathing and blinking. The process of building the virtual human consisted of three steps. First, Averager open-source software provided by Face Research Lab enabled me to create a composite photo of a human face that was both gender neutral and racially ambiguous

without being distasteful. It was important to create a virtual human that had unclear attributes to avoid confounding significant effects of virtual human presence with effects due to similarity between the virtual human and the respondent (Zanbaka, et al., 2004). In the second step, the composite photo became the input into Haptik's PeoplePutty software that created the three dimensional virtual human. Third, I used Camtasia Studio for Windows to record the virtual human to make it display lifelike behaviors such as breathing and blinking. The video file of the virtual human, saved as a graphics interchange formatted (GIF) file, could subsequently be programmed into the survey. Figure 2 presents two screenshots of the virtual human condition.

Survey Content

To generate data adequate for testing my hypotheses about CR, I used a collection of measures to create a lengthy yet realistic survey. Berry and colleagues (1992) found a tendency for CR to occur near the end of long surveys. To ensure CR had an opportunity to take place, numerous survey items filled 6 web pages with 75 items per page. The seven measures that comprised the survey were filler in the sense that their purpose was to lengthen the survey. The survey contained items that were CR indicators, and such items were critical for testing the hypotheses. Details about CR indicators appear in the section titled *Measures to test hypotheses*. Responses were on 7-point Likert scales with radio buttons. Participants responded to self-report questions about their awareness of the different messages in the instructions and virtual presence to check the manipulations. Table 2 presents manipulation check items. I used Qualtrics survey software to administer

the Internet-based survey. I arranged Qualtrics settings so that respondents were required to take the survey in one sitting without the option of saving and returning to the survey later. The following measures comprised a survey of sufficient length to induce CR.

Survey measures to induce CR.

External organizational justice. External Organizational Justice (EOJ) measures employee perceptions “of the degree to which her or his organization behaves fairly, equitably, and ethically when interacting with entities outside of the organization” (Toaddy & Pond, 2012, p. 9). According to Toaddy and Pond (2012) the 11-item EOJ scale measures three types of EOJ: four items measuring distributive external justice, four items measuring procedural external justice, and three items measuring interactional external justice.

Personality. The 300-item International Personality Item Pool (IPIP; Goldberg, 1999) is a measure of the Five Factor Model (McCrae & Costa, 1987) of personality and was a large part of the survey. The survey also included a 6-item measure of other orientation (De Dreu & Nauta, 2009) and a 40-item measure of narcissism (Raskin & Terry, 1988).

Social desirability. The survey included two social desirability measures. Social desirability scales contain items that would make the respondent seem incredibly virtuous if he or she endorsed a large number of the items. The survey included the 33-item Marlowe-Crowne social desirability scale (Crowne & Marlowe, 1960) with an adapted response scale to match the same seven-point response format used for other items in the survey. The

second measure of social desirability in the survey was the IPIP social desirability scale (Goldberg, 1999).

Measures to test hypotheses. I used the following six measures to assess CR in respondent data. The values of each CR indicator were used to assess the aforementioned hypotheses.

Instructed-response items. Instructed-response items directed the respondent to use a particular response option as the correct response option. An example of an instructed-response item is, “Select ‘strongly disagree’ for this item.” Item interpretation for instructed-response items is literal and unambiguous. Furthermore, the metric is clear for scoring correct or incorrect responses on instructed-response items (Meade & Craig, 2012). A total of six instructed response items appeared in the survey, with one instructed response item in every segment of 75 items. Adding more instructed-response items could agitate the respondents, potentially resulting in more CR or even prompting them to quit responding to the survey altogether (Meade & Craig, 2012). The survey included a total of six instructed-response items, with possible values between zero (correctly answered all items) and six (incorrectly answered all items).

Data preparation of Instructed-response items. Responses to instructed response items are either correct or incorrect, making this CR indicator a dichotomous outcome variable. I recoded instructed response items to a score of “0” for responses that matched the instructed choice, and “1” for responses that did not match the instructions. The sum of the six instructed response items for each participant became the total instructed response

score for every participant. The sum, hereafter referred to as instructed response sum, created a continuous dependent variable for inclusion in the MANOVA. Possible values for the total instructed response scores ranged from zero to six, indicating all correct responses to all incorrect responses respectively. For example, people who answered five of the six instructed response items incorrectly received a score of five for their total. High scores on the total instructed response variable indicates more CR.

Self-reported single item (SRSI) indicator. The survey included a single item measure adapted from Meade and Craig (2012). Respondents read instructions that said, “It is vital to our study that we only include responses from people that devoted their full attention to this study. Otherwise, years of effort (the researchers’ and the time of other participants) could be wasted. You will receive credit for this study no matter what. In your honest opinion, should we use your data in our analyses in this study?” (p. 442). In this survey, the response scale for this item was a seven-point Likert scale ranging from strongly disagree to strongly agree. This item collected self-report data of respondents’ judgments of the quality of their own responses with possible values ranging between one and seven.

Response time. The response time for each respondent was a record of the time it took a respondent to complete the entire survey. Qualtrics settings recorded the times every respondent started and finished the survey. I calculated the total survey duration by subtracting the start time from the finish time. Larger values in response time meant respondents spent more minutes completing the survey.

Outlier analysis. I used Mahalanobis distance measures to identify CR by calculating each respondent's distance from the average response pattern. This value is calculated from the series of responses for each scale. Like Meade and Craig (2012), I calculated one Mahalanobis distance measure for each of the five personality factors (60 items per factor), giving every respondent five Mahalanobis distance measures. The average of those Mahalanobis distance measures produced a single Mahalanobis distance value per respondent, with higher values indicating more CR.

Consistency indicator. Logically, participants who paid attention to the survey should have chosen equivalent response options across similar items. For example, participants responding to items measuring cheerfulness will show consistent agreement or disagreement to cheerfulness items if they are paying attention. Agreeing strongly and disagreeing strongly to items that assess the same construct would not make sense logically. An *Even-Odd Consistency* measure of CR assessed the extent to which participants chose equivalent response options to items measuring similar constructs. To calculate the *Even-Odd Consistency* measure, unidimensional scales were divided into two subscales based on the order in which items appeared. The first item was assigned the number one, the next item was assigned number two, and so on. Thus, one subscale consisted of even-numbered items and the other consisted of odd items (Jackson, 1977). This split in the scale created two subscales and every respondent received a score for each subscale. A respondent's subscale score for even items was the average of that person's responses on all even items. The average of the person's responses on all odd items was that person's score for the odd

subscale. Finally, a within-person correlation from the even and odd subscale scores for the respondent was that respondent's *Even-Odd Consistency* value. The range of possible values for the *Even-Odd Consistency* indicator is between negative one and one.

Data preparation of Even-Odd Consistency. Qualtrics survey software randomized survey items within surveys for each individual respondent. Consequently, each respondent viewed survey items in a unique order. This became problematic for calculating *Even-Odd Consistency* as stipulated by Jackson (1977). Simply using scale order to split scales into even and odd subscales would have completely ignored viewing order. Thus, I rearranged items by viewing order for each respondent because the items on unidimensional scales measure one construct.

Exploratory factor analysis tested the dimensionality of the EOJ scale, IPIP scales, Marlowe-Crowne social desirability scale, and narcissism scale. Scales with insufficient alphas were excluded from further EOC calculations. Items with non-significant loadings and cross-loadings, were removed. Additionally, factors with three or fewer items were removed because subscale averages could not have been calculated. All unidimensional scales achieved adequate reliability (.70 or higher) with the exceptions of the sympathy scale ($\alpha = .68$) and single factor subscales of activity level ($\alpha = .67$) and Marlowe-Crowne social desirability scale ($\alpha = .69$; Crowne & Marlowe, 1960). Refining scales for the calculations of the *Even-Odd Consistency* occurred independently of calculating Mahalanobis distance and thus, did not reduce variance in other CR indicators. The 34

refined single factor scales entered into calculations of an *Even-Odd Consistency* value for each respondent.

Response pattern. It was possible for a participant to show high levels of consistency across items while responding carelessly, i.e., answering all survey items with the same response option. In this study *LongString* values identified response patterns where participants repeatedly chose the same response option. Johnson (2005) recommended using the *LongString* indicator as a way to detect highly consistent, yet inaccurate responding. Meade and Craig (2012) found *Maximum LongString* values effectively screened data for this type of CR. For the current study, a Visual Basic for Applications program in Microsoft Excel computed for each survey web page the maximum number of consecutive items with identical responses. The *Maximum LongString* value for one respondent was the largest *LongString* value found on any of the survey web pages. Every respondent included in the analyses received a *Maximum LongString* value based on his or her responses, with the possible range of values between zero (no consecutive identical responses) and 75 (all responses on the webpage were identical).

Results

Analysis of the Fidelity of Manipulations

Prior to testing the hypotheses of this study, I ran a series of manipulation checks to verify that participants perceived differences among the types of instructions and among the types of virtual presence. Participants indicated their levels of agreement to items that described the instructions such as, “I will receive feedback about the quality of my survey

responses.” Items also checked that participants perceived the intended virtual presence, e.g. “There was an animated picture of a person on the upper left side of my survey display.” With the exception of normal instructions, respondents perceived levels of instructions and virtual presence as intended.

Manipulation checks for instructions. To determine if respondents accurately perceived the content of the different instruction conditions (normal, feedback, and warning) I ran one-way fixed effects ANOVAs on six manipulation check items. Table 2 presents results of the analyses. The first item assessed the normal instructions conditions by stating, “The instructions for this survey asked me to answer honestly and accurately.” Contrary to the expectation that participants who read normal instructions would show the highest agreement to this item, there were no significant differences in agreement among participants who read normal instructions ($M = 6.42$, $SD = 0.80$), participants who read feedback instructions ($M = 6.48$, $SD = 0.80$), and participants who read warning instructions ($M = 6.55$, $SD = 0.63$). The second item verified that participants could recall the instructions (“I remember what the instructions said for this survey”) rather than assessing a particular level of instructions. Participants who read normal instructions ($M = 4.83$, $SD = 1.59$), feedback instructions ($M = 4.95$, $SD = 1.49$), or warning instructions ($M = 4.99$, $SD = 1.61$) all indicated on average that they remembered the instructions; there were no significant differences among groups.

Two items verified that participants accurately perceived the message of feedback instructions; these were the third and fourth items. The third item stated, “I will find out if

my data can be used after I complete this survey.” As expected, participants who read feedback instructions showed significantly higher agreement to this item ($M = 5.65$, $SD = 1.71$) than participants who read warning instructions ($M = 4.14$, $SD = 1.43$) or normal instructions ($M = 4.26$, $SD = 1.61$). The level of agreement was not significantly different between the participants who read warning instructions and participants who read normal instructions

For the following items, the Levene’s test of the assumption of homogeneity of variance was significant. Thus, for those items I used the non-parametric Welch’s test because of the test’s robustness to violations of the homogeneity of variance. Item four stated, “I will receive feedback about the quality of my survey responses.” As expected, participants who read feedback instructions showed significantly higher agreement to this item ($M = 5.80$, $SD = 1.39$) than participants who read warning instructions ($M = 4.27$, $SD = 1.72$) or normal instructions ($M = 3.94$, $SD = 1.76$). The level of agreement was not significantly different between the participants who read warning instructions and participants who read normal instructions.

The final two items verified that participants accurately perceived the message of warning instructions. Item five stated, “My survey will be flagged for low quality if I didn't complete this survey carefully.” As expected, participants who read warning instructions showed significantly higher agreement to this item ($M = 5.93$, $SD = 1.46$) than participants who read feedback instructions ($M = 5.28$, $SD = 1.80$) or normal instructions ($M = 4.84$, SD

= 1.80). The level of agreement was not significantly different between the participants who read feedback instructions and participants who read normal instructions.

Item six stated, “Statistical control methods will evaluate the quality of my survey data.” As expected, participants who read warning instructions showed significantly higher agreement to this item ($M = 6.03$, $SD = 1.05$) than participants who read feedback instructions ($M = 5.61$, $SD = 1.38$) or normal instructions ($M = 5.28$, $SD = 1.41$). The level of agreement was not significantly different between the participants who read feedback instructions and participants who read normal instructions. Table 2 presents ANOVA results for instruction manipulation checks.

Manipulation checks for virtual presence. I conducted a series of one-way fixed effects ANOVAs to assess whether participants perceived the virtual presence conditions (none, animated shape, and virtual human) accurately. One item verified that participants noticed the presence of the animated shape (“There was a fading red circle on the upper left side of my survey display”). The Levene’s test of the assumption of homogeneity of variance was significant. Thus, for this item I used the non-parametric Welch’s test because of the test’s robustness to violations of homogeneity of variance. As expected, participants in the animated shape condition showed significantly higher agreement ($M = 6.50$, $SD = .97$) than participants in the virtual human condition ($M = 2.20$, $SD = 1.43$), or participants in the no virtual presence condition ($M = 2.17$, $SD = 1.26$).

Three items verified that participants accurately perceived the virtual human and assessed how participants interpreted the virtual human. The first of these three items

stated, “There was an animated picture of a person on the upper left side of my survey display.” As expected, participants in the virtual human condition showed significantly higher agreement ($M = 6.45$, $SD = 1.05$) than participants in the animated shape condition ($M = 1.73$, $SD = 1.20$), or participants in the no virtual presence condition ($M = 1.96$, $SD = 1.33$). The next item stated, “There was a computer program monitoring my survey activity through an animated picture of a person.” As expected, participants in the virtual human condition showed significantly higher agreement ($M = 4.85$, $SD = 1.90$) than participants in the animated shape condition ($M = 2.21$, $SD = 1.52$), or participants in the no virtual presence condition ($M = 1.98$, $SD = 1.19$). The last item stated, “There was a person monitoring my survey activity through an animated picture of a person.” As expected, participants in the virtual human condition showed significantly higher agreement ($M = 3.98$, $SD = 2.09$) than participants in the animated shape condition ($M = 1.95$, $SD = 1.22$), or participants in the no virtual presence condition ($M = 1.97$, $SD = 1.20$). The Levene’s test of the assumption of homogeneity of variance was significant for the last two items; thus, I used Welch’s test. Table 2 presents ANOVA results for instruction manipulation checks. All ANOVAs were significant at $p < .001$ showing strong evidence that respondents clearly perceived the levels of virtual presence.

Analysis of the Hypotheses

I performed a 3 x 3 MANOVA using SAS 9.3. The between subjects variables were instructions (normal, feedback, or warning) and virtual presence (no presence, animated shape, or virtual human). The dependent variables (i.e., instructed-response sum, SRSI

indicator, response time, Mahalanobis distance, *Even-Odd Consistency*, and *Maximum LongString*) formed the multivariate composite of CR. Table 3 presents correlations among dependent variables that consisted of the six aforementioned CR indicators.

A series of data checks verified assumptions of MANOVA in each of the nine conditions (Tabachnick & Fidell, 2013). Examinations of linearity and multicollinearity were satisfactory. Sample sizes in each condition were large enough to ensure robustness to violations of normality. New values converted back from z scores replaced the values of 47 univariate outliers. The results from Box's M-test ($\chi^2 = 662.32, p < .0001$) and variance ratios (greater than 10:1) warrant using the less powerful Pillai's criterion because it is more robust to violations of the homogeneity of variance assumption (Tabachnick & Fidell, 2013).

With the use of Pillai's criterion, the multivariate composite of CR was significantly affected by instructions, $F(12, 816) = 1.83, p = .04$, and but not by the interaction of instructions and virtual presence, $F(24, 1640) = 1.16, p = .27$, or by virtual presence alone, $F(12, 816) = 1.43, p = .15$ (although Roy's greatest root criterion did reach significance $F(6, 408) = 2.64, p = .02$). Instructions explained 1.32% (partial $\eta^2 = .01$) of the variance in the multivariate composite of CR. Table 4 presents means and standard deviations of every CR indicator by level of instructions. The interaction of instructions and virtual presence explained 1.04% (partial $\eta^2 = .01$). Virtual presence explained 1.67% of the variance in the multivariate composite of CR (partial $\eta^2 = .02$). Table 5 presents means and standard deviations of every CR indicator by level of virtual presence.

Hypothesis one stated that respondents who read instructions warning them that statistical methods would be used to evaluate their responses, would score significantly lower on a multivariate composite of CR than respondents who read normal instructions. The overall MANOVA ($F(12, 816) = 1.83, p = .04$) indicated that groups significantly differed in CR due to level of instruction. Lower values on CR indicators indicate less CR, with the exception of extremely small values of total response time. The results of follow-up univariate ANOVAs did not show significant differences due to warning instructions. These failed to support hypothesis one. Table 6 presents means and standard deviations for each dependent variable by condition, and Table 7 presents ANOVA results for each dependent variable.

Hypothesis two stated that respondents who receive instructions that promise that they will be given feedback about the quality of their survey responses would score significantly lower on a multivariate composite of CR than respondents who read normal instructions. The overall MANOVA ($F(12, 816) = 1.83, p = .04$) indicated that groups significantly differed in CR due to level of instruction. The results of follow-up univariate ANOVAs showed significant differences in total response time ($F(2, 412) = 3.97, p = .02, \eta^2 = .02$ with 95% confidence limits from 0.14 to 21.58). Planned comparisons showed that on average participants who read feedback instructions spent significantly longer on the survey ($M = 61.05$) than participants who read normal instructions ($M = 51.07$). All other follow-up ANOVAs were not significant. Table 5 presents ANOVA results for conditions

in which participants read feedback instructions. These results show marginal support for hypothesis two.

Hypothesis three stated that respondents who saw a virtual human displayed on an Internet-based survey would score significantly lower on a multivariate composite of CR as compared to respondents who do not see the virtual human displayed. The nonsignificant results of the overall MANOVA model of virtual presence ($F(12, 816) = 1.43, p = .15$) failed to support hypothesis three.

Hypothesis four stated that the interaction between virtual human display and warning instructions will lower scores on a multivariate composite of CR to a greater extent than all other conditions. The nonsignificant results of the overall MANOVA model of the interaction of instructions and virtual presence ($F(24, 1640) = 1.16, p = .27$) failed to support hypothesis four.

Discussion

The hypotheses in this study predicted that respondents would exhibit less CR due to manipulations of instructions and virtual presence. Analyses showed marginal support for the proposition that manipulating the content of survey instructions could affect CR. Warning instructions failed to reduce CR more than other types of instruction. Thus, results failed to support hypothesis one that predicted that warning instructions would reduce CR compared with normal instructions. Promising feedback via instructions influenced respondents to spend more time on survey completion. However, promising feedback failed to reduce CR on all other types of indicators. These findings provide some initial support

for hypothesis two, which predicted that feedback instructions would reduce CR compared with normal instructions. The significant main effect of survey instructions on CR is consistent with previous findings (Huang et al., 2012; Meade & Craig, 2012). This study makes a unique contribution to the literature by finding that promising feedback via instructions is a potentially useful way to obviate CR.

Results failed to support the prediction in hypothesis three that displaying a virtual human would reduce CR. Likewise, results failed to support hypothesis four, which stated that the largest reduction in CR would result from the combination of displaying both a virtual human and instructions warning respondents of negative consequences of CR.

There are a few possible explanations for the lack of support for the effect of virtual presence on CR. First, virtual humans affect outcomes differently based on specific attributes of the virtual human (Park, 2009). The virtual human for this study was race neutral and gender neutral and that ambiguity may have reduced the influence of the virtual human on CR. Similarity between the virtual human and participant may strengthen the effect (Behrend & Thompson, 2011). Second, participants may need to perceive the virtual human as an evaluative presence. In this study, participants showed the highest agreement for the interpretation that the virtual human was just an animated picture of a person. Post hoc ANOVAs showed no significant differences in interpretations of the virtual human based on instructions. Changing the design of the virtual human may give the impression that the virtual human is evaluating participant behaviors during the survey. Consequently, this may induce a stronger effect and obviate CR.

One limitation warrants consideration when interpreting the findings from this study. The sample consisted solely of undergraduate students and as a result, findings may not generalize to applied settings such as annual performance review surveys in organizations. In other words, employees completing organizational surveys presumably are more concerned about the results than students completing a survey for course credit. Ostensibly, promising feedback via instructions may be more effective when respondents are employees who appreciate receiving feedback.

Future Directions

Organizations may be constrained in what they can include in survey instructions to improve the quality of survey data. In some contexts, organizations may not be able to promise feedback to survey respondents. For example, promising feedback would not be appropriate when surveying employees about sensitive topics such as counterproductive work behaviors; doing so would likely decrease perceived anonymity. Organizations that need to optimize the quality of their survey data but are constrained in terms of instruction content, could benefit from further investigation into the ways virtual presence decreases CR. Future studies may be able to specify elements of virtual presence that are necessary for reducing CR. Furthermore, research would do well to delineate the limits of inducing perceptions of accountability in the virtual context. Research investigating effects of specific characteristics of virtual humans should focus on increasing the saliency of the virtual human in the context of Internet-based surveys given the nonsignificant findings from using a gender- and race-neutral virtual human in this study (Behrend & Thompson,

2011; Park, 2009; Park & Catrambone, 2007). Enabling participants to control features of the virtual human is one way to potentially increase saliency of the virtual human (Behrend & Thompson, 2012). Identifying optimal ways of using virtual humans to increase personal accountability and thereby reduce CR is an essential first step towards increasing data quality on Internet-based surveys.

In addition to determining possible effects of virtual presence, future research needs to determine optimal survey instructions given the particular context, e.g. typical respondents, purpose, etc., of a survey. This study offered feedback to participants as a way to reduce CR in a more positive way than warning participants of repercussions for CR. This presumed that participants cared about received feedback about their survey data, and the marginal support for this presumption suggests that it is at most partially true in student populations. Students may not have perceived such feedback as especially interesting and therefore, they were not motivated to attentively respond to survey items. Presumably, the ability of promised feedback to affect CR applies to a greater extent in job contexts where respondents may have more invested in the outcomes of the survey than students completing a survey for course credit. Future studies should examine the effects of promised feedback in organizational contexts to determine the generalizability of the results of this study.

In sum, this study demonstrates the utility of crafting instructions to reduce CR in data collected with Internet-based surveys. In practice, general trends show that the prevalence of Internet-based surveys in organizations will continue (Barak & English,

2002; Berta, 2006; Marx, 2012) and this study has implications for both theory and practice in regard to this mode of inquiry. In the work context, modifying instructions is a very low-cost way to influence the quality of survey responses. My results showed some initial support for the effectiveness of feedback instructions compared with normal instructions. Yet organizations can be constrained by the influence they can leverage via instructions. Furthermore, it is probably harder to construct and deliver meaningful feedback instructions for employees than for college students. As such, finding ways to increase accountability by modifying other aspects of the design of the survey itself seems necessary. This study is one of the first to investigate the impact of including a virtual human to increase respondent accountability. Although results show that a generic virtual human may not create a virtual social facilitation effect, they suggest that instructions can affect CR on Internet-based surveys.

REFERENCES

- Acquavita, S. P. (2009). Personal and organizational diversity factors' impact on social workers' job satisfaction: Results from a national internet-based survey. *Administration in Social Work*, 33, 151. doi:10.1080/03643100902768824
- Ahn, S. J., Fox, J., & Bailenson, J. N. (2012). Avatars. In Bainbridge, W. S. (Ed.), *Leadership in Science and Technology: A Reference Handbook*. SAGE Publications.
- Aiello, J.R, & Svec, C. M. (1993). Computer monitoring of work performance: Extending the social facilitation framework to electronic presence. *Journal of Applied Social Psychology*, 23, 537-548. doi:10.1111/j.1559-1816.1993.tb01102.x
- Anderson, A. L. F. (2010). Assessment of qualifications needed by environmental health graduates entering private-sector employment. (cover story). *Journal of Environmental Health*, 72(9), 14.
- Averager [Computer Software]. University of Aberdeen, Aberdeen, UK: Face Research Lab.
- Barak, A., & English, N. (2002). Prospects and limitations of psychological testing on the internet. *Journal of Technology in Human Services*, 19, 65-89.
- Behrend, T., & Foster Thompson, L. (2012). Using animated agents in learner-controlled training: the effects of design control. *The International Journal of Training and Development*.

- Behrend, T., & Foster Thompson, L. (2011). Similarity effects in online training: Effects with computerized trainer agents. *Computers in Human Behavior*, 27, 1201-1206.
- Berry, D. T. R., Wetter, M. W., Baer, R. A., Larsen, L., Clark, C., & Monroe, K. (1992). MMPI-2 random responding indices: Validation using a self-report methodology. *Psychological Assessment*, 4, 340-345. doi:10.1037/1040-3590.4.3.340
- Berta, D. (2006, November 13). Caribou Coffee Co. refines hiring process with online applicant-screening program. Penton Media, Inc.
- Carrier, L. M., Cheever, N. A., Rosen, L. D., Benitez, S., & Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of americans. *Computers in Human Behavior*, 25, 483-489. doi:10.1016/j.chb.2008.10.012
- Colarelli, S. M., Dean, R. A., & Konstans, C. (1987). Comparative effects of personal and situational influences on job outcomes of new professionals. *Journal Of Applied Psychology*, 72, 558-566. doi:10.1037/0021-9010.72.4.558
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Curran, P. G., Kotrba, L., & Denison, D. (2010). *CR in surveys: Applying traditional techniques to organizational settings*. Paper presented at the 25th Annual Conference of the Society for Industrial/Organizational Psychology, Atlanta, GA.
- De Dreu, C. W., & Nauta, A. (2009). Self-interest and other-orientation in organizational behavior: Implications for job performance, prosocial behavior, and personal

- initiative. *Journal Of Applied Psychology*, 94, 913-926.
- Ehlers, C., Greene-Shortridge, T.M. Weekley, J.A. & Zaiack, M.D. (2009). *The exploration of statistical methods in detecting random responding*. Poster session presented at the annual meeting for the Society for Industrial/Organizational Psychology, Atlanta, GA.
- Feinberg, J. M., & Aiello, J. R. (2006). Social Facilitation: A Test of Competing Theories. *Journal of Applied Social Psychology*, 36, 1087-1109. doi:10.1111/j.0021-9029.2006.00032.x
- Feinberg, J. M., & Aiello, J. R. (2010). The effect of challenge and threat appraisals under evaluative presence. *Journal of Applied Social Psychology*, 40, 2071-2104. doi:10.1111/j.1559-1816.2010.00651.x
- Field, A. & Miles, J. (2010). *Discovering statistics using SAS statistics*. Sage.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, 7, 7-28.
- Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M., van der Werf, R. J., & Morency, L. P. (2007). Can virtual humans be more engaging than real ones? Paper presented at the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments, Beijing, China.
- Griffith, T. L. (1993). Monitoring and performance: A comparison of computer and

- supervisor monitoring. *Journal of Applied Social Psychology*, 23, 549-572.
doi:10.1111/j.1559-1816.1993.tb01103.x
- People Putty [Computer Software]. Freedom, CA: Hapttek, Inc. Retrieved from
<http://www.hapttek.com/?url=http://www.hapttek.com/products/peopleputty/>
- Hardré, P. L., Crowson, H. M., & Xie, K. (2012). Examining contexts-of-use for web-based and paper-based questionnaires. *Educational and Psychological Measurement*. doi:10.1177/0013164412451977
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, 51, 59–67.
doi:10.1016/S0006-3223(01)01330-0
- Huang, J., Curran, P., Keeney, J., Poposki, E., & DeShon, R. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27, 99–114. doi:10.1007/s10869-011-9231-8
- Jackson, D. N. (1977). *Jackson Vocational Interest Survey manual*. Port Huron, MI: Research Psychologists Press.
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality*, 39, 103-129.
doi:10.1016/j.jrp.2004.09.009
- Klehe, U., Anderson, N., & Hoefnagels, E. A. (2007). Social facilitation and inhibition during maximum versus typical performance situations. *Human Performance*, 20, 223-239. doi:10.1080/08959280701333040

- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254-284.
- Marx, J. (2012). Nonprofit governance: Improving performance in troubled economic times. *Administration in Social Work*, 36, 40. doi:10.1080/03643107.2010.550670
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90. doi:10.1037/0022-3514.52.1.81.
- Mervis, J. (2007). NSF survey of applicants finds a system teetering on the brink. *Science*, 317, 880-881.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17, 437–455. doi:10.1037/a0028085
- Mowday, R. T., Steers, R. M., Porter, L. W. (1979). The measurement of organizational commitment. *Journal of vocational behavior*, 14, 224–247.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, CHI '94 (pp. 72–78). New York, NY, USA: ACM. doi:10.1145/191666.191703
- Nichols, D. S., Greene, R. L., & Schmolck, P. (1989). Criteria for assessing inconsistent

- patterns of item endorsement on the MMPI: Rationale, development, and empirical trials. *Journal of Clinical Psychology*, 45, 239-250. doi:10.1002/1097-4679(198903)45:2<239::AID-JCLP2270450210>3.0.CO;2-1
- Northcraft, G. B., Schmidt, A. M., & Ashford, S. J. (2011). Feedback and the rationing of time and effort among competing tasks. *Journal of Applied Psychology*, 96, 1076-1086. doi:10.1037/a0023221
- Park, S. (2009). *Social Responses to Virtual Humans: The Effect of Human-Like Characteristics*. (Unpublished doctoral dissertation). Georgia Institute of Technology, Atlanta, Georgia.
- Park, S., & Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49, 1054 - 1060. doi:10.1518/001872007X249910
- Patrick, J. (2012). Being in one's chosen job determines pre-training attitudes and training outcomes. *Journal of Occupational & Organizational Psychology*, 85, 245. doi:10.1111/j.2044-8325.2011.02027.x
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609. Price, J. L., & Mueller, C. W. (1986). *Handbook of organizational measurement*. doi: 10.1108/01437729710182260
- Raskin, R., & Terry, H. (1988). A principal-components analysis of the narcissistic

- personality inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology*, 54, 890-902.
- Reeves, B., & Nass, C. I. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. Stanford, California: Cambridge University Press. Retrieved from <http://catalog.lib.ncsu.edu/record/NCSU923425>
- Rickenberg, R., & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '00* (pp. 49–56). New York, NY, USA: ACM. doi:10.1145/332040.332406
- Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, 9, 367–373. doi:10.1177/014662168500900405
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93–105. doi:10.1037/0003-066X.54.2.93
- Stanton, J. M., Sinar, E. F., Balzer, W. K., Julian, A. L., Thoresen, P., Aziz, S., Fisher, G.G., & Smith, P.C. (2002). Development of a compact measure of job satisfaction: The abridged Job Descriptive Index. *Educational and Psychological Measurement*, 62, 173-191.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.

Toaddy, S. R., & Pond, S. B. (2012). *Validation of a Measure of External Organizational Justice*. (Unpublished doctoral dissertation). North Carolina State University, Raleigh, North Carolina.

Woods, C. (2006). CR to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, 28, 186–191. doi:10.1007/s10862-005-9004-7

Zanbaka, C., Ulinski, A., GoolkasianP., & Hodges, L. F. (2004). Effects of virtual human presence on task performance. *Proceedings of International Conference on Artificial Reality and Telexistence (ICAT)* (pp. 174–181).

Table 1.
Number of Participants, Mean Age (Standard Deviations), and Sex for each condition.

Type of Instruction	Type of Virtual Presence				Total
	None	Animated Shape	Virtual Human		
Normal	<i>N</i>	67	57	43	167
	Age	19.24 (1.34)	19.04 (1.13)	19.21 (1.42)	19.16 (1.29)
	Sex	35 F	34 F	30 F	99
		32 M	20 M	13 M	68
Promised Feedback	<i>N</i>	33	29	47	109
	Age	19.24 (1.64)	19.10 (1.11)	18.89 (0.95)	19.06 (1.24)
	Sex	21 F	15 F	34 F	70
		12 M	14 M	13 M	39
Warning	<i>N</i>	45	45	58	148
	Age	19.18 (1.05)	18.87 (0.92)	19.36 (1.79)	19.16 (1.37)
	Sex	30 F	29 F	39 F	98
		15 M	16 M	19 M	50
Total	<i>N</i>	145	131	148	424
	Age	19.22 (1.33)	18.99 (1.06)	19.17 (1.47)	19.13 (1.30)
	Sex	86 F	78 F	103 F	267 F
		59 M	53 M	45 M	157 M

Table 2.
Univariate ANOVAs to Check the Efficacy of Manipulations.

Items	Independent Variables	<i>df</i>	<i>F</i>	<i>p</i>	Partial η^2
1	The instructions for this survey asked me to answer honestly and accurately.	2	1.23	0.29	0.01
2	I remember what the instructions said for this survey.	2	0.43	0.65	0.00
3	I will find out if my data can be used after I complete this survey.	2	32.98	0.00	0.14
4	I will receive feedback about the quality of my survey responses.	2	53.91	0.00	0.17
5	My survey will be flagged for low quality if I didn't complete this survey carefully.	2	17.71	0.00	0.07
6	Statistical control methods will evaluate the quality of my survey data.	2	15.06	0.00	0.06
7	There was a fading red circle on the upper left side of my survey display.	2	696.31	0.00	0.72
8	There was an animated picture of a person on the upper left side of my survey display.	2	710.02	0.00	0.77
9	There was a computer program monitoring my survey activity through an animated picture of a person.	2	127.04	0.00	0.42
10	There was a person monitoring my survey activity through an animated picture of a person.	2	58.54	0.00	0.27

Note. Estimates reported are based on Type III sum of squares. Boldface indicates value based on Welch's test.

Table 3.

Correlations Among Careless Responding Indicators with Standard Deviations on Diagonal.

	Mahalanobis Distance	Even-Odd Consistency	Maximum LongString	Response Time	Instructed Response Sum	SRSI Indicator
Mahalanobis Distance	18.96					
Even-Odd Consistency	0.40***	0.06				
Maximum LongString	-0.17**	-0.06	5.58			
Response Time	-0.08	-0.06	-0.09	33.34		
Instructed Response Sum	0.04	0.20***	0.11*	-0.02	1.09	
Use Me SRSI	0.00	0.21***	0.34***	-0.16**	0.29***	0.93

Note. Even-Odd consistency was reversed so that larger values indicate more CR. * $p < .05$ ** $p < .01$ *** $p < .001$

Table 4.

Means and Standard Deviations of Dependent Variables by Level of Instructions.

Independent Variable		Dependent Variables											
Instructions	Mahalanobis Distance		Even-Odd Consistency		Maximum LongString		Response Time		Instructed Response Sum		SRSI Indicator		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Normal	60.81	20.7	0.1	0.05	8.57	6.48	51.07	27.6	0.32	0.87	1.68	0.94	
Feedback	59.56	18.91	0.1	0.06	7.91	5.97	60.74	44.91	0.57	1.35	1.72	1.05	
Warning	59.64	16.93	0.1	0.06	7.09	3.85	49.66	28.23	0.39	1.09	1.66	0.82	

Table 5.

Means and Standard Deviations of Dependent Variables by Level of Virtual Presence.

Independent Variable		Dependent Variables										
Instructions	Mahalanobis Distance		Even-Odd Consistency		Maximum LongString		Response Time		Instructed Response Sum		SRSI Indicator	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Normal	60.26	17.5	0.1	0.05	7.92	5.62	54.83	31.02	0.41	1	1.74	0.98
Feedback	58.25	21.15	0.09	0.04	8.36	7.58	50.06	31.42	0.52	1.34	1.7	1.02
Warning	61.52	18.27	0.11	0.07	7.43	2.73	53.99	37.04	0.32	0.92	1.62	0.77

Table 6.
Means and Standard Deviations of Dependent Variables by Condition.

Independent Variables		Dependent Variables											
Virtual Presence	Instructions	Mahalanobis Distance		Even-Odd Consistency		Maximum LongString		Response Time		Instructed Response Sum		SRSI Indicator	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
None	Normal	59.93	18.18	0.09	0.05	8.56	6.43	54.53	27.79	0.46	1.09	1.72	1.01
None	Feedback	60.17	17.93	0.1	0.05	7.33	2.69	59.45	34.29	0.35	0.76	1.64	0.74
None	Warning	60.82	16.51	0.1	0.05	7.39	5.9	51.9	30.31	0.36	1.03	1.84	1.09
Animated Shape	Normal	62.92	24.98	0.09	0.04	9.04	8.27	44.88	20.5	0.29	0.83	1.73	1.07
Animated Shape	Feedback	51.68	17.37	0.08	0.04	10.31	10.51	64.16	44.44	0.76	1.72	1.88	1.36
Animated Shape	Warning	56.57	16.61	0.09	0.05	6.24	2.13	47.52	30.86	0.65	1.54	1.56	0.62
Virtual Human	Normal	59.38	18.25	0.1	0.05	7.98	3	53.87	34.03	0.15	0.44	1.55	0.55
Virtual Human	Feedback	63.98	19.31	0.11	0.08	6.82	2.51	59.55	52.01	0.61	1.43	1.69	1.02
Virtual Human	Warning	61.11	17.48	0.11	0.07	7.52	2.65	49.57	21.36	0.21	0.57	1.61	0.68

Table 7.
Univariate ANOVA follow-up tests to the two-way MANOVA.

Dependent Variables	Independent Variables	<i>df</i>	<i>F</i>	<i>p</i>	Partial η^2
Mahalanobis Distance	Instructions	2	0.54	0.58	0
	Virtual Presence	2	1.69	0.19	0.01
	Interaction	4	2.06	0.09	0.02
Even-Odd Consistency	Instructions	2	0.26	0.77	0
	Virtual Presence	2	3.65	0.03*	0.02
	Interaction	4	0.7	0.59	0.01
Maximum LongString	Instructions	2	2.94	0.05	0.01
	Virtual Presence	2	1.38	0.25	0
	Interaction	4	2.07	0.08	0.02
Response Time	Instructions	2	3.97	0.02*	0.02
	Virtual Presence	2	0.3	0.74	0
	Interaction	4	0.59	0.67	0.01
Instructed Response Sum	Instructions	2	1.95	0.14	0.01
	Virtual Presence	2	1.83	0.16	0.01
	Interaction	4	1.46	0.21	0.01
SRSI Indicator	Instructions	2	0.21	0.81	0
	Virtual Presence	2	0.65	0.52	0
	Interaction	4	0.82	0.51	0.01

Note. Estimates reported are based on Type III sum of squares. * $p < .05$

Please respond to each statement or question as honestly and accurately as you can.

Your answers will be kept strictly confidential.

Click the button below to indicate that you have read these instructions and that you understand them well.

Please rate the extent to which you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Agree	Strongly Agree
The degree to which my organization considers everyone's needs when determining how to treat	☐	☐	☐	☐	☐	☐	☐

Figure 1. Depiction of the animated shape condition. The color of red changed continuously between the two shades depicted.

Your honest and thoughtful responses are important to us and to the study.

To ensure the quality of survey data, your responses will be subject to sophisticated statistical control methods. Responding without much effort will be flagged for low-quality data.

Your answers will be kept strictly confidential.

Click the button below to indicate that you have read these instructions and that you understand them well.

Please rate the extent to which you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Agree	Strongly Agree
I am satisfied with the way my organization gives out money to other groups outside of itself.	☐	☐	☐	☐	☐	☐	☐
I feel good about the procedures my organization uses in determining how to treat other groups outside of itself.	☐	☐	☐	☐	☐	☐	☐
My organization uses fair procedures to decide how to treat other groups outside of itself.	☐	☐	☐	☐	☐	☐	☐

Figure 2. Depiction of the virtual human condition. Minor changes occurred in facial expressions, posture, and blinking for the duration of the survey.

APPENDIX

Statement of the Problem

Advances in technology have spurred the extensive use of Internet-based surveys, assessments, and measures. In academia, Internet-based surveys collect student feedback as a way to improve pedagogy and assess grant-funding processes (Mervis, 2007). In research, job satisfaction (Acquavita, 2009), KSAO assessment (Anderson, 2010), and training (Patrick, 2012) are examples of areas in industrial/organizational psychology that use survey data. In practice, online surveys screen out job applicants (Berta, 2006) and inform administrative recommendations (Marx, 2012). In sum, Internet-based survey data inform knowledge creation in research and justify work activities in many organizations.

Survey administrators are increasingly computerizing paper-and-pencil surveys for several reasons (Barak & English, 2002, p. 70). Compared to paper-and-pencil surveys, Internet-based surveys are more convenient for both respondents and researchers. “An Internet-based test is convenient to construct, revise, distribute, and standardize, and it offers a researcher the opportunity to gather a great deal of data in a relatively short amount of time. This type of test is also convenient to take, use, score, and receive informational feedback for the user” (Barak & English, 2002, p. 74). Increased speed and convenience propel the popularity of Internet-based surveys.

While there are decided advantages to Internet-based surveys, this mode of administration comes with its own set of challenges. For example, respondents cannot ask researchers questions when taking Internet-based surveys. Also, the convenience of being able to respond to a survey wherever and whenever Internet connectivity is available

introduces a variety of settings into the data collection process that, in some circumstances, may influence responses (Hardré, Crowson, & Xie, 2012). Respondents may ignore important parts of the survey, submit multiple times, and may be more apt to respond carelessly (Barak & English, 2002).

The prevalence of survey methodology implies that those using Internet-based surveys to collect data believe that the responses they are collecting are coming from attentive respondents. However, because Internet-based surveys collect data in various places at various times predicated unknown context effects on responses, respondents to this kind of survey may often exhibit what has been described as careless responding (CR; Barak & English, 2002; Berry et al., 1992; Curran, Kotrba, & Denison, 2010; Hardré, Crowson, & Xie, 2012; Johnson, 2005; Meade & Craig, 2012). That is, survey respondents may intentionally or unintentionally respond to the survey in a manner that does not accurately reflect their true scores. Research needs to address this problem because the advantages of Internet-based surveys all but ensure the longevity of this methodology.

The purpose of this study is to examine two approaches that address the problem of CR. The first approach, and primary concern of this study, promises to increase the attentiveness among respondents with better survey design. The second approach involves refining protocols to better estimate this form of responding in the first place.

Understanding facets of Internet-based surveys that tend to encourage attentiveness will increase the quality of data collected. Moreover, inferring from cleaner datasets leads to sounder conclusions. In an indirect way, deterring and identifying CR in Internet-based

surveys improves both research and practice because research conclusions derived from higher quality data supports better theory development and its subsequent application. A potential contribution of this research is to improve Internet-based research methodology by avoiding the problem of CR altogether.

Psychometric Problems Associated with Internet-based Surveying

The problem of careless responding – what is CR? Careless responding is responding with or without regard to item content that reflects inaccuracy rather than respondents' true scores. 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 refers to when one responds to a survey without consideration of item content. Random responding would fall under this category, although some content nonresponsive response patterns 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 indicative of CR as a random response pattern (Johnson, 2005; Meade & Craig, 2012).

How to detect CR. There are two general ways to identify CR. First, when researchers build the survey they can add special survey items that would indicate CR.

Some examples include: self-report items asking the respondents to rate their engagement during the survey, self-report items asking if their survey data is of sufficient quality for research use, and instructed response items 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 is adequate for research use are more effective than other types of self-report items (Meade & Craig, 2012). The present study uses two types of survey items as CR indicators, namely instructed response items and self-report items asking respondents if their data is 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 survey data itself. Effective CR indicators that researchers can derive from survey data include: Mahalanobis distance (Ehlers et al., 2009), Even-Odd consistency (Jackson, 1977), and response patterns such as the same response option selected consecutively, i.e. *LongString* (Johnson, 2005). Ehlers and colleagues (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, large within-person variance on the subsets of even and odd items would indicate CR. The last CR indicator is *LongString*, which is the maximum number of consecutive items with the same

answer choice selected (Johnson, 2005). Although more CR indicators exist, 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. For example, self-reported CR shows low to moderate correlations with other indicators of CR. This means that self-report items likely insufficiently indicate CR when used alone (Meade & Craig, 2012). Therefore, the current study will use the previously mentioned CR indicators because they measure CR more effectively and comprehensively than any single CR indicator. A more detailed discussion of the CR indicators in the current study appears in the *Measures* section.

CR indicators are the criteria assessing effects of survey design interventions. If respondents' values on CR indicators decrease after changing the survey design, then one might conclude that the intervention was effective assuming that lowering CR was the goal. The proposed study investigates two methods of obviating CR in order avoid the detrimental consequences of CR's presence in datasets.

Importance of finding solutions to CR. Finding a solution to CR should be a priority to researchers and practitioners for two reasons. First, both research findings and data-driven decision making heavily rely on raw survey data for justifying decisions and actions based on those decisions. Turnover intentions (Price & Mueller, 1986), job satisfaction (Stanton et al., 2002), and organizational commitment (Mowday, Steers, &

Porter, 1979) are just a few organizational constructs of interest to practitioners that survey data assess. Evidence-based conclusions rely on clean datasets assimilated through research.

Although inaccurate responding is a longstanding issue in survey methodology, few studies examine the quality of responses to surveys beyond typical data cleaning procedures. Among studies that employ Internet-based survey methodology few if any use metrics to estimate CR as a means of screening out inattentive respondents. Recent research is starting to fill this gap by investigating methods, beyond typical data cleaning procedures that identify careless respondents (Meade & Craig, 2012).

In studies that do report prevalence estimates of CR, the values span a wide 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 and colleagues (1992) found that a *majority* of respondents responded carelessly to one or more items. Berry and colleagues (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 careless responding. Depending on the criteria by which they defined inattentive responding to a job satisfaction questionnaire, Curran and colleagues (2010) estimated the rates of careless responding 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 the presence of CR can create psychometric issues. 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 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. Unfortunately, very few studies have explored data-cleaning options specific to Internet-based surveys. 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 is to omit data from certain respondents. Each respondent receives values on CR indicators. If the data from a respondent give extreme values beyond a cutoff score, then further analysis excludes that respondent's data (Tobachnick & Fidell, 2007). Keeping low-quality data is an undesirable alternative to throwing out respondent data altogether. Although necessary, correctly extracting data contributed by careless respondents is a limited solution to CR issues.

Trimming respondents is a reactive approach that, if perfectly executed, can lead to a host of other problems. 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 affects

random sampling and increases method-bound problems of a survey. Therefore, it is imperative to find ways of preventing CR in addition to correctly identifying it after it happens.

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 in paper and pencil surveys. Researchers have posited that less direct interaction between the administrator and participant (Johnson, 2005), more environmental distractions (Carrier, Cheever, Rosen, Benitez, & Chang, 2009), lower participant interest (Schwartz, 1999), and survey length (Berry et al., 1992) may elicit CR in Internet-based surveys. Johnson (2005) suggests the distance between respondent and survey administrator and increased ease of survey submission via the Internet increase the likelihood and prevalence of respondents rushing through surveys. Carefully constructing survey instructions can discourage respondents from rushing through surveys at the expense of data quality. The varied environments in which respondents can complete Internet-based surveys can increase distracters and influence responses (Hardré, Crowson, & Xie, 2012). Increasing the perceived presence of a survey administrator could improve and maintain the ability of respondents to focus on an Internet-based survey. The current study proposes changing Internet-based survey design to identify CR and maintain respondent attention during Internet-based surveys.

How to Make Respondents Pay Attention

Thus far, research in CR and Internet-based surveys has largely focused on identifying and extracting CR. As previously mentioned, the problems of removing careless respondents necessitate the search for preventing CR. The proposed study investigates ways survey instructions and virtual humans can obviate CR and maintain the integrity of the sampling process.

Manipulating instructions. Regardless of the method of administration, surveys often start with instructions that give respondents information about the survey they are about to complete. Survey instructions typically emphasize honesty, accuracy, and anonymity (Huang et al., 2012). Instructions may also give an example of valid or invalid responses to items. These instructions are sometimes referred to as “normal” instructions. While they provide necessary information for survey completion, these instructions may also affect the amount of CR among respondents.

Changing survey instructions may influence the attentiveness of respondents. For example, Huang and colleagues (2012) studied the effects of different instructions on the quality of survey responses. In the control group, respondents read instructions asking for honesty and clarifying that there were no correct or incorrect answers. In what these researchers termed the warning-instructions condition, respondents read that statistical control methods would check the validity of their responses, and that if flagged for insufficient effort, the respondents would not receive participation credit. Respondents who read the warning instructions showed markedly less CR than respondents in the control

group who read the “normal” instructions. Specifically, respondents in the warning group scored higher on indicators of CR. For ease of interpretation with other CR indicators, Huang et al. (2012) reverse scored their *LongString* values so high *LongString* values indicate careful responding. Respondents in the warning group showed higher individual reliability and reverse-scored *LongString* values. In other words, respondents who read the “warning” instructions showed higher reliability and consistency in their answers. Huang and colleagues’ (2012) significant findings show that survey instructions can influence survey responses.

In addition to survey instructions that emphasize honest responding or that warn against not responding carefully, Meade and Craig (2012) investigated the impact of instructions that reduced the respondents’ perceived anonymity. Meade and Craig studied the effects of three types of instruction they described as “normal,” “identified,” or “warning” instructions. “Normal” instructions confirmed the anonymity of responses. ‘Identified’ instructions confirmed the confidentiality of responses and required respondents to enter their names at the bottom of each web page of the survey. ‘Warning’ instructions reminded participants that their responses were subject to the institution’s academic integrity policy. Respondents who read ‘warning’ instructions had to type their name next to a statement to indicate that they had answered all questions with care and honesty before they could move to the next page of the survey. The statement was similar to an integrity statement one might sign after completing an exam. Thus, ‘warning’ instructions implied that responding carelessly would be in violation of the academic

integrity policy (i.e., it would be like cheating on an assignment) and might be subject to university sanctions.

Like Huang and colleagues (2012), Meade and Craig (2012) found that manipulating instructions had a significant main effect on two CR indicators. First, instructions affected the number of times respondents agreed with items when agreement was logically impossible. For example, if a respondent agreed with an item that said leprechauns pay their salaries then a data analyst might reasonably infer CR (Meade & Craig, 2012). Second, the amount of attention respondents reported varied by the type of instructions they read. When instructions required respondents to add their names to each survey page, respondents reported higher levels of attention than the control group that read instructions confirming anonymity. Instructions warning respondents that their answers would be subject to an academic integrity policy had no significant effect on self-reported levels of attention (Meade & Craig, 2012).

In summary, manipulating the content and length of survey instructions influences CR, but the nuances of the relationship between survey instructions and CR remain unknown (Huang et al., 2012; Meade & Craig, 2012). Instruction manipulation impacted CR indicators differently across studies (Huang et al., 2012; Meade & Craig, 2012). Although notable effects came from instructions hinting at punitive repercussions, these instructions also diminished respondents' attitudes toward the survey (Meade & Craig, 2012). A positive approach relative to vague threats of punishment could be beneficial. Meade and Craig proposed that instructions telling respondents that they would receive

feedback about their surveys results might incentivize respondents to respond more attentively (2012). There is evidence that feedback can improve task performance (Northcraft, Schmidt, & Ashford, 2011) and job outcomes (Colarelli, Dean, & Konstans, 1987). It is possible that the beneficial effects feedback can have on job performance might extend to respondent performance on surveys. Promising feedback about the quality of respondents' survey data may increase the attentiveness of respondents. A search through the literature revealed no protocols about feedback following Internet-based surveys. To date, no studies have investigated the effect on performance of survey instructions that include a statement promising respondents that they will receive feedback on the quality of their Internet-based survey responses. The current study predicts that telling respondents that they will receive feedback about the quality of their survey data will influence the level of attentiveness each respondent demonstrates through their response option selections.

Manipulating virtual human presence. In addition to manipulating instructions, research suggests that a second way to impact CR is to include virtual humans in Internet-based surveys to affect the attention of respondents as they complete surveys (Aiello & Svec, 1993; Behrend & Thompson, *in press*; Behrend & Thompson, 2011; Park, 2009; Park & Catrambone, 2007; Rickenberg & Reeves, 2000; Zambaka, Ulinski, Goolkasian, & Hodges, 2004). Virtual humans evolved from avatars, which historically were simply visual markers of users. Now, “users have greater freedom to build not just a graphical marker...but *virtual humans* with distinctive personalities, unique appearance, and

individualized behavioral patterns” (Ahn, 2012, p. 3). In other words, virtual humans bear a closer resemblance to real human beings as computer technology has advanced.

The increasing similarity between humans and virtual humans has enabled virtual humans to elicit the same kinds of social responses from people that one would expect from human-to-human interactions (Gratch, et al., 2007; Haxby, Hoffman, & Gobbini, 2002; Park, 2009). Social interactions involve perceptions and judgments of others as well as relating to and influencing others (Park, 2009). These defining features of social interactions are largely present in exchanges between humans and virtual humans (Nass, Steuer, & Tauber, 1994; Reeves & Nass, 1996).

Adding virtual humans to Internet-based surveys could increase perceived interaction between survey administrators and respondents (Park & Catrambone, 2007). Unlike in-person surveys, Internet-based surveys severely limit interaction between survey administrators and respondents. As a result, there is limited opportunity, if any, for the effects of social influence to hold respondents accountable for choosing responses carefully. The lack of social interaction in Internet-based surveys and the evidence suggesting that virtual humans might simulate social interaction suggest an avenue that survey developers could use to improve the design of Internet-based surveys. Adding virtual humans to Internet-based surveys may reduce CR by increasing the respondents’ perceptions of social interaction with the survey administrators. Seeing a virtual human throughout the survey will likely increase the attention respondents give to completing a survey and may increase

their sense of personal accountability to the process. To the best of my knowledge, however, research has neither confirmed nor rejected these possible outcomes.

Research has shown that task performance may change in the presence of a virtual human similarly to performance changes resulting from the presence of a real human (Aiello & Svec, 1993; Griffith, 1993; Park & Catrambone, 2007; Park, 2009; Rickenberg & Reeves, 2000; Zambaka et al., 2004). Griffith (1993) studied performance on a data entry task that participants completed using a computer. Participants either entered data while alone, electronically monitored by the computer itself, or in the physical presence of a human supervisor. In the electronic monitoring condition, the computer screen displayed messages telling the participant that the computer was monitoring their progress on the task. Griffith found that performance improved when participants were electronically monitored, but the increase did not reach statistical significance. Other researchers found that the presence of virtual humans significantly affected performance on experimental tasks of known difficulty levels, i.e., tasks that can be sorted into easy, medium, and difficult categories based on prior research (Aiello & Svec, 1993; Park & Catrambone, 2007; Park, 2009; Rickenberg & Reeves, 2000; Zambaka et al., 2004). To the best of my knowledge, researchers have not conducted studies examining how the presence of virtual humans influences survey performance. The current study starts a research program that should help to uncover ways of incorporating virtual humans into survey design to improve data collection.

The literature suggests that displaying virtual humans is likely to decrease CR. Likewise, instructions that warn respondents of evaluation are likely to reduce CR (Huang et al., 2012; Meade & Craig, 2012). It follows that combining the presence of a virtual human and a warning of response evaluations may reduce CR to a greater extent than either change individually.

Taken together, the prevalence of CR throughout Internet-based surveys and the incomplete understanding of CR indicators call for research exploring ways to obviate CR. In addition to identifying CR in a survey after it occurs, this study will investigate instructional manipulation and virtual human presence as potential buffers against CR.

Method

Participants

Past estimates of CR effect size range from .02 to .25 (Huang et al., 2012; Meade & Craig, 2012). This study takes a conservative approach by splitting the difference. Using a .10 effect size, G*Power2 software indicated a minimum sample size of 93 will have sufficient power to test the hypotheses proposed in this study. I will draw at least this number of voluntary respondents from a pool of students enrolled in introductory psychology courses at a large Southeastern university. Instructors of introductory psychology courses will recruit respondents by giving students the option of participating in research to fulfill a course requirement. The instructors of respondents' introductory psychology courses will direct them toward a website called Experimentrix that displays available studies. Respondents can opt into this study by selecting it from the list of

available research projects, consenting to participate, and completing an Internet-based survey.

Study Design

This study will use a 3x3 between-subjects experimental design where virtual presence (no presence, animated shape, and virtual human) and type of instruction (normal, warning, and promised feedback) will be the independent variables. Indicators of CR will be the dependent variables. Random assignment will place participants into the control group (no virtual presence with normal instructions) or one of the experimental conditions in which respondents will see some combination of a virtual presence and instruction content on the survey pages. I will use Haptek's PeoplePutty software (Haptek, 2003) to create a virtual human that will look at the participant's activity and display lifelike behaviors such as breathing and blinking. Demographic items will assess the age, sex, ethnicity, and native language of each respondent (see Table 1).

The following hypotheses will be tested:

Hypothesis 1: Respondents who receive instructions warning them that statistical methods will be used to evaluate their responses will score significantly lower on a multivariate composite of CR than respondents who do not receive these instructions.

Hypothesis 2: Respondents who receive instructions that promise that they will be given feedback about the quality of their survey responses will score significantly

lower on a multivariate composite of CR than respondents who do not receive these instructions.

Hypothesis 3: Respondents who see a virtual human displayed on an Internet-based survey will score significantly lower on a multivariate composite of CR as compared to respondents who do not see the virtual human displayed.

Hypothesis 4: Compared to all other conditions, respondents exposed to the combination of a virtual human display and warning instructions will score significantly lower on a multivariate composite of CR. Compared to respondents who do not receive warning instructions or the presence of a virtual human.

Procedure

Participants will receive via e-mail a hyperlink to webpage hosted by the survey administrator after signing up for the study and agreeing to the conditions specified in an informed consent statement. At the webpage I will use a JavaScript routine (see Table 2) to randomly assign respondents to one of a combination of the three levels of virtual presence and three levels of instruction for a total of nine experimental conditions (including the control condition). The JavaScript routine automatically directs respondents to the web page corresponding with the appropriate survey. After completing the survey all respondents will read a note of gratitude for their participation and receive a debriefing immediately after they complete the survey.

Survey Instruction Conditions

The survey will display one of three types of instruction sets described in more detail below. Table 3 contains each of the three instruction conditions.

Normal instruction. Respondents in the anonymous instruction condition will see basic survey instructions that are adapted from Huang and colleagues (2012). Respondents in the anonymous instruction condition will serve as the control group for this manipulation.

Warning instruction. Respondents assigned to the warning instructions condition will be told that while their confidentiality will remain protected, statistical methods will evaluate the quality of their survey responses.

Promised feedback. Respondents assigned to the promise feedback instructions condition will see a message at the beginning of their survey that confirms the confidentiality of their responses but also informs participants that they will receive feedback about the quality of their survey responses upon completion.

Virtual Presence Conditions

No virtual presence. Participants in this condition will serve as the control group, as they will be taking the Internet-based survey under typical conditions, i.e., without any virtual human or geometric shape displayed on screen.

Animated shape presence. In the geometric shape condition, an animated geometric shape will appear at the beginning of the survey and remain visible until

completion. The amount of movement the shape exhibits will roughly match the amount of movement that the virtual human will exhibit.

Virtual human presence. Participants assigned to the third condition will see a virtual human throughout their surveys. In this condition, the virtual human will exhibit movements such as breathing and blinking.

Survey Content

To generate data upon which to test my hypotheses about survey design and CR, I will use a collection of measures to create a lengthy yet realistic survey. Berry and colleagues (1992) found a tendency for CR to occur near the end of long surveys. To ensure CR has an opportunity to take place, numerous survey items will be spread across 6 web pages with 75 items per page. The seven measures that comprise the survey are filler in the sense that their purpose is to lengthen the survey. The survey will contain items that are CR indicators, and such items are critical to testing the hypotheses. Details about all measures appear in the *Measures* section. I will use 7-point Likert scales with drop-down boxes from which participants will select responses. Participants will respond to self-report questions about their awareness of the virtual human's presence to check the manipulation. Qualtrics survey software will administer the Internet-based survey that will span 6 web pages with 75 items per page (see Table 4). Qualtrics settings will require respondents to take the survey in one sitting without the option of saving and returning to the survey later. I will include the following measures to comprise a survey that is long enough to induce CR.

Survey measures to induce CR.

External organizational justice. External Organizational Justice (EOJ) is “[perceptions] by an employee of the degree to which her or his organization behaves fairly, equitably, and ethically when interacting with entities outside of the organization” (Toaddy & Pond, 2012). According to Toaddy and Pond (2012) the EOJ scale has 11 items measuring three types of EOJ: four items measuring distributive external justice, four items measuring procedural external justice, and three items measuring interactional external justice. See Table 5 for the EOJ scale.

Personality. The 300-item International Personality Item Pool (IPIP; Goldberg, 1999) is a measure of the Five Factor Model (McCrae & Costa, 1987) of personality and will be a large part of the survey. See Table 6 for IPIP items. I will also include a 26-item measure of psychopathy (Levenson, Kiehl, & Fitzpatrick, 1995) and a 40-item measure of narcissism (Raskin & Terry, 1988). See Table 7 and Table 8 respectively for the psychopathy and narcissism items.

Social desirability. I will include four social desirability measures in the survey. Social desirability scales contain items that would make the respondent seem incredibly virtuous if he or she endorsed a large number of the items. The survey will include the 33-item Marlowe-Crowne social desirability scale (Crowne & Marlowe, 1960) with an adapted response scale to match the same seven-point response format used for other items in the survey (see Table 9). The second measure of social desirability in the survey will be the

IPIP social desirability scale (Goldberg, 1999). See Table 10 for the IPIP social desirability items.

Measures to test hypotheses. The following six measures assess CR in respondent data. The values of each CR indicator will be used to either support or fail to support the aforementioned hypotheses. See Table 11 for a summary table of all CR indicators in this study.

Instructed-response items. Instructed-response items indicate one response option as the correct response option. An example of an instructed-response item is, “Select ‘strongly disagree’ for this item.” Item interpretation for instructed-response items is literal and unambiguous. Furthermore, the metric is clear for scoring correct or incorrect responses on instructed-response items (Meade & Craig, 2012). Throughout the survey, one instructed response item appears in every segment of 75 items. Adding more instructed-response items could agitate the respondents, potentially resulting in more CR or even prompting them to quit responding to the survey altogether (Meade & Craig, 2012).

Self-reported single item (SRSI) indicator. The survey will include a single item measure adapted from Meade and Craig (2012). Respondents will read instructions that say, “It is vital to our study that we only include responses from people that devoted their full attention to this study. Otherwise, years of effort (the researchers’ and the time of other participants) could be wasted. You will receive credit for this study no matter what” (p. 442). Then respondents will report 1=“yes” or 0=“no” to the following item, “In your

honest opinion, should we use your data in our analyses in this study?” (p. 442). This item will collect self-report data of respondents’ judgments of the quality of their own responses.

Response time. The response time for each respondent will be completion time for the entire survey. Both low and high total response times can indicate CR, meaning there is a non-linear relationship between response time and CR. The minimum and maximum response-time cutoff will be set to 1.5 standard deviations from the mean response time in the sample. Respondents who clock in at times farther than 1.5 standard deviations from the mean most likely responded carelessly. Such extreme completion times will identify respondents for CR.

Outlier analysis. Mahalanobis distance measures will be used to identify CR by calculating each respondent’s distance from the mean. This value is calculated from the series of responses for each scale. The Mahalanobis values that are greater than critical values will flag respondents as likely to have responded carelessly. Frequency distributions will signify appropriate cutoff values. Like Meade and Craig (2012) I will calculate one Mahalanobis distance measure for each of the five personality factors (60 items per factor), giving every respondent five Mahalanobis distance measures. The average of those Mahalanobis distance measures will produce a single Mahalanobis distance value per respondent.

Consistency indicator. Logically, participants who pay attention to the survey should choose equivalent response options across similar items. The *Even-Odd Consistency* measure divides unidimensional scales into two groups based on the order in which items

appear. One group consists of even-numbered items and the other group consists of odd items (Jackson, 1977). This split in the scale creates two subscales and every respondent gets a score for each subscale. A respondents' subscale score for even items is the average of that person's responses on all even items. The average of the person's responses on all odd items is that person's score for the odd subscale. If coefficient alpha values are sufficiently high ($>.70$), I will form subscales from the Narcissism scale (Raskin & Terry, 1988), Psychopathy scale (Levenson et al., 1995), and IPIP scales (Goldberg, 1999). Finally, a within-person correlation from the even and odd subscale scores for the respondent is that respondent's *Even-Odd Consistency* value.

Response pattern. It is possible for a participant to show high levels of consistency across items while responding carelessly, i.e., answering all survey items with the same response option. A legitimate response set consisting of the same response option is unlikely, and in most surveys it is impossible given reverse-worded items. In the proposed study *LongString* values will identify response patterns where participants repeatedly choose the same response option. Johnson (2005) identified 3.5% of response sets as invalid using the *LongString* indicator and recommended it as a way to detect highly consistent, yet inaccurate responding. Meade and Craig (2012) found *Maximum LongString* values effectively screened data for this type of CR. For the proposed study, a Visual Basic for Applications program in Microsoft Excel will compute for each survey web page the maximum number of consecutive items with identical responses. The *Maximum LongString* value for one respondent is the largest *LongString* value found on any of the survey web

pages. Every respondent will get a *Maximum LongString* value based on his or her responses.

Analyses

After collecting the data, I will clean the data using the SAS computer program and subsequently perform a 3 x 3 MANOVA. The between subjects variables will be instruction (normal, warning, or promised feedback) and virtual presence (no presence, animated shape, or virtual human). The dependent variables (instructed response items, SRSI indicator, response time, Mahalanobis distance, *Even-Odd Consistency*, and *Maximum LongString*) will become a multivariate composite of CR. Data analysis will treat CR indicators as continuous variables for two reasons. First, there are numerous values that the CR indicators can take. Second, although some research has explored cutoff scores (Huang et al., 2012; Johnson, 2005), studies still need to verify those values.

Significant main effects and interactions in the MANOVA output will indicate support or lack thereof for the aforementioned hypotheses. The first two hypotheses predict instruction type will reduce the number of respondents showing CR. The direction and significance level of the main effect of instruction on CR will address the first two hypotheses. Hypothesis three predicts that the presence of a virtual human will reduce the number of respondents flagged for CR. The direction and significance of the main effect of virtual human presence on CR will address hypothesis three. The additive effects of instruction and virtual presence will address hypothesis four, specifically that the

combination of virtual human presence and warning instructions will reduce CR significantly more than any other condition

I will follow up the 3 x 3 MANOVA in two ways. First I will conduct one-way ANOVAs. Second, I will conduct a discriminant function analysis to see how the CR indicators discriminate among conditions and which groups are discriminated by different variate functions. By performing these follow-up tests I will capture any underlying dimensions in the dependent variables.

References

- Acquavita, S. P. (2009). Personal and organizational diversity factors' impact on social workers' job satisfaction: Results from a national internet-based survey. *Administration in Social Work*, 33, 151. doi:10.1080/03643100902768824
- Ahn, S. J., Fox, J., & Bailenson, J. N. (2012). Avatars. In Bainbridge, W. S. (Ed.), *Leadership in Science and Technology: A Reference Handbook*. SAGE Publications.
- Aiello, J.R, & Svec, C. M. (1993). Computer monitoring of work performance: Extending the social facilitation framework to electronic presence. *Journal of Applied Social Psychology*, 23, 537-548. doi:10.1111/j.1559-1816.1993.tb01102.x
- Anderson, A. L. F. (2010). Assessment of qualifications needed by environmental health graduates entering private-sector employment. (cover story). *Journal of Environmental Health*, 72(9), 14.
- Barak, A., & English, N. (2002). Prospects and limitations of psychological testing on the internet. *Journal of Technology in Human Services*, 19, 65-89.
- Behrend, T., & Foster Thompson, L. (*in press*). Using animated agents in learner-controlled training: the effects of design control. *The International Journal of Training and Development*.
- Behrend, T., & Foster Thompson, L. (2011). Similarity effects in online training: Effects with computerized trainer agents. *Computers in Human Behavior*, 27, 1201-1206.
- Berry, D. T. R., Wetter, M. W., Baer, R. A., Larsen, L., Clark, C., & Monroe, K. (1992).

- MMPI-2 random responding indices: Validation using a self-report methodology. *Psychological Assessment*, 4, 340-345. doi:10.1037/1040-3590.4.3.340
- Berta, D. (2006, November 13). Caribou Coffee Co. refines hiring process with online applicant-screening program. Penton Media, Inc.
- Carrier, L. M., Cheever, N. A., Rosen, L. D., Benitez, S., & Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of americans. *Computers in Human Behavior*, 25, 483-489. doi:10.1016/j.chb.2008.10.012
- Colarelli, S. M., Dean, R. A., & Konstans, C. (1987). Comparative effects of personal and situational influences on job outcomes of new professionals. *Journal Of Applied Psychology*, 72, 558-566. doi:10.1037/0021-9010.72.4.558
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Curran, P. G., Kotrba, L., & Denison, D. (2010). *CR in surveys: Applying traditional techniques to organizational settings*. Paper presented at the 25th Annual Conference of the Society for Industrial/Organizational Psychology, Atlanta, GA.
- Ehlers, C., Greene-Shortridge, T.M. Weekley, J.A. & Zaiack, M.D. (2009). *The exploration of statistical methods in detecting random responding*. Poster session presented at the annual meeting for the Society for Industrial/Organizational Psychology, Atlanta, GA.

- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. D. Fruyt & F. Ostendorf (Eds.), *Personality Psychology in Europe* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.
- Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M., Van Der Werf, R. J., & Morency, L.-P. (2007). *Can virtual humans be more engaging than real ones?* Paper presented at the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments, Beijing, China.
- Griffith, T. L. (1993). Monitoring and performance: A comparison of computer and supervisor monitoring. *Journal of Applied Social Psychology*, 23, 549-572. doi:10.1111/j.1559-1816.1993.tb01103.x
- Hardré, P. L., Crowson, H. M., & Xie, K. (2012). Examining contexts-of-use for web-based and paper-based questionnaires. *Educational and Psychological Measurement*. doi:10.1177/0013164412451977
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, 51, 59–67. doi:10.1016/S0006-3223(01)01330-0
- Huang, J., Curran, P., Keeney, J., Poposki, E., & DeShon, R. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27, 99–114. doi:10.1007/s10869-011-9231-8.
- Jackson, D. N. (1977). *Jackson Vocational Interest Survey manual*. Port Huron, MI:

- Research Psychologists Press.
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality*, 39, 103-129.
doi:10.1016/j.jrp.2004.09.009
- Levenson, M.R. Kiehl, K.A., & Fitzpatrick, C.M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, 68, 151-158.
- Marx, J. (2012). Nonprofit governance: Improving performance in troubled economic times. *Administration in Social Work*, 36, 40. doi:10.1080/03643107.2010.550670
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90. doi:10.1037/0022-3514.52.1.81.
- Mervis, J. (2007). NSF survey of applicants finds a system teetering on the brink. *Science*, 317, 880-881.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17, 437–455. doi:10.1037/a0028085
- Mowday, R. T., Steers, R. M., Porter, L. W. (1979). The measurement of organizational commitment. *Journal of vocational behavior*, 14, 224–247.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating*

- interdependence*, CHI '94 (pp. 72–78). New York, NY, USA: ACM.
doi:10.1145/191666.191703
- Nichols, D. S., Greene, R. L., & Schmolck, P. (1989). Criteria for assessing inconsistent patterns of item endorsement on the MMPI: Rationale, development, and empirical trials. *Journal of Clinical Psychology*, 45, 239-250. doi:10.1002/1097-4679(198903)45:2<239::AID-JCLP2270450210>3.0.CO;2-1
- Northcraft, G. B., Schmidt, A. M., & Ashford, S. J. (2011). Feedback and the rationing of time and effort among competing tasks. *Journal of Applied Psychology*, 96, 1076-1086. doi:10.1037/a0023221
- Park, S. (2009). *Social Responses to Virtual Humans: The Effect of Human-Like Characteristics* (Unpublished doctoral dissertation). Georgia Institute of Technology, Atlanta, Georgia.
- Park, S., & Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49, 1054-1060. doi:10.1518/001872007X249910
- Patrick, J. (2012). Being in one's chosen job determines pre-training attitudes and training outcomes. *Journal of Occupational & Organizational Psychology*, 85, 245. doi:10.1111/j.2044-8325.2011.02027.x
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H.

- I. Price, J. L., & Mueller, C. W. (1986). *Handbook of organizational measurement*. Marshfield, MA: Pitman.
- Price, J. L., & Mueller, C. W. (1986). *Handbook of organizational measurement*. doi: 10.1108/01437729710182260
- Raskin, R., & Terry, H. (1988). A principal-components analysis of the narcissistic personality inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology*, 54, 890-902.
- Reeves, B., & Nass, C. I. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. Stanford, California: Cambridge University Press. Retrieved from <http://catalog.lib.ncsu.edu/record/NCSU923425>
- Rickenberg, R., & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '00 (pp. 49–56). New York, NY, USA: ACM. doi:10.1145/332040.332406
- Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, 9, 367–373. doi:10.1177/014662168500900405
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93–105. doi:10.1037/0003-066X.54.2.93
- Stanton, J. M., Sinar, E. F., Balzer, W. K., Julian, A. L., Thoresen, P., Aziz, S., Fisher,

- G.G., & Smith, P.C. (2002). Development of a compact measure of job satisfaction: The abridged Job Descriptive Index. *Educational and Psychological Measurement*, 62, 173-191.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.
- Toaddy, S. R., & Pond, S. B. (2012). *Validation of a Measure of External Organizational Justice*. (Unpublished doctoral dissertation). North Carolina State University, Raleigh, North Carolina.
- Woods, C. (2006). CR to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, 28, 186–191. doi:10.1007/s10862-005-9004-7
- Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2004). Effects of virtual human presence on task performance. *Proceedings of International Conference on Artificial Reality and Telexistence (ICAT)* (pp. 174–181).

Table 1.

Demographic items

Items
What is your age?
What is your sex?
What is your grade level?
Freshman
Sophomore
Junior
Senior
Other
If you selected other in the question above, please specify your grade level here.
What is your ethnicity?
African and/or African American
Asian and/or Asian American
Caucasian and/or European American
Hispanic
Native American and/or Alaskan native
Native Hawaiian and/or Pacific Islander
Other
If you selected other in the question above, please specify your ethnicity here.
What is your nationality?
Is English your first language?
If English is not your first language, please indicate what is.

Table 2.

JavaScript routine that will randomly assign respondents to an experimental condition.

```
// JavaScript Document
```

```
<!-- Paste this code into an external JavaScript file named: feelLucky.js -->
```

```
/* This script and many more are available free online at  
The JavaScript Source :: http://javascript.internet.com  
Created by: Will Bontrager :: http://www.bontragerconnection.com/ */
```

```
// Leave next line as is.  
var Lucky = new Array();
```

```
// The cookie to keep track of which "lucky" destinations  
// have already been visited needs a name. Okay to change  
// the cookie name.
```

```
var FeelLuckyCookieName = "Already_Visited";
```

```
// When a "lucky" destination has been decided upon, shall  
// the browser open a new window with the destination URL?  
// (Specify "y" or "yes" if yes new window; otherwise "").  
var NewWindow = "";
```

```
// Specify your lucky destination URLs here. The first is  
// assigned to Lucky[0], the next to Lucky[1], and so  
// forth, in numerical sequence -- as many as you want.
```

```
Lucky[0] = "http://surveylinktoNovirtualpresenceNormalinstructions";  
Lucky[1] = "http://surveylinktoNovirtualpresenceFeedbackinstructions";  
Lucky[2] = "http://surveylinktoNovirtualpresenceWarninginstructions";  
Lucky[3] = "http://surveylinktoShapevirtualpresenceNormalinstructions";  
Lucky[4] = "http://surveylinktoShapevirtualpresenceFeedbackinstructions";  
Lucky[5] = "http://surveylinktoShapevirtualpresenceWarninginstructions";  
Lucky[6] = "http://surveylinktoVirtualhumanpresenceNormalinstructions";  
Lucky[7] = "http://surveylinktoVirtualhumanpresenceFeedbackinstructions";  
Lucky[8] = "http://surveylinktoVirtualhumanpresenceWarninginstructions";
```

```
// No additional JavaScript customizations are required. //
```

```
var TabChar = String.fromCharCode(9);
```

```

var CurrentCookie = "";
function GetLuckyCookie() {
    var cookiecontent = "";
    if(document.cookie.length > 0) {
        var cookiename = FeelLuckyCookieName + '=';
        var cookiebegin = document.cookie.indexOf(cookiename);
        var cookieend = 0;
        if(cookiebegin > -1) {
            cookiebegin += cookiename.length;
            cookieend = document.cookie.indexOf(";",cookiebegin);
            if(cookieend < cookiebegin) { cookieend = document.cookie.length; }
            cookiecontent = document.cookie.substring(cookiebegin,cookieend);
        }
    }
    return cookiecontent;
}

function PutLuckyCookie(value) {
    if(CurrentCookie.length > 0) { value = CurrentCookie + TabChar + value; }
    value = escape(value);
    document.cookie = FeelLuckyCookieName + "=" + value;
}

function YesMakeMeLucky() {
    CurrentCookie = GetLuckyCookie();
    CurrentCookie = unescape(CurrentCookie);
    if(CurrentCookie == '.') { CurrentCookie = ""; }
    var LuckyVisitedList = CurrentCookie.split(TabChar);
    if(LuckyVisitedList.length >= Lucky.length) {
        document.cookie = FeelLuckyCookieName + "=". ";";
        CurrentCookie = "";
        LuckyVisitedList = Array();
    }
    for(var i = 0; i < LuckyVisitedList.length; i++) { Lucky[LuckyVisitedList[i]] = "; }
    var LuckyL = new Array();
    for(var i = 0; i < Lucky.length; i++) {
        if(Lucky[i].length > 0) { LuckyL.push(" + i + TabChar + Lucky[i]); }
    }
    var LuckyDestinationNumber = 0;
    if(LuckyL.length > 1) { LuckyDestinationNumber = Math.ceil((Math.random() *
LuckyL.length) - 1); }
    var LuckyNumberPlace = new Array();

```

```
LuckyNumberPlace = LuckyL[LuckyDestinationNumber].split(TabChar);  
PutLuckyCookie(LuckyNumberPlace[0]);  
NewWindow = NewWindow.toLowerCase();  
if(NewWindow.substr(0,1) == "y") { window.open(LuckyNumberPlace[1]); }  
else { document.location = LuckyNumberPlace[1]; }  
}
```

Table 3.
Summary of Instruction Conditions

Condition	Instructions Displayed
Normal	“There are no correct or incorrect answers on this survey. Please respond to each statement or question as honestly and accurately as you can. Your answers will be kept strictly anonymous and confidential.”
Warning	“Your honest and thoughtful responses are important to us and to the study. To ensure the quality of survey data, your responses will be subject to sophisticated statistical control methods. Responding without much effort will be flagged for low-quality data. Your answers will be kept strictly anonymous and confidential.”
Promised Feedback	“Your honest and thoughtful responses are important to us and to the study. You will receive feedback about the quality of your survey responses and whether we can use the information that you provided to us upon completion of the survey. Your answers will be kept strictly anonymous and confidential.”

Table 4.

Section of the survey respondents will complete showing an instructed-response item.

I would do almost anything on a dare.
<input type="text" value="Neither Agree nor Disagree"/>
Select "Agree" for this item.
<div><div><div>Strongly Disagree</div><div>Disagree</div><div>Somewhat Disagree</div><div>Neither Agree nor Disagree</div><div>Somewhat Agree</div><div>Agree</div><div>Strongly Agree</div></div></div>
I wish somebody would someday write my biography.
<input type="text"/>
I get upset when people don't notice how I look when I go out in public.
<input type="text"/>
I am more capable than other people.
<input type="text"/>
I am an extraordinary person.

Table 5.
External Organizational Justice (EOJ) scale.

Constructs	Items
Distributive External Organizational Justice	<p>I am satisfied with the way my organization gives out money to other groups outside of itself.</p> <p>I feel good about the way my organization gives out money to other groups outside of itself.</p> <p>I feel good about the way my organization distributes resources to other groups outside of itself.</p> <p>My organization gives out money to other groups outside of itself justly.</p>
Procedural External Organizational Justice	<p>My organization uses fair procedures to decide how to treat other groups outside of itself.</p> <p>The degree to which my organization considers everyone's needs when determining how to treat other groups outside of itself is just.</p> <p>I feel good about the procedures my organization uses in determining how to treat other groups outside of itself.</p> <p>I feel good about the policies that my organization has in place to determine how to treat other groups outside of itself.</p>
Interactional External Organizational Justice	<p>I feel good about the amount of honesty that my organization displays when interacting with other groups outside of itself.</p> <p>I am satisfied with the way my organization gives explanations for its actions to outside groups.</p> <p>I feel good about the way my organization gives explanations for its actions to outside groups.</p>

Note. Converted to a 7-point Likert-type response scale from the original 5-point Likert-type scale with anchors at "Strongly Disagree" and "Strongly Agree".

Table 6.
International Personality Item Pool (IPIP) scale.

Facets	Items	Scoring
Anxiety	Worry about things.	
	Fear for the worst.	
	Am afraid of many things.	
	Get stressed out easily.	
	Get caught up in my problems.	
	Am not easily bothered by things.	
	Am relaxed most of the time.	Reverse
	Am not easily disturbed by events.	Reverse
	Don't worry about things that have already happened.	Reverse
	Adapt easily to new situations.	Reverse
Anger	Get angry easily.	
	Get irritated easily.	
	Get upset easily.	
	Am often in a bad mood.	
	Lose my temper.	
	Rarely get irritated.	
	Seldom get mad.	Reverse
	Am not easily annoyed.	Reverse
	Keep my cool.	Reverse
	Rarely complain.	Reverse
Depression	Often feel blue.	
	Dislike myself.	
	Am often down in the dumps.	
	Have a low opinion of myself.	
	Have frequent mood swings.	
	Feel desperate.	
	Feel that my life lacks direction.	
	Seldom feel blue.	
	Feel comfortable with myself.	Reverse
	Am very pleased with myself.	Reverse
Self-Consciousness	Am easily intimidated.	
	Am afraid that I will do the wrong thing.	
	Find it difficult to approach others.	
	Am afraid to draw attention to myself.	
	Only feel comfortable with friends.	

	Stumble over my words.	
	Am not embarrassed easily.	
	Am comfortable in unfamiliar situations.	Reverse
	Am not bothered by difficult social situations.	Reverse
	Am able to stand up for myself.	Reverse
Immoderation	Often eat too much.	
	Don't know why I do some of the things I do.	
	Do things I later regret.	
	Go on binges.	
	Love to eat.	
	Rarely overindulge.	
	Easily resist temptations.	Reverse
	Am able to control my cravings.	Reverse
	Never spend more than I can afford.	Reverse
	Never splurge.	Reverse
Vulnerability	Panic easily.	
	Become overwhelmed by events.	
	Feel that I'm unable to deal with things.	
	Can't make up my mind.	
	Get overwhelmed by emotions.	
	Remain calm under pressure.	
	Can handle complex problems.	Reverse
	Know how to cope.	Reverse
	Readily overcome setbacks.	Reverse
	Am calm even in tense situations.	Reverse
Friendliness	Make friends easily.	
	Warm up quickly to others.	
	Feel comfortable around people.	
	Act comfortably with others.	
	Cheer people up.	
	Am hard to get to know.	
	Often feel uncomfortable around others.	Reverse
	Avoid contacts with others.	Reverse
	Am not really interested in others.	Reverse
	Keep others at a distance.	Reverse
Gregariousness	Love large parties.	
	Talk to a lot of different people at parties.	

Assertiveness	Enjoy being part of a group.	
	Involve others in what I am doing.	
	Love surprise parties.	
	Prefer to be alone.	
	Want to be left alone.	Reverse
	Don't like crowded events.	Reverse
	Avoid crowds.	Reverse
	Seek quiet.	Reverse
	Take charge.	
	Try to lead others.	
Activity Level	Can talk others into doing things.	
	Seek to influence others.	
	Take control of things.	
	Wait for others to lead the way.	Reverse
	Keep in the background.	Reverse
	Don't like to draw attention to myself.	Reverse
	Hold back my opinions.	Reverse
	Am always busy.	
	Am always on the go.	
	Do a lot in my spare time.	
Excitement-Seeking	Like to take it easy.	Reverse
	Like to take my time.	Reverse
	Like a leisurely lifestyle.	Reverse
	Let things proceed at their own pace.	Reverse
	React slowly.	Reverse
	Love excitement.	
	Seek adventure.	
	Love action.	
	Enjoy being part of a loud crowd.	
	Enjoy being reckless.	
Cheerfulness	Act wild and crazy.	
	Would never go hang gliding or bungee jumping.	Reverse
	Dislike loud music.	Reverse
	Radiate joy.	
	Have a lot of fun.	
	Express childlike joy.	
	Laugh my way through life.	
	Love life.	

Imagination	Look at the bright side of life.	
	Amuse my friends.	
	Am not easily amused.	Reverse
	Seldom joke around.	Reverse
	Have a vivid imagination.	
	Enjoy wild flights of fantasy.	
	Love to daydream.	
	Like to get lost in thought.	
	Indulge in my fantasies.	
	Spend time reflecting on things.	
Artistic Interest	Seldom daydream.	Reverse
	Do not have a good imagination.	Reverse
	Seldom get lost in thought.	Reverse
	Have difficulty imagining things.	Reverse
	Believe in the importance of art.	
	Like music.	
	See beauty in things others might not notice.	
	Love flowers.	
	Enjoy the beauty of nature.	
	Do not like art.	Reverse
Emotionality	Do not like poetry.	Reverse
	Do not enjoy going to art museums.	Reverse
	Do not like concerts.	Reverse
	Do not enjoy watching dance performances.	Reverse
	Experience my emotions intensely.	
	Feel others' emotions.	
	Am passionate about causes.	
	Enjoy examining myself and my life.	
	Try to understand myself.	
	Seldom get emotional.	Reverse
Adventurousness	Am not easily affected by my emotions.	Reverse
	Rarely notice my emotional reactions.	Reverse
	Experience very few emotional highs and lows.	Reverse
	Don't understand people who get emotional.	Reverse
	Prefer variety to routine.	
	Like to visit new places.	

Intellect	Am interested in many things.	
	Like to begin new things.	
	Prefer to stick with things that I know.	Reverse
	Dislike changes.	Reverse
	Don't like the idea of change.	Reverse
	Am a creature of habit.	Reverse
	Dislike new foods.	Reverse
	Am attached to conventional ways.	Reverse
	Like to solve complex problems.	
	Love to read challenging material.	
	Have a rich vocabulary.	
	Can handle a lot of information.	
	Enjoy thinking about things.	
	Am not interested in abstract ideas.	Reverse
	Avoid philosophical discussions.	Reverse
Liberalism	Have difficulty understanding abstract ideas.	Reverse
	Am not interested in theoretical discussions.	Reverse
	Avoid difficult reading material.	Reverse
	Tend to vote for liberal political candidates.	
	Believe that there is no absolute right and wrong.	
	Believe that criminals should receive help rather than punishment.	
	Believe in one true religion.	Reverse
	Tend to vote for conservative political candidates.	Reverse
	Believe that too much tax money goes to support artists.	Reverse
	Believe laws should be strictly enforced.	Reverse
Trust	Believe that we coddle criminals too much.	Reverse
	Believe that we should be tough on crime.	Reverse
	Like to stand during the national anthem.	Reverse
	Trust others.	
	Believe that others have good intentions.	
	Trust what people say.	
	Believe that people are basically moral.	
	Believe in human goodness.	
	Think that all will be well.	

Morality	Distrust people.	Reverse
	Suspect hidden motives in others.	Reverse
	Am wary of others.	Reverse
	Believe that people are essentially evil.	Reverse
	Would never cheat on my taxes.	
	Stick to the rules.	
	Use flattery to get ahead.	Reverse
	Use others for my own ends.	Reverse
	Know how to get around the rules.	Reverse
	Cheat to get ahead.	Reverse
	Put people under pressure.	Reverse
	Pretend to be concerned for others.	Reverse
	Take advantage of others.	Reverse
	Obstruct others' plans.	Reverse
Altruism	Make people feel welcome.	
	Anticipate the needs of others.	
	Love to help others.	
	Am concerned about others.	
	Have a good word for everyone.	
	Look down on others.	Reverse
	Am indifferent to the feelings of others.	Reverse
	Make people feel uncomfortable.	Reverse
	Turn my back on others.	Reverse
	Take no time for others.	Reverse
Cooperation	Am easy to satisfy.	
	Can't stand confrontations.	
	Hate to seem pushy.	
	Have a sharp tongue.	Reverse
	Contradict others.	Reverse
	Love a good fight.	Reverse
	Yell at people.	Reverse
	Insult people.	Reverse
	Get back at others.	Reverse
	Hold a grudge.	Reverse
Modesty	Dislike being the center of attention.	
	Dislike talking about myself.	
	Consider myself an average person.	
	Seldom toot my own horn.	
	Believe that I am better than others.	Reverse

Sympathy	Think highly of myself.	Reverse
	Have a high opinion of myself.	Reverse
	Know the answers to many questions.	Reverse
	Boast about my virtues.	Reverse
	Make myself the center of attention.	Reverse
	Sympathize with the homeless.	
	Feel sympathy for those who are worse off than myself.	
	Value cooperation over competition.	
	Suffer from others' sorrows.	
	Am not interested in other people's problems.	Reverse
	Tend to dislike soft-hearted people.	Reverse
	Believe in an eye for an eye.	Reverse
	Try not to think about the needy.	Reverse
	Believe people should fend for themselves.	Reverse
Self-Efficacy	Can't stand weak people.	Reverse
	Complete tasks successfully.	
	Excel in what I do.	
	Handle tasks smoothly.	
	Am sure of my ground.	
	Come up with good solutions.	
	Know how to get things done.	
	Misjudge situations.	Reverse
	Don't understand things.	Reverse
	Have little to contribute.	Reverse
	Don't see the consequences of things.	Reverse
Orderliness	Like order.	
	Like to tidy up.	
	Want everything to be "just right."	
	Love order and regularity.	
	Do things according to a plan.	
	Often forget to put things back in their proper place.	Reverse
	Leave a mess in my room.	Reverse
	Leave my belongings around.	Reverse
	Am not bothered by messy people.	Reverse
	Am not bothered by disorder.	Reverse
Dutifulness	Try to follow the rules.	
	Keep my promises.	

	Pay my bills on time.	
	Tell the truth.	
	Listen to my conscience.	
	Break rules.	Reverse
	Break my promises.	Reverse
	Get others to do my duties.	Reverse
	Do the opposite of what is asked.	Reverse
	Misrepresent the facts.	Reverse
Achievement- Striving	Go straight for the goal.	
	Work hard.	
	Turn plans into actions.	
	Plunge into tasks with all my heart.	
	Do more than what's expected of me.	
	Set high standards for myself and others.	
	Demand quality.	
	Am not highly motivated to succeed.	Reverse
	Do just enough work to get by.	Reverse
	Put little time and effort into my work.	Reverse
Self-Discipline	Get chores done right away.	
	Am always prepared.	
	Start tasks right away.	
	Get to work at once.	
	Carry out my plans.	
	Find it difficult to get down to work.	Reverse
	Waste my time.	Reverse
	Need a push to get started.	Reverse
	Have difficulty starting tasks.	Reverse
	Postpone decisions.	Reverse
Cautiousness	Avoid mistakes.	
	Choose my words with care.	
	Stick to my chosen path.	
	Jump into things without thinking.	Reverse
	Make rash decisions.	Reverse
	Like to act on a whim.	Reverse
	Rush into things.	Reverse
	Do crazy things.	Reverse
	Act without thinking.	Reverse
	Often make last-minute plans.	Reverse

Note. Items marked “reverse” will be reverse scored.

Table 7.
Psychopathy scale.

Scales	Items
Primary Characteristics	<p>Success is based on survival of the fittest; I am not concerned about the losers.</p> <p>For me, what's right is whatever I can get away with.</p> <p>In today's world, I feel justified in doing anything I can get away with to succeed.</p> <p>My main purpose in life is getting as many goodies as I can.</p> <p>Making a lot of money is my most important goal.</p> <p>I let others worry about higher values; my main concern is with the bottom line.</p> <p>People who are stupid enough to get ripped off usually deserve it.</p> <p>Looking out for myself is my top priority.</p> <p>I tell other people what they want to hear so that they will do what I want them to do.</p> <p>I would be upset if my success came at someone else's expense.</p> <p>I often admire a really clever scam.</p> <p>I make a point of trying not to hurt others in pursuit of my goals.</p> <p>I enjoy manipulation other people's feelings.</p> <p>I feel bad if my words or actions cause someone else to feel emotional pain.</p> <p>Even if I were trying very hard to sell something, I wouldn't lie about it.</p> <p>Cheating is not justified because it is unfair to others.</p>
Secondary Characteristics	<p>I find myself in the same kinds of trouble, time after time.</p> <p>I am often bored.</p> <p>I find that I am able to pursue one goal for a long time.</p> <p>I don't plan anything very far in advance.</p> <p>I quickly lose interest in tasks I start.</p> <p>Most of my problems are due to the fact that other people just don't understand me.</p> <p>Before I do anything, I carefully consider the possible consequences.</p> <p>I have been in a lot of shouting matches with other people.</p> <p>When I get frustrated, I often "let off steam" by blowing my top.</p> <p>Love is overrated.</p>

Table 8.
Narcissism scale.

Items
I can usually talk my way out of anything.
I like to be the center of attention.
I will be a success.
I think I am a special person.
I see myself as a good leader.
I am assertive.
I like to have authority over other people.
I find it easy to manipulate people.
I have a natural talent for influencing people.
I insist upon getting the respect that is due me.
I like to display my body.
I can read people like a book.
I like to take responsibility for making decisions.
I want to amount to something in the eyes of the world.
I like to look at my body.
I am apt to show off if I get the chance.
I always know what I am doing.
Modesty doesn't become me.
I rarely depend on anyone else to get things done.
Everybody likes to hear my stories.
I expect a great deal from other people.
I will never be satisfied until I get all that I deserve.
I like to be complimented.
I have a strong will to power.
I like to start new fads and fashions.
I like to look at myself in the mirror.
I really like to be the center of attention.
I can live my life in any way I want to.
People always seem to recognize my authority.
I would prefer to be a leader.
I am going to be a great person.
I can make anybody believe anything I want them to.
I would do almost anything on a dare.
I am a born leader.
I wish somebody would someday write my biography.
I get upset when people don't notice how I look when I go out in public.

I am more capable than other people.

I am an extraordinary person.

I know that I am good because everybody keeps telling me so.

If I ruled the world it would be a much better place.

Table 9.
Marlowe-Crowne social desirability scale.

Items	Socially-desirable responses
Before voting I thoroughly investigate the qualifications of all the candidates.	Agreement
I never hesitate to go out of my way to help someone in trouble.	Agreement
It is sometimes hard for me to go on with my work if I am not encouraged.	Disagreement
I have never intensely disliked anyone.	Agreement
On occasion I have had doubts about my ability to succeed in life.	Disagreement
I sometimes feel resentful when I don't get my way.	Disagreement
I am always careful about my manner of dress.	Agreement
My table manners at home are as good as when I eat out in a restaurant.	Agreement
If I could get into a movie without paying and be sure I was not seen I would probably do it.	Disagreement
On a few occasions, I have given up doing something because I thought too little of my ability.	Disagreement
I like to gossip at times.	Disagreement
There have been times when I felt like rebelling against people in authority even though I knew they were right.	Disagreement
No matter who I'm talking to, I'm always a good listener.	Agreement
I can remember "playing sick" to get out of something.	Disagreement
There have been occasions when I took advantage of someone.	Disagreement
I'm always willing to admit it when I make a mistake.	Agreement
I always try to practice what I preach.	Agreement
I don't find it particularly difficult to get along with loud mouthed, obnoxious people.	Agreement
I sometimes try to get even rather than forgive and forget.	Disagreement
When I don't know something I don't at all mind admitting it.	Agreement
I am always courteous, even to people who are disagreeable.	Agreement
At times I have really insisted on having things my own way.	Disagreement
There have been occasions when I felt like smashing things.	Disagreement
I would never think of letting someone else be punished for my wrong- doings.	Agreement
I never resent being asked to return a favor.	Agreement

I have never been irked when people expressed ideas very different from my own.	Agreement
I never make a long trip without checking the safety of my car.	Agreement
There have been times when I was quite jealous of the good fortune of others.	Disagreement
I have almost never felt the urge to tell someone off.	Agreement
I am sometimes irritated by people who ask favors of me.	Disagreement
I have never felt that I was punished without cause.	Agreement
I sometimes think when people have a misfortune they only got what they deserved.	Disagreement
I have never deliberately said something that hurt someone's feelings.	Agreement

Note. The original scale items were set up on the true or false response options. In this study the item responses will be on a 7-point Likert scale from "Strongly Disagree" to "Strongly Agree". Items marked with "disagreement" in the table will be reverse scored so larger values correspond with a higher degree of socially desirable responding.

Table 10.
IPIP Social Desirability Scale.

Factors	Items
Unlikely Virtues	Always admit it when I make a mistake. Never give up hope. Know that anyone who tries can get a job. Always know why I do things. Never give up. Know immediately what to do. Believe there is never an excuse for lying. Always know what I am doing. Am always ready to start afresh. Have never engaged in gossip. Will do anything for others. Am always prepared. Don't always practice what I preach. Have some bad habits. Have sometimes had to tell a lie. Am not always honest with myself. Am not always what I appear to be.
Impression Management	Would never take things that aren't mine. Would never cheat on my taxes. Believe there is never an excuse for lying. Always admit it when I make a mistake. Rarely talk about sex. Return extra change when a cashier makes a mistake. Try to follow the rules. Easily resist temptations. Tell the truth. Rarely overindulge. Have sometimes had to tell a lie. Use swear words. Use flattery to get ahead. Am not always what I appear to be. Break rules. Cheat to get ahead.

Don't always practice what I preach.

Misuse power.

Get back at others.

Am likely to show off if I get the chance.

Table 11.

Summary of CR Indicators.

Indicator	Description
Total Instructed Response	Sum of six survey items that direct respondents to select a “correct” response option.
SRSI Indicator	Self-reported single item 7-point agreement scale as to whether respondents feel their data should be used for analysis.
Response Time	Total time to complete the survey.
Mahalanobis D	Multivariate distance between respondent’s response vector and the vector of sample means.
Even Odd Consistency	Within-person correlation across subscales formed by even-odd split of unidimensional scales.
Max LongString	Maximum of LongString values. LongString is the maximum number, i.e. string, of identical consecutive responses on a webpage.

Note. Adapted from Meade and Craig (2012).