St. Cloud State University

The Repository at St. Cloud State

11-2013

The Meaning of Response Latency in Personality Assessment: Exploring the Impact of Faking, Difficulty, and Social Desirability

Mark A. Roebke

THE MEANING OF RESPONSE LATENCY IN PERSONALITY ASSESSMENT:

EXPLORING THE IMPACT OF FAKING, DIFFICULTY, AND

SOCIAL DESIRABILITY

by

Mark A. Roebke

B.S., North Dakota State University, 2009

14000209

A Thesis

Submitted to the Graduate Faculty

of

St. Cloud State University

in Partial Fulfillment of the Requirements

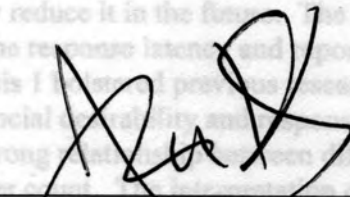for the Degree

Master of Science

St. Cloud, Minnesota

November, 2013

This thesis submitted by Mark A. Roebke in partial fulfillment of the requirements for the Degree of Master of Science at St. Cloud State University is hereby approved by the final evaluation committee.
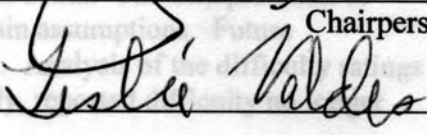
_____

_____ Chairperson

_____

Month   Year

_____

Patricia Hughes
Dean
School of Graduate Studies

THE MEANING OF RESPONSE LATENCY IN PERSONALITY ASSESSMENT: EXPLORING THE IMPACT OF FAKING, DIFFICULTY, AND SOCIAL DESIRABILITY
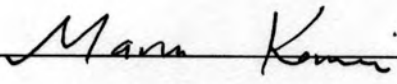
Mark A. Roebke

This literature review has shown that people are distorting their answers during personality assessments in varying contexts. This distortion arises as both a function of items and individual differences. While it is easy to identify, understanding the underlying processes of faking is of vital importance to the field. Investigators have identified response latency as one of the better tools available to aid in understanding faking. By using this method, it is possible to analyze differences in faking by items and situation. Understanding differences in both items and condition should help expand our understanding of faking and ultimately reduce it in the future. The current study examines how social desirability relates to the response latency and reported response difficulty of personality items. Hypothesis 1 bolstered previous research demonstrating a non-linear relationship between social desirability and response latency. An analysis of covariance revealed the strong relationship between difficulty and response latency when controlling for character count. The interpretation of hypothesis 2 is very difficult to make without certain assumptions. Future investigation will be required to explain this effect. Analysis of the difficulty ratings showed that when I told subjects to answer honestly, reported difficulty no longer covaried with social desirability.
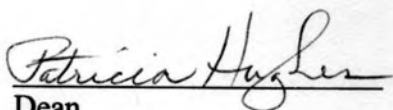
Month	Year	Approved by Research Committee:
		John Kulas Chairperson

ACKNOWLEDGMENTS

I would like to thank my wife, Kayla, for all of her love and support during this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Self-ratings in personality assessment implicitly assume that if an assessor asks a person a question, they will be willing and able to accurately answer it. Self-report questionnaires are commonly used to measure personality because psychologists have accepted the idea that they can assess psychological constructs, record the answers, and build a score that would allow them to interpret the results in a meaningful way (Wagner-Menghin, 2006). However, some researchers acknowledge the possibility that respondents may distort responses.

When individuals respond to questionnaire items in a systematic fashion based on something other than the specific item content, they are demonstrating response bias (Paulhus, 1991). Paulhus (1991) considered response bias, "…a disturbing issue in psychological assessment" (p. 17) that can distort accurate content measurement. As research on the measurement of personality expanded, questions about validity and the potential for distortion moved to the forefront. To address this, a panel of former editors from *Personnel Psychology* and *Journal of Applied Psychology* came together at the 2004 conference for the Society for Industrial and Organizational Psychology to discuss the use of personality assessments in employment settings. In an article derived from that discussion, the authors determined that a series of substantial issues

exist related to the accuracy and use of self-report personality tests (Morgeson et al., 2007).

Morgeson et al. (2007) expressed concern with "faking" on personality tests and how falsification affected assessment outcomes. Understanding the concept of faking is an important first step in determining the sources of response inaccuracy. Faking has been conceptualized in many ways such as self-enhancement, socially desirable responding, self-deception, and impression management (Ellingson, Sackett, & Connelly, 2007; Hogan, Barrett, & Hogan, 2007; Holden & Passey, 2010; Ones, Reiss, & Viswesvaran 1996; Paulhus 1986). Utilizing the various labels that categorize faking allows researchers to explore different operationalizations of the construct. Unfortunately, the lack of clear terminology has also lead to inconclusive or conflicting findings (Griffith, Malm, English, Yoshita, & Gujar, 2006). To resolve this, Griffith and McDaniel (2006) suggest that the common theme throughout these manifestations is the enhancement of personality inventory scores under motivated conditions. Participants may know their true answer to an item, but modify their reported response to attain a more positive score on the latent construct that they perceive the item is measuring. This elevation is the source of concern that drives much of the research on response distortion. Despite a lack of clear operational definition of faking (Zickar, Gibby, & Robie, 2004), researchers have sought to determine whether psychological assessments are susceptible to response manipulation.

In an effort to better understand this intentional form of distortion, researchers have made numerous attempts to induce faking on assessments. The procedure used by studies that assess whether faking is possible generally employ a controlled experiment design where a researcher instructs subjects to fake-good and then compares scale scores with those obtained through instructions to respond honestly (Mahar, Cologon, & Duck, 1995; Tett et al., 2006). Thus, as Griffith et al. (2006) propose, faking can be defined as a difference between scale scores obtained across response conditions. Assessment in an unmotivated condition provides an approximation of a true score and the observed deviation from that score under motivated conditions (e.g., selection) represents faking (Griffith et al., 2006). Addressing one of the key questions of Morgeson et al's (2007) discussion, the authors reviewed almost 40 studies finding that scores in every study were susceptible to response distortion. They identified that many of the most commonly employed personality assessments can be manipulated, such as the MMPI, Eysenck Personality Inventory, Sixteen Personality Factor Questionnaire, Edwards Personality Inventory, and the California Psychological Inventory (Morgeson et al., 2007). Conducting a meta-analysis on over 50 studies, Viswesvaran and Ones (1999) confirmed that individuals prove able to fake responses to personality inventories and that it is possible across all of the Big Five personality factors. With resounding literary support that shows a person can fake when instructed, research has shifted to determining if respondents actually fake on his or her own.

Establishing whether people fake on personality assessments remains complicated because faking can depend on things such as the test, the testing situation, and the stakes of the test (Arthur, Glaze, Villado, & Taylor, 2010; Luther & Thornton, 1999). Morgeson et al. (2007) reported mixed results, with seven reviewed articles showing that applicants produced higher scores than non-applicants, indicating faking. Four other sources indicated that faking occurred (although less than anticipated) and three found that it did not, with applicants having similar scores to non-applicants (Morgeson et al., 2007). Additionally, Tett et al. (2006) pointed out numerous authors that claim faking is not pervasive while Griffith et al. (2006) suggest five more studies that show a significant portion of applicants fake. Although there is a lack of consensus on the extent that it occurs, there is a body of research that indicates at least some people do fake responses when taking pre-employment personality assessments. However, if there is "no right answer" to personality assessment items, how are people able to fake responses to these tests? Literature on the subject has indicated that people respond in a socially desirable manner.

Socially Desirable Responding

A fundamental issue with social desirability is a lack of clear conceptualization (Holtgraves, 2004). One of the most commonly utilized frameworks is Paulhus' (1986) two-factor interpretation that breaks socially desirable responding into self-deception and impression management. While a general consensus exists that the construct includes two aspects, these dimensions have been given dozens of different labels and researchers have applied subtle usage differences (Tett et al., 2006). The

lack of clear conceptualization is exemplified by the absence of convergent validity in three of the most common social desirability measures (Holden & Fekken, 1989). Despite conflicting viewpoints, agreement on a two-facet model stems from the idea that responding in a socially desirable manner is a product of both unconscious (unintentional) and conscious (deliberate) factors (Tett et al., 2006; Zerbe & Paulhus, 1987). The unconscious act of self-deception refers to self-presentation that is distorted in an overly positive manner, while the conscious use of impression management involves altering responses to form a socially positive image (Holtgraves, 2004). Whether conscious or not, socially desirable responding involves choosing responses based on what is perceived to be desirable by other people (Kuncel & Tellegen, 2009).

Crowne and Marlowe (1964) were the first to suggest that social desirability was not simply a response set, but rather a significant personality trait. They saw socially desirable responding as a representation of an individual's need for approval and then later interpreted it as avoidance of disapproval (Crowne, 1979; Crowne & Marlowe, 1964). With this change in perspective, Furnham (1986) took stock of the literature in the following two decades and found evidence that suggested social desirability is a stable bi-dimensional trait. Since then, researchers have found numerous significant correlations between social desirability measures and Big Five personality factors (Barrick & Mount, 1996; Rosse, Stecher, Miller, & Levin, 1998). Self-esteem was another personality construct related to social desirability (Winters & Neale, 1985). These and other examples led Tett et al. (2006) to suggest that social

desirability's link with several individual difference variables indicates that social desirability itself is an individual difference variable. Noting social desirability's relationship with Neuroticism and Conscientiousness, Smith and Ellingson (2002) also advance the idea that it represents a unique component of personality.

In addition to the propensity to respond in a socially desirable manner as an individual characteristic, the concept of social desirability also manifests itself in varying degrees throughout personality measures. Edwards (1957) was one of the first researchers to point out that personality items differed in social desirability content, with some items being more or less desirable than others. By asking raters to judge the desirability of items, Edwards (1957) found that ratings were consistent across wide varieties of people (e.g,. genders, nationalities). Dalen, Stanton, and Roberts (2001) explain that, "…there is a common notion of what is 'good' or 'bad' on personality traits." Morgeson et al. (2007) echoed the assumption that the "right" or most positive answer could be apparent to respondents. The idea that item desirability varies, is central to the development of some of the most common social desirability measures. Many social desirability assessments were developed by creating scales made up of items that were rated on their level of social desirability (e.g., Edwards Social Desirability Scale, Marlowe-Crowne Social desirability Scale, Responding Desirably on Attitudes and Opinions; Paulhus, 1991). This usage also highlights the relationship between faking and socially desirable items, because higher scores on these socially desirable items are a sign of faking (Edwards, 1957). Desirability can obscure the content of an item and provide applicants information that allows them to

distort their answers in a positive way (i.e., fake). Because of the constant threat of response distortion and the influence of social desirability, researchers have worked tirelessly to determine the cognitive processes involved in faking. The present paper aims to add to the literature on distortion within the broader response process.

Response Process

While there have been many faking response models proposed (e.g., McFarland & Ryan 2006; Mueller-Hanson, Heggestad, & Thornton, 2006; Tett & Simonet, 2011; Ziegler, 2011), researchers have not achieved a consensus on the most accurate. This lack of consensus has resulted in calls for a focus on theory and process in applicant faking (Griffith & Peterson, 2011; Kuncel, Goldberg, & Kiger, 2011). Two common methods used to assess the cognitive processes of respondents are "think aloud" protocols and measuring response latencies (Amelang, Eisenhut, & Rindermann, 1991). Verbal protocols have been used by Robie, Brown, and Beaty (2007) and Ziegler (2011) to gain insight into the thoughts subjects were having while completing a personality assessment. While providing a glimpse at the cognitive process, think aloud protocols are not without drawbacks. Nisbett and Wilson (1977) suggest that subjects are unable to lend insight to the higher order cognitions that they are experiencing. To address some of these shortcomings, researchers have turned to response latencies.

For decades, response latencies have been utilized outside of faking research to assess cognitive processes. At the most basic, response latencies are related to the length of the item stem, with longer items requiring more reading and comprehension

(Casey & Tryon, 2001; Dunn, Lushene, & O'Neil, 1972). There are individual differences in reading speed, item comprehension, and the physical response processes that also effect latencies (Tatsuoka & Tatsuoka, 1980). When adjusting for individual differences, response latencies become a measure of the subject's cognitive effort (Sudman, Bradburn, & Schwarz, 1996). Introspectively determining the frequency a participant engages in a particular behavior or the mental accessibility of the information required to respond, are examples of cognitive effort that can influence response times (Johnson, 2004; Mervielde, 1988). Finally, the distance between an item's position on a trait continuum and the respondent's trait level has been shown to affect response latencies (Ferrando & Lorenzo-Seva, 2007; Kuncel, 1973; Kuncel & Fiske, 1974). As an item approaches a respondent's response threshold for that trait, it takes longer for them to respond to that item.

Hsu, Santelli, and Hsu (1989) conducted one of the earliest studies on faking and response latencies. The authors found that response latencies have a faking detection ability because participants in faking conditions consistently respond faster, differentiating them from non-faking respondents. To explain this phenomenon, Hsu et al. (1989) attributed faster responses to the use of cognitive shortcuts that did not require accessing self-referenced thoughts. Holden, Fekken, and Cotton (1991) proposed that when a respondent answers a test item, they are comparing that item with a cognitive schema. Respondents have various schemas based on the traits that make up their personalities (Fiske & Taylor, 1984). Building on the concept of schemas as response mechanisms, Holden, Kroner, Fekken, and Popham (1992)

developed a model of faking response where responses are faster to items that fit the participant's self-schema and slower when those items conflict with their identity. Holden et al. (1992) also proposed that subjects could create faking schemas that allow them to respond quickly when the item is desirable. By adopting a fake good strategy, a participant could create a schema where favorable impression drives their responses. Holden and Kroner (1992) and Holden and Hibbs (1995) replicated these results, with subjects responding slower when items do not match their active schema. This theory was also supported by Brunetti, Schlottmann, Scott, and Hollrah (1998) in a study where subjects were instructed to fake psychopathology. While this body of research provides support for the idea that faking creates faster response times, most of the subjects in these studies were simply instructed to fake or respond in a certain manner. In a study where social desirability was induced without instruction, subjects under an increased social desirability condition had slower response times than participants who were allowed to respond anonymously, presumably lowering their concern for social desirability (Holtgraves, 2004). This suggests that if there is a realistic reason to manipulate scores, it may take longer to respond.

A recent study by van Hooft and Born (2012) tried to shed additional light on the faking response process by using a reaction time study that incorporated eye tracking as a way to increase the detection of fakers. Van Hooft and Born (2012) used a within-subjects design, having participants complete a personality measure under normal and "fake good" instructions. Reaction times were recorded as well as eye movement information in an attempt to explain more of the cognitive process active in

faking response (van Hooft & Born, 2012). The authors duplicated Hsu et al.'s (1989)

result of faster response latencies in their faking condition but also produced

additional faking detection with their eye tracking measures (van Hooft & Born,

2012). Eye tracking was also used as a parallel indicator of cognitive processing, with

longer fixations representing thought that is more complex. The length of fixations

improved faking classification beyond response latencies supporting the theoretical

position that faking is less cognitively complex and that it elicits less cognitive

demand (van Hooft & Born, 2012).

Because of the pervasive influence of social desirability on personality items

and its relationship with faking, an accurate measure of item social desirability could

provide additional information about the response process. Several studies have

shown that response times are faster to socially desirable items (e.g., Hanley, 1962;

Holden, Fekken, & Jackson, 1985; Kulas & Stachowski, 2012). Because of social

desirability's role in the response process and its effect on latencies, including this

metric may tease apart differences across items and provide additional information on

what factors are considered during the decision process. By following Holden et al.'s

(1992) protocol of using positively keyed vs. negatively keyed items as a proxy for

desirability, researchers are potentially missing valuable information about what

participants do when the desirability of an item is not clear. As Kulas and Stachowski

(2012) showed, the relationship between response time and desirability is not linear,

with responses that are faster for both desirable and undesirable items, but slower for

responses where the desirability is ambiguous or neutral. Limiting interpretation to a

two-facet view of desirability is unnecessarily restrictive and may result in

unexplained variance.

Another limitation of previous research on faking and social desirability is the

use of instructed faking. As pointed out by van Hooft and Born (2012), researchers

that experimentally induce faking could yield results quite different from those who

simply instruct participants to fake. Creating a situation where response enhancement

is beneficial but not instructed could activate complicated cognitive processes, thus

influencing response times. Research by Zickar et al. (2004) shows that faking in

actual applicant settings may have multiple influences and is not as simple as a binary

"fake" or "don't fake" condition.

Finally, researchers should attempt to resolve the incongruence between the

theoretical conclusion that faking is easier than honest responding and the results of

the "perceived task difficulty" items that indicate faking is harder. Literature on the

subject suggests that response latency can be used as a measure of the subject's

difficulty in responding (Hanley, 1962, 1965; Nowakowska, 1983). While subjects

rated faking as harder through a global measure, it is possible that individual items

caused a lingering sense of difficulty in van Hooft and Born's (2012) experiment. By

assessing response difficulty per item, Rogers (1973) found a strong correlation

between an item's difficulty rating and its "controversiality" (operationalized as

neutral desirability by Edwards, 1953). This may indicate that participants consider

the social desirability of items while they respond. Cone (1971) also found that

subjects report greater ease in answering socially desirable items. When items were

obviously desirable or undesirable, subjects reported little difficulty in choosing their answer. Directly measuring item difficulty as subjects respond along with response time information would allow for direct assessment of the meaning of response latencies as well as the possible meaning of their association with social desirability.

This literature review has shown that people are distorting their answers during personality assessments in varying contexts. This distortion arises as both a function of items and individual differences. While it is easy to identify, understanding the underlying processes of faking is of vital importance to the field. Investigators have identified response latency as one of the better tools available to aid in understanding faking. By using this method, it is possible to analyze differences in faking by items and situation. Understanding differences in both items and condition should help expand our understanding of faking and ultimately reduce it in the future. Before this can happen, several inconsistencies in our knowledge need to be resolved.

The Current Study

When the response process is studied, social desirability's contamination of personality items should be acknowledged and taken into account. This permeating force possibly affects the perceived difficulty of responding to personality assessment items (which in turn may be revealed via response latencies). By investigating how these three variables interact during personality measurement, the field can move one-step closer toward modeling the personality response process. The current study goes beyond the work of van Hooft and Born (2012) by incorporating these considerations into the experimental design. The current study examines how social desirability

relates to the response latency and reported response difficulty of personality items. Hypothesis 1 predicts that items with neutral social desirability have larger response latencies and will also be rated as more difficult to answer. Hypothesis 2 is non-directional and incorporates experimental conditions of faking, being honest, and a control condition: If Holden et al.'s (1992) theory that schemas are used by faking participants is accurate, honest condition participants will have greater response latencies and higher reported difficulty in responding than participants in the other conditions. Alternatively, if Holtgrave's (2004) theory that social desirability engages an editing process is correct, response latencies and difficulty will be less influenced by social desirability for participants in the honest condition.

METHOD

Participants

Undergraduate psychology students were recruited using Sona-System at Saint Cloud State University. Seventy-five students completed the experimental study, with 50 identifying as female and 25 as male. Participants' age ranged from 17 to 56 with a mean age of approximately 22 ($M = 22.24$, $SD = 5.14$).

Materials

Personality items were selected from the International Personality Item Pool (IPIP) within scales that followed the big five-factor structure (Goldberg, 1992). I selected items because of the accessibility of social desirability ratings for these items from a previous study. In the previous study, undergraduate students rated 300 IPIP items on social desirability. I selected 75 of these 300 items for inclusion in the current study. In order to identify the 75, I identified each of the 300 items into their respective Big Five facets and sorted them by social desirability. Three items with the highest desirability ratings and three items with the lowest ratings as well as items with the most neutral ratings were selected first (this process was repeated within each of the Big Five dimensions). Next, I included three items that fell half way between

the lowest ratings and the most neutral as well as three more items that were half way

between the most neutral and highest ratings. This resulted in 15 items for each of the

Big Five constructs with a stratified distribution of social desirability ratings. A

computer randomly presented these items to each participant. The 75 item stems and

their average social desirability rating can be found in the Appendix.

Procedure

I randomly assigned participants to one of three experimental groups. All

participants completed the questionnaires in an on-campus psychology lab. The

surveys were administered in a counterbalanced manner to reduce order effects. When

given the computer-administered assessment, participants sat in a private lab room

with a computer running E-Prime. After reading simple instructions explaining to

respond "as quickly but accurately as possible to the following personality items",

participants were given five practice items to become familiar with the layout of the

response options and the format of the questionnaire. After each response, the item

would be removed from the screen and the program asked subjects to rate "how

difficult was it to make your selection" on that item. Upon completion of the practice

items, participants were read one of the following instruction sets depending on what

condition they were randomly assigned.

I provided the following instructions to subjects in the experimental "faking"

group: "In this experiment, you will be responding to a series of personality

questions. Please answer as quickly but accurately as possible but be aware that

participants with the best personality score will receive this 50 dollars. Are there any

questions?" After this was read, a crisp new 50-dollar bill was placed on the computer desk in front of the participant.

I instructed subjects in the "honest" group as follows: "In this experiment, you will be responding to a series of personality questions. Please answer as quickly but accurately as possible regarding whether the statement IS or IS NOT descriptive of you. It is typical with these assessments for people to present themselves in a socially desirable manner. We want you to try very hard to not do this and instead answer completely honestly. Are there any questions?" A large mirror was present in the room positioned in such a way that participants could see themselves while completing the assessment during this condition.

Finally, subjects in the control group were only provided with the instructions needed to complete the questionnaire: "In this experiment, you will be responding to a series of personality questions. Please answer as quickly but accurately as possible regarding whether the statement IS or IS NOT descriptive of you. Are there any questions?"

If participants did not have any questions, they would begin the 75 randomly presented items. The computer recorded response latencies on the personality items from the time the item stem was presented until a response option was clicked.

Participants provided direct ratings along a five-point graphic rating scale assessing level of agreement: 1-Strongly Disagree, 2-Disagree, 3-Niether Agree nor Disagree, 4-Agree, and 5-Strongly Agree. Response options were equally spaced around the cursor start position and the cursor returned to the center of the screen after

the participants gave each difficulty rating. An example of the screen participants saw

when responding with this scale is presented in Figure 1.

Figure 1

Personality Response Screen

To determine how difficult the response choice was, each item was also rated

on its difficulty to make the decision with a four-point scale: 1-Very Easy, 2- Slightly

Easy, 3-Slightly Difficult, and 4-Very Difficult. The difficulty response screen is

demonstrated in Figure 2. The items were administered by computer through E-Prime

software version 2.0 to record response choices and so that response latencies could be

recorded in milliseconds.

How difficult was it to make your decision on the previous item?

| Very Easy | Slightly Easy | Slightly Difficult | Very Difficult |

Figure 2

Difficulty Response Screen

In addition to completing the computer based personality measure, subjects

also completed a paper copy of the BIDR and a paper based NEO-FFI to assess

personality on the Big Five facets. Once all three had been administered, subjects

completed a manipulation check to assess their conscious level of honesty.

Participants were asked to rate how true the following statements were: "I responded

completely honestly on the **computer** based assessment," and "I tried to make myself appear better than I truly am on the **computer** based assessment." The same questions were also asked with "paper" in place of the bolded "computer." Finally, participants were debriefed and given credit that is used in undergraduate psychology classes. The following debriefing was read to each participant: "The study that you participated in was interested in how you respond to personality items using computer based administration. Because we are interested in why people sometimes do not give completely honest self-evaluations, we told some participants that the best personality score would receive 50 dollars. To be fair, all participants, including you, will have equal opportunity for the $50. Please do not tell other students or participants about the $50 because it will likely influence the results of the study. Do you have any questions? Thank you." Upon termination of the experiment, one participant was randomly chosen and given the $50.

RESULTS

External Scales

To assess the convergent validity of the IPIP items with the Big Five

personality factors, scale scores for IPIP and NEO were calculated and then correlated.

Each NEO FFM dimension was found to correlate at least $r_{xy} = .68$ with its respective

IPIP scale (See Table 1). Cronbach's alpha was used to assess the internal consistency

of each scale, and these estimates are located in the diagonal of Table 1. Table 2

includes each scale's mean and standard deviation across conditions.

Table 1

Personality Scale Correlations

Scale	Neuro NEO	Extra NEO	Open NEO	Agree NEO	Consc NEO	Neuro IPIP	Extra IPIP	Open IPIP	Agree IPIP	Consc IPIP
NeuroNEO	.84									
ExtraNEO	-.32*	.77								
OpenNEO	.18	-.11	.78							
AgreeNEO	-.18	.28*	-.01	.76						
ConscNEO	-.28*	.24	-.27*	.28*	.86					
NeuroIPIP	.72*	-.33*	.05	-.36*	-.29*	.85				
ExtraIPIP	-.39*	.74*	.10	.23	.17	-.49*	.68			
OpenIPIP	.23	.13	.68*	.22	-.18	.03	.33*	.55		
AgreeIPIP	.04	.18	-.11	.69*	.40*	-.16	.14	.15	.68	
ConscIPIP	-.14	.25*	-.12	.41*	.76*	-.30*	.25*	-.03	.60*	.74

Note: Convergent validity coefficients are bolded. *p < 0.05. Values on the diagonal are within-scale Chronbach's Alphas.

Table 2

Mean and Standard Deviation

Scale Means					NEO				
Condition	Neuro		Exra		Open		Agree		Consc
Overall	2.70 (0.67)		3.67 (0.49)		3.47 (0.57)		3.79 (0.49)		3.79 (0.62)
Faking	2.67 (0.61)		3.63 (0.53)		3.57 (0.38)		3.77 (0.55)		3.74 (0.59)
Honest	2.79 (0.69)		3.65 (0.45)		3.37 (0.73)		3.70 (0.53)		3.74 (0.69)
Control	2.77 (0.74)		3.64 (0.49)		3.51 (0.51)		3.81 (0.46)		3.81 (0.54)

					IPIP				
Condition	Neuro		Exra		Open		Agree		Consc
Overall	2.26 (0.55)		3.57 (0.44)		3.71 (0.41)		3.83 (0.41)		3.83 (0.46)
Faking	2.17 (0.61)		3.63 (0.44)		3.74 (0.39)		3.81 (0.41)		3.78 (0.46)
Honest	2.32 (0.58)		3.50 (0.42)		3.62 (0.45)		3.72 (0.38)		3.71 (0.44)
Control	2.33 (0.52)		3.57 (0.47)		3.76 (0.38)		3.90 (0.46)		3.87 (0.52)

Note: Standard deviations in parenthesis.

Response Latencies

Response latencies were examined and outliers were removed. First, any

response latency that was three times the mean latency for that item was removed to

eliminate responses where the participant did not engage in the question. Then, any

response that was faster than 200ms or slower than 3 standard deviations from the

mean of that item was removed to further clean the responses where a participant had

accidentally clicked without reading or became distracted and lingered on a question.

Between-subjects t-tests indicate that there were significant differences in reaction

time between English and the 8 non-English speakers (see the Appendix). Non-

English speakers (n = 8) had significantly slower reaction times on over half of the

items (k = 42). Although these participants had slower reaction times, none of the

22

analyses of interest changed significantly when they were included in the sample. For this reason, I retained ESL participants' responses.

Social Desirability on Response Latencies

Because item length is known to influence latency, I calculated character count for each item. Item length was entered in the first step of a hierarchical moderated regression as a control and was found to be a significant predictor of response latency ($R^2 = .54$, $F(1, 223) = 259.75$, $p < .05$). The main effects of social desirability and condition were entered in step 2. These predictors accounted for significant variance beyond item length (R^2 change $= .02$, $F(3, 220) = 70.60$, $p < .05$). I entered the linear interactions between social desirability and condition into step 3. These interaction terms did not add significant variance (R^2 change $= .00$, $F(2, 218) = .73$, ns). Because social desirability's effect on latency was expected to be curvilinear, a squared social desirability term was entered into step 4 of the regression. This squared term added significantly to the model (R^2 change $= .06$, $F(1, 217) = .34.64$, $p < .05$). This significance of this polynomial demonstrates the curvilinear relationship between item social desirability and response latency supporting hypothesis 1. The standardized coefficients for predictors in this model are shown in Table 3. It takes longer for people to answer in the honest condition, confirming the first part of hypothesis 2. Figure 3 demonstrates the difference in regression coefficients by condition.

Table 3

Hierarchical Moderated Regression Analysis Predicting Job Satisfaction

Step	Predictor variable	β	ΔR^2	R^2_{total}
1	Character Count	.64*	.54*	.54
2	Honest Condition	.10*	.02*	.56
	Faking Condition	.03		
	Social Desirability	-.13*		
3	SD * Condition 1		.00	.56
	SD * Condition 2			
4	SD^2		.06*	.63

Note: SD = Social Desirability. *p < .05. Conditions were dummy coded.
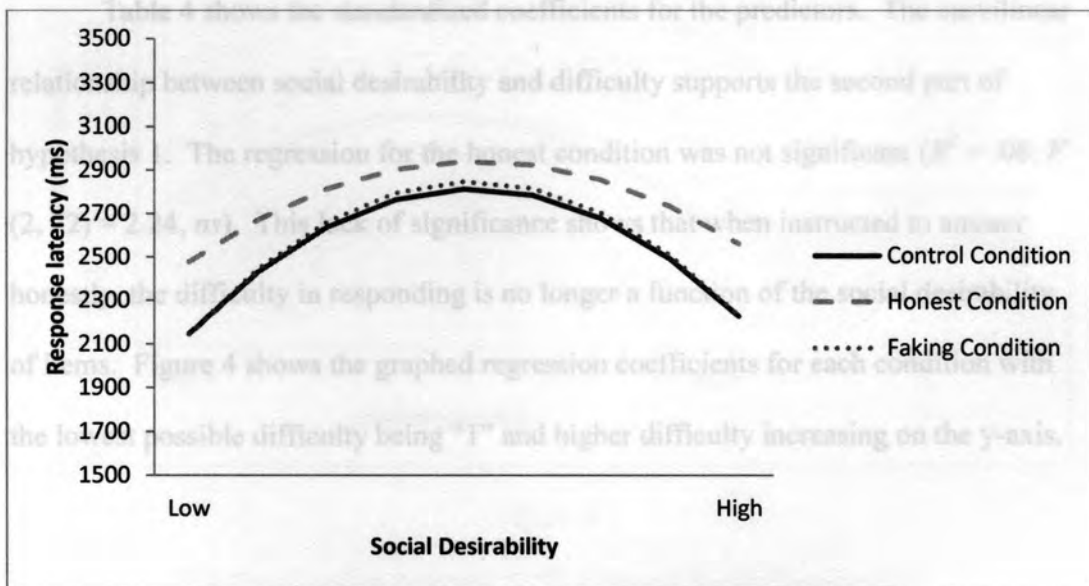
Figure 3

The Regression Relationship between Social Desirability and Response Latency
Moderated by Experimental Condition

Social Desirability on Difficulty Ratings

No significant differences accrued in average difficulty ratings across all items by condition ($F (2, 64) = .776$, *ns*). Because a curvilinear relationship was expected between social desirability and difficulty ratings, I calculated a hierarchical moderated regression model. The main effects of social desirability and condition were entered in step 1 of the regression and found to be significant ($R^2 = .11$, $F (3, 221) = 8.619$, $p <$ .05). The interaction between social desirability and condition was entered in step 2 but did not add significantly to the variance (R^2 *change* $= .01$, $F (2, 219) = .97$, *ns*). To test the curvilinear effect of social desirability, a squared social desirability term was entered in step 3 and found to add significant variance to the model (R^2 *change* $=$ .10, $F (1, 218) = 28.51$, $p < .05$).

Table 4 shows the standardized coefficients for the predictors. The curvilinear relationship between social desirability and difficulty supports the second part of hypothesis 1. The regression for the honest condition was not significant ($R^2 = .06$. F $(2, 72) = 2.24$, *ns*). This lack of significance shows that when instructed to answer honestly, the difficulty in responding is no longer a function of the social desirability of items. Figure 4 shows the graphed regression coefficients for each condition with the lowest possible difficulty being "1" and higher difficulty increasing on the y-axis.

Table 4

Hierarchical Moderated Regression Analysis Predicting Difficulty

Step	Predictor variable	β	ΔR^2	R$^2_{total}$
1	Honest Condition	.33*	.11*	.11
	Faking Condition	.01		
	Social Desirability	1.50*		
2	SD * Condition 1		.01	.12
	SD * Condition 2			
3	SD^2		.10*	.22

Note: SD = Social Desirability. *p < .05. Conditions dummy coded.
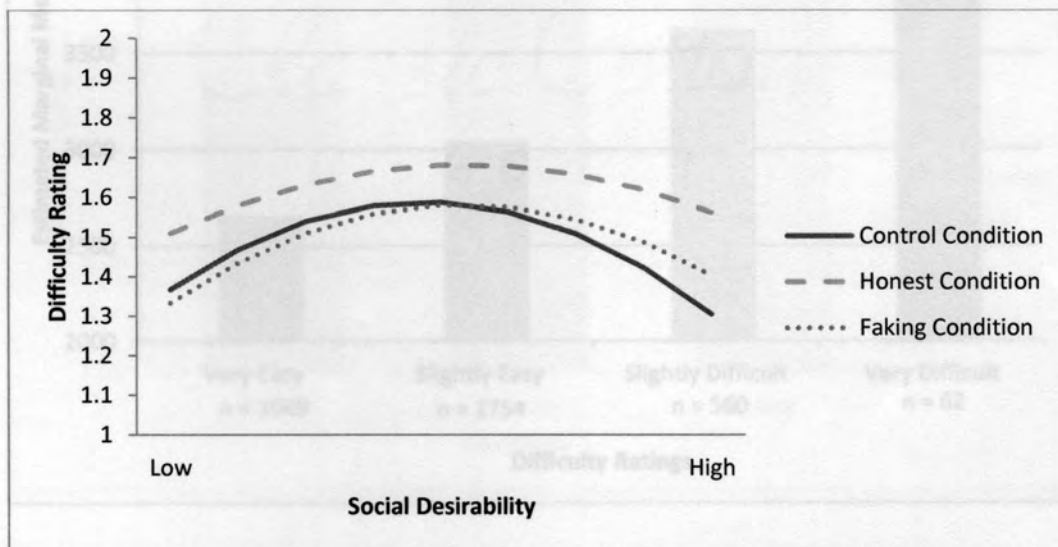
Figure 4

The Relationship between Difficulty and Social Desirability
Moderated by Experimental Condition

Difficulty and Response Latency

To assess the relationship between response latency, I conducted an ANCOVA to control for the effect of item length on response latency. A significant effect of difficulty accrued on response latency after controlling for the effect of character count, $F(1, 3) = 131.79$, $p < .05$. The covariate of character count was significantly related to response latency, $F(1, 3) = 684.311$, $p < .05$. The estimated marginal means of response latency are reported in Figure 5.
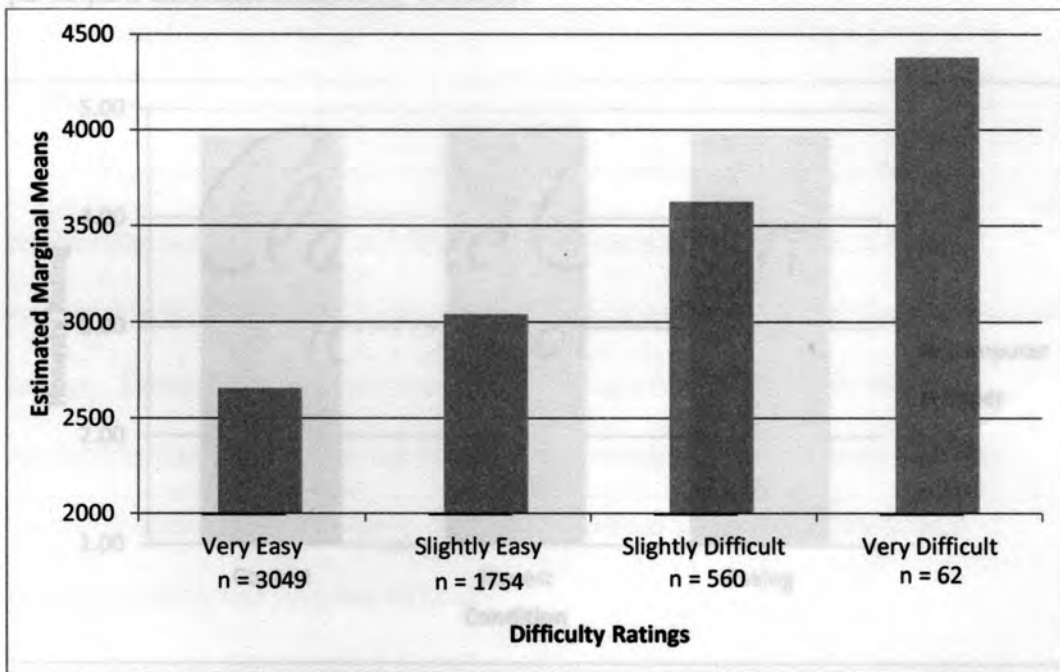
Figure 5

Estimated Marginal Means of Response Latency. (The number of endorsements for each difficulty option is listed under each bar.)

Manipulation Checks

Manipulation checks were used to assess the conscious level of honesty participants had while participating in the study. I detected no significant differences in reported honesty between the computer administered IPIP questions and the paper based NEO (t (128) = .739, *ns*). No significant differences occurred in reported difficulty between conditions (F (2, 62) = .327, *ns*). Figure 6 shows the lack of reported honesty difference between conditions with 5 on the y-axis representing the participant answered completely honestly.
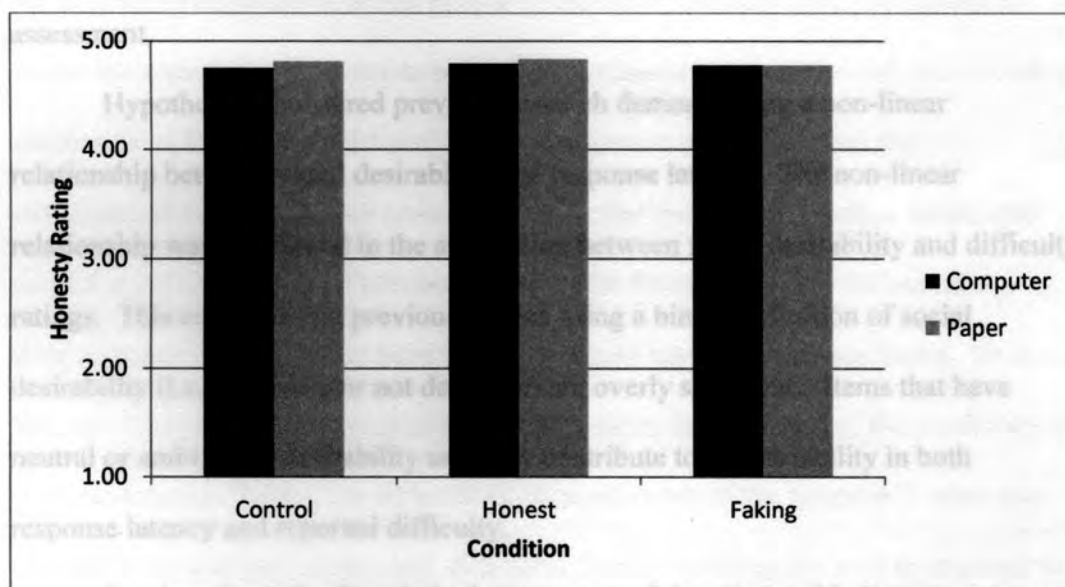
Figure 6

The Reported Honesty for the Computer and Paper Based Personality Assessments

Chapter IV

DISCUSSION

The present study sought to examine the role of item social desirability in the assessment of personality. This study also attempted to expand the understanding of cognitive processes active during personality assessment. This was done by assessing item response latency and self-reported difficulty while completing a personality assessment.

Hypothesis 1 bolstered previous research demonstrating a non-linear relationship between social desirability and response latency. The non-linear relationship was also found in the association between social desirability and difficulty ratings. This confirms that previous studies using a binary definition of social desirability (i.e., desirable or not desirable) are overly simplistic. Items that have neutral or ambiguous desirability uniquely contribute to the variability in both response latency and reported difficulty.

Previous investigations lacked assessment of the relationship between response latency and difficulty ratings. An analysis of covariance revealed the strong relationship between difficulty and response latency when controlling for character count. This shows response latency has a direct relationship with response difficulty,

28

something previously only assumed. This finding reinforces the use of response

latency as an indicator of cognitive processing.

The interpretation of hypothesis 2 is very difficult to make without certain

assumptions. If participants in the control condition are using a schema of themselves

to answer the survey, the matching response latency pattern of the faking condition

could be interpreted as the use of a schema. This would indicate that the increase in

response latency in the honest condition is due to actual introspection requiring

increased cognitive processing, confirming Holden et al.'s (1992) theory on schemas.

An alternative explanation is that people in the faking condition did not attempt to

distort their answers either due to lack of motivation or the fear of some sort of trick or

retribution. I found no significant differences in scale scores between the

experimental conditions, so it remains possible that similarities between faking and

control conditions occurred because people in the faking condition did not change

their responses or cognitions from how they would have normally answered. Despite

this, no differences occurred as a function of honesty between any of the conditions in

the manipulation checks, yet differences occurred between the honest and other two

conditions in response latency and difficulty. Future investigation will be required to

explain this effect.

Analysis of the difficulty ratings showed that when I told subjects to answer

honestly, reported difficulty no longer covaried with social desirability. While both

the faking and control condition had significant curvilinear relationships between

difficulty and social desirability, I observed no linear or non-linear relationship

between difficulty and social desirability. This finding suggests that the social desirability of an item plays an important role a respondent's response process. In both the control and faking conditions, participants use the desirability of items to help determine their response. A very desirable or very undesirable item produces quicker responding, probably because people generally like to present themselves positively. When the desirability of an item is ambiguous, participants cannot rely on that part of the item to generate a response leading to higher difficulty ratings. When told to respond honestly, the desirability of an item no longer predicts how difficult it is to answer the item. By telling them to answer honestly, participants become introspective and individual differences in personality become the source of difficulty ratings. This conclusion uniquely demonstrates social desirability's role in the personality response process.

Limitations

There were some shortcomings of this study. First, the sample size was smaller than ideal. With only 75 participants, each condition had only 25 subjects. While this still allows for statistical analysis, it lowers the possible effect sizes and increases the likelihood of an error. Follow-up investigations should have a greater sample size to increase the validity of findings.

Another limitation was the weakness of the manipulations. There were several informative differences, but they were somewhat weak. There were no significant differences in reported honesty for the participants despite that being the variable the study attempted to manipulate. The experimental condition was a novel attempt to

induce actual faking. Similar to a real situation, it could have been viewed as a trick or trap to catch fakers. It is possible that chance for 50 dollars was not that motivating. It also could have been unclear how to fake, (e.g., what is best) as opposed to in a real life hiring situation where the desirable responses may be more obvious.

Further Directions

Future studies would benefit from investigating and developing other manipulations that induce faking without simply telling participants to fake. The existing research is scarce on alternative manipulations that can induce faking without outright asking for it. Applying this study's methodology in a situation where actual hiring will take place may expose additional effects not seen in the lab.

The larger study intent would have included individual differences in self-deception and impression management as well as personality (as measured via the NEO). Although these individual difference variables were not incorporated into the current study, they may play an important role in understanding personality differences that relate to the response process. For example, participants with higher scores on the BIDR may utilize item social desirability more rigorously to determine item response than people with lower impression management and self-deception levels.

Another direction for future research would be to determine what is difficult about responding to personality items. This study documents the relationship between social desirability and difficulty, but was not designed to induce difficulty with neutral

social desirability. Experimental studies that are created to directly assess this relationship would be more informative and expansive.

While items themselves have varying levels of desirability, Kuncel and Tellegen (2009) have shown that the responses to items have varying levels of desirability as well. Analyzing response latency and difficulty in regards to response desirability instead of item desirability may provide additional insight into the personality response process.

REFERENCES

Amelang, M., Eisenhut, K., & Kindermann, H. (1991). Responding to adjective check list items: A reaction time analysis. *Personality and Individual Differences*, *12*(6), 523-533.

Arthur, W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2010). The magnitude and extent of cheating and response distortion effects on unproctored internet-based tests of cognitive ability and personality. *International Journal of Selection and Assessment*, *18*(1), 1-16.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, *81*, 261-272.

REFERENCES

Brunetti, D. G., Schlottmann, R. S., Scott, A. B., & Hollrah, J. L. (1998). Instructed faking and MMPI-2 response latencies: The potential for assessing response validity. *Journal of Clinical Psychology*, *54*(2), 143-153.

Casey, M. M., & Tryon, W. W. (2001). Validating a double-press method for computer administration of personality inventory items. *Psychological Assessment*, *13*(4), 521-530.

Cone, J. D. (1971). Social desirability scale values and ease of responding to personality statements. *Proceedings of the Annual Convention of the American Psychological Association*, *6*, 119-120.

Crowne, D. P. (1979). *The experimental study of personality*. Hillsdale, NJ: Erlbaum.

Crowne, D. P., & Marlowe, D. (1964). *The approval motive*. New York: Wiley.

Dalen, L. H., Stanton, N. A., & Roberts, A. D. (2001). Faking personality questionnaires in personnel selection. *Journal of Management Development*, *20*(8), 729-742.

Dunn, T. G., Lushene, R. E., & O'Neil, H. F. (1972). Complete automation of the MMPI and a study of its response latencies. *Journal of Consulting and Clinical Psychology*, *39*, 381-387.

REFERENCES

Amelang, M., Eisenhut, K., & Rindermann, H. (1991). Responding to adjective check list items: A reaction time analysis. *Personality and Individual Differences, 12*(6), 523-533.

Arthur, W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2010). The magnitude and extent of cheating and response distortion effects on unproctored internet-based tests of cognitive ability and personality. *International Journal of Selection and Assessment, 18*(1), 1-16.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81,* 261- 272.

Brunetti, D. G., Schlottmann, R. S., Scott, A. B., & Hollrah, J. L. (1998). Instructed faking and MMPI-2 response latencies: The potential for assessing response validity. *Journal of Clinical Psychology, 54*(2), 143-153.

Casey, M. M., & Tryon, W. W. (2001). Validating a double-press method for computer administration of personality inventory items. *Psychological Assessment, 13*(4), 521-530.

Cone, J. D. (1971). Social desirability scale values and ease of responding to personality statements. *Proceedings of the Annual Convention of the American Psychological Association, 6,* 119-120.

Crowne, D. P. (1979). *The experimental study of personality*. Hillsdale, NJ: Erlbaum.

Crowne, D. P., & Marlowe, D. (1964). *The approval motive*. New York: Wiley.

Dalen, L. H., Stanton, N. A., & Roberts, A. D. (2001). Faking personality questionnaires in personnel selection. *Journal of Management Development, 20*(8), 729-742.

Dunn, T. G., Lushene, R. E., & O'Neil, H. F. (1972). Complete automation of the MMPI and a study of its response latencies. *Journal of Consulting and Clinical Psychology, 39,* 381- 387.

34

Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of applied Psychology, 37,* 90-93.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* New York: Dryden.

Ellingson, J. E., Sackett, P. R., & Connelly, B. S. (2007). Personality assessment across selection and development contexts: Insights into response distortion. *Journal of Applied Psychology, 92*(2), 386-395.

Ferrando, P. J., & Lorenzo-Seva, U. (2007). A measurement model for Likert responses that incorporates response time. *Multivariate Behavioral Research, 42*(4), 675-706.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition.* Reading, MA. Addison-Wesley.

Furnham, A. (1986). Response bias, social desirability, and dissimulation. *Personality and Individual Differences, 7,* 385-400.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26-42.

Griffith, R. L., Malm, T., English, A., Yoshita, Y., & Gujar, A. (2006). Applicant faking behavior: Teasing apart the influence of situational variance, cognitive biases, and individual differences. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 1-19). Greenwich, CT: Information Age Publishing.

Griffith, R. L., & McDaniel, M. (2006). The nature of deception and applicant faking behavior. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 1-19). Greenwich, CT: Information Age Publishing.

Griffith, R. L., & Peterson, M. H. (2011). One piece at a time: The puzzle of applicant faking and a call for theory. *Human Performance, 24*(4), 291-301.

Hanley, C. (1962). The "difficulty" of a personality inventory item. *Educational and Psychological Measurement, 22*(3), 577-584.

Hanley, C. (1965). Personality item difficulty and acquiescence. *Journal of Applied Psychology, 49*(3), 205.

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92*(5), 1270-1285.

Holden, R. R., & Fekken, G. C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers? *Journal of Research in Personality, 23*(2), 180-191.

Holden, R. R., Fekken, G. C., & Cotton, D. H. (1991). Assessing psychopathology using structured test-item response latencies. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(1), 111-118.

Holden, R. R., Fekken, G. C., & Jackson, D. N. (1985). Structured personality test item characteristics and validity. *Journal of Research in Personality, 19*(4), 386-394.

Holden, R. R., & Hibbs, N. (1995). Incremental validity of response latencies for detecting fakers on a personality test. *Journal of Research in Personality, 29*(3), 362-372.

Holden, R. R., & Kroner, D. G. (1992). Relative efficacy of differential response latencies for detecting faking on a self-report measure of psychopathology. *Psychological Assessment, 4*(2), 170-173.

Holden, R. R., Kroner, D. G., Fekken, G. C., & Popham, S. M. (1992). A model of personality test item response dissimulation. *Journal of Personality and Social Psychology, 63*(2), 272-279.

Holden, R. R., & Passey, J. (2010). Socially desirable responding in personality assessment: Not necessarily faking and not necessarily substance. *Personality and Individual Differences, 49*(5), 446-450.

Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin, 30*(2), 161-172.

Hsu, L. M., Santelli, J., & Hsu, J. (1989). Faking detection validity and incremental validity of response latencies to MMPI subtle and obvious items. *Journal of Personality Assessment, 53*(2), 278-295.

Johnson, M. (2004). Timepieces: Components of survey question response latencies. *Political Psychology, 25*, 679-702.

Kulas, J. T., & Stachowski, A. A. (2012). Social desirability in personality assessment: A variable item contamination perspective. *The International Journal of Educational and Psychological Assessment, 11*(1), 23-42.

Kuncel, N. R., Goldberg, L. R., & Kiger, T. (2011). A plea for process in personality prevarication. *Human Performance, 24*(4), 373-378.

Kuncel, N. R., & Tellegen, A. (2009). A conceptual and empirical reexamination of the measurement of the social desirability of items: Implications for detecting desirable response style and scale development. *Personnel Psychology, 62*(2), 201-228.

Kuncel, R. B. (1973). Response processes and relative location of subject and item. *Educational and Psychological Measurement, 33*(3), 545-563.

Kuncel, R. B., & Fiske, D. W. (1974). Stability of response process and response. *Educational and Psychological Measurement, 34*(4), 743-755.

Luther, N. J., & Thornton III, G. C. (1999). Does faking on employment tests matter. *Employment Testing Law and Policy Reporter, 8*, 129-136.

Mahar, D., Cologon, J., & Duck, J. (1995). Response strategies when faking personality questionnaires in a vocational selection setting. *Personality and Individual Differences, 18*(5), 605-609.

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*(4), 979-1016.

Mervielde, I. (1988). Cognitive processes and computerized personality assessment. *European Journal of Personality, 2*(2), 97-111.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*(3), 683-729.

Mueller-Hanson, R., Heggestad, E. D., & Thornton, G. C. (2006). Individual differences in impression management: An exploration of the psychological processes underlying faking. *Psychology Science, 48*(3), 288-312.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review; Psychological Review, 84*(3), 231.

Nowakowska, M. (1983). *Quantitative psychology: Some chosen problems and new ideas.* Amsterdam: North-Holland.

Ones, D. S., Reiss, A. D., & Viswesvaran, C. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*(6), 660-679.

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner, & J. S. Wiggins (Eds.), *Personality assessment via questionnaires: Current issues in theory and measurement* (pp. 143-165). New York: Springer-Verlag.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 17-59). San Diego, CA: Academic Press.

Robie, C., Brown, D. J., & Beaty, J. C. (2007). Do people fake on personality inventories? A verbal protocol analysis. *Journal of Business and Psychology, 21*(4), 489-509.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied* Psychology, 83, 634-644.

Rogers, T. B. (1973). Toward a definition of the difficulty of a personality item. *Psychological Reports, 33*(1), 159-166.

Smith, D. B., & Ellingson, J. E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology, 87, 211-219.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology.* San Francisco: Jossey-Bass Publishers.

Tatsuoka, K. K., & Tatsuoka, M. M. (1980). A model for incorporating response-time data in scoring achievement tests. In D. J. Weiss (Ed.), *Proceedings of the 1979 Computerized Adaptive Testing Conference* (pp. 236-256). Minneapolis: University of Minnesota, Department of Psychology.

Tett, R. P., Anderson, M. G., Ho, C. Yang, T. S., Huang, L., & Hanvongse, A. (2006). Seven nested questions about faking on personality tests: An overview and interactionist model of item-level response distortion. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 1-19). Greenwich, CT: Information Age Publishing.

Tett, R. P., & Simonet, D. V. (2011). Faking in personality assessment: A "multisaturation" perspective on faking as performance. *Human Performance, 24*(4), 302-321.

van Hooft, E. A., & Born, M. P. (2012). Intentional response distortion on personality tests: Using eye-tracking to understand response processes when faking. *Journal of Applied Psychology, 97*(2), 301-316.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*(2), 197-210.

Wagner-Menghin, M. (2006). The mixed-rasch model: An example for analyzing the meaning of response latencies in a personality questionnaire. *Journal of Applied Measurement, 7*(2), 225-237.

Winters, K. C., & Neale, J. M. (1985). Mania and low self-esteem. *Journal of Abnormal Psychology, 94*(3), 282.

Zerbe W. J., & Paulhus D. L. (1987). Socially desirable responding in organizational behavior: A reconceptualization. *Academy of Management Review*, 12, 250-264.

Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods, 7(2),* 168-190.

Ziegler, M. (2011). Applicant faking: A look into the black box. *The Industrial-Organizational Psychologist, 49*(1), 29-36.

Item Characteristics

APPENDIX

Item Characteristics

Item stem	SD	t	% Very Easy	% Slightly Easy	% Slightly Difficult	% Very Difficult
Take advantage of others.	1.21	1.24	68.9	23	8.1	0
Insult people.	1.36	1.78	72.2	18.1	9.7	0
Turn my back on others.	1.43	2.43	73.3	20	4	2.7
Break my promises.	1.50	3.63	60.8	29.7	8.1	1.4
Get angry easily.	1.50	3	63.4	22.5	11.3	2.8
Do the opposite of what is asked.	1.86	4.52	56.8	28.4	14.9	0
Put little time and effort into my work.	1.86	2.91	66.2	31.1	2.7	0
Lose my temper.	2.00	1.13	59.5	28.4	12.2	0
Dislike myself.	2.07	4.29	63.8	20.3	13	2.9
Am not really interested in others.	2.43	2.25	69.9	27.4	2.7	0
Often feel uncomfortable around others.	2.71	2.9	46.6	41.1	11	1.4
Avoid contacts with others.	2.71	2.31	60.3	34.2	5.5	0
Feel desperate.	2.86	2.02	63	23.3	12.3	1.4
Often feel blue.	3.00	1.25	50.7	31.9	14.5	2.9
Am often down in the dumps.	3.07	2.07	57.5	31.5	11	0
Tend to dislike soft-hearted people.	3.07	1.96	52.7	29.7	17.6	0
Am not interested in other people's problems.	3.21	1.03	58.9	26	15.1	0
Have difficulty starting tasks.	3.50	2.82	44.4	40.3	12.5	2.8
Break rules.	3.50	2.35	61.6	27.4	11	0
Am wary of others.	3.50	2.13	45.1	45.1	9.9	0
Do crazy things.	3.57	0.88	44.6	37.8	17.6	0
Don't like the idea of change.	3.86	2.17	52.1	37	8.2	2.7
Don't understand people who get emotional.	3.86	1.89	54.8	32.9	11	1.4
Dislike changes.	3.86	1.77	59.5	31.1	8.1	1.4
React slowly.	4.14	0.73	44.4	47.2	5.6	2.8
Would never go hang gliding or bungee jumping.	4.21	2.42	45.1	39.4	9.9	5.6

Keep in the background.	4.21	1.79	63.9	23.6	11.1	1.4
Am afraid that I will do the wrong thing.	4.43	1.72	52.7	31.1	16.2	0
Have difficulty imagining things.	4.50	2.33	54.2	36.1	8.3	1.4
Do not enjoy watching dance performances.	4.50	1.72	51.4	40.5	8.1	0
Believe in one true religion.	4.57	1.95	66.2	18.9	12.2	2.7
Am afraid to draw attention to myself.	4.64	0.53	52.1	34.2	12.3	1.4
Seek quiet.	4.71	1.83	53.5	42.3	4.2	0
Know how to get around the rules.	4.79	1.6	51.4	34.7	11.1	2.8
Hold back my opinions.	4.86	1.59	44.4	37.5	18.1	0
Am not bothered by disorder.	4.86	1.18	45.1	42.3	8.5	4.2
Don't like to draw attention to myself.	4.93	4.87	47.9	35.6	15.1	1.4
Stick to my chosen path.	4.93	2.19	45.8	45.8	6.9	1.4
Am very pleased with myself.	4.93	1.52	59.7	31.9	8.3	0
Seldom daydream.	4.93	0.77	52.8	34.7	12.5	0
Enjoy wild flights of fantasy.	5.00	3.19	50	34.7	12.5	2.8
Tend to vote for conservative political candidates.	5.00	2.76	56.8	20.3	13.5	9.5
Am not highly motivated to succeed.	5.00	2.67	78.9	18.3	2.8	0
Consider myself an average person.	5.07	2.59	50	37.8	10.8	1.4
Dislike talking about myself.	5.21	4.61	49.3	38.7	10.7	1.3
Like music.	5.75	0.02	84.3	12.9	2.9	0
Am not easily affected by my emotions.	5.86	1.7	46.6	37	15.1	1.4
Interested in many things.	6.00	2.59	73.2	18.3	4.2	4.2
Laugh my way through life.	6.36	3.78	46.7	38.7	14.7	0
Can talk others into doing things.	6.43	3.41	45.9	41.9	10.8	1.4
Involve others in what I am doing.	6.64	4.16	47.8	39.1	11.6	1.4
Remain calm under pressure.	7.07	0.87	46.4	44.9	8.7	0
Handle tasks smoothly.	7.14	2.77	48.6	31.9	19.4	0

Believe that people are basically moral.	7.14	2.47	39.7	43.8	16.4	0
Can handle a lot of information.	7.14	1.31	54.1	23	21.6	1.4
Am not easily bothered by things.	7.14	0.33	45.1	39.4	15.5	0
Can handle complex problems.	7.21	1.66	49.3	34.2	16.4	0
Set high standards for myself and others.	7.29	4.09	56.8	36.5	6.8	0
Stick to the rules.	7.29	3.01	54.1	37.8	8.1	0
Choose my words with care.	7.29	2.8	54.3	35.7	10	0
Anticipate the needs of others.	7.29	2.61	64.7	30.9	4.4	0
Am passionate about causes.	7.50	2.62	64.4	27.4	6.8	1.4
Feel others' emotions.	7.64	2.29	68.5	24.7	5.5	1.4
Have a good word for everyone.	8.00	3.12	59.7	31.9	6.9	1.4
Keep my cool.	8.00	2.43	52.7	31.1	16.2	0
Cheer people up.	8.14	1.8	61.4	30	7.1	1.4
Readily overcome setbacks.	8.21	2.28	38.4	42.5	16.4	2.7
Radiate joy.	8.21	0.56	52.1	35.6	12.3	0
Complete tasks successfully.	8.36	4.95	53.6	31.9	13	1.4
Make friends easily.	8.36	3.15	56.9	38.9	4.2	0
Feel comfortable with myself.	8.36	2.68	52.1	36.6	11.3	0
Tell the truth.	8.50	3.08	67.1	27.1	5.7	0
Make people feel welcome.	8.50	1.56	65.3	31.9	2.8	0
Keep my promises.	8.86	1.03	79.5	19.2	1.4	0
Love to help others.	8.86	0.98	65.2	26.1	5.8	2.9

Note: SD = Average social desirability rating. Test of English Vs. Non-English differences in response latencies with t > 2.02 as significant p < .05. Percentage of difficulty ratings for each item are listed.