

3-23-2017

# Can Selection Tests Administered via Video Games Reduce Faking?

Philip Scott Ramsay

*University of South Florida*, [psramsay@mail.usf.edu](mailto:psramsay@mail.usf.edu)

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [Psychology Commons](#)

---

## Scholar Commons Citation

Ramsay, Philip Scott, "Can Selection Tests Administered via Video Games Reduce Faking?" (2017). *Graduate Theses and Dissertations*. <http://scholarcommons.usf.edu/etd/6750>

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).

Can Selection Tests Administered via Video Games Reduce Faking?

by

Philip S. Ramsay

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Arts  
with a concentration in Industrial-Organizational Psychology  
Department of Psychology  
College of Arts and Sciences  
University of South Florida

Major Professor: Wendy L. Bedwell, Ph.D.  
Walter Borman, Ph.D.  
Doug Rohrer, Ph.D.  
James K. Summers, Ph.D.

Date of Approval:  
September 6, 2016

Keywords: personality, measurement, response distortion, validity

Copyright © 2016, Philip S. Ramsay

## TABLE OF CONTENTS

List of Tables .....	iii
List of Figures .....	iv
Abstract .....	v
Chapter One: Introduction .....	1
Personality as a Predictor of Job Performance .....	3
Can it be measured? .....	4
Does it matter? .....	4
Will applicants respond accurately? .....	5
Faking in Selection Contexts .....	6
Video Games in Selection Contexts .....	13
Video Games and Surveys .....	15
Enjoyment .....	16
Cognitive Load Theory .....	17
Chapter Two: Methods .....	19
Participants .....	19
Materials .....	19
Game platform .....	19
Measures .....	21
Demographics .....	21
Personality .....	21
Procedure .....	21
Time 1 .....	22
Time 2 .....	22
Chapter Three: Results .....	23

Hypothesis 1a: Intentions to Fake Influence Faking Conscientiousness Scores .....	25
Hypothesis 1b: Intentions to Fake Influence Faking Extraversion Scores .....	25
Hypothesis 2a: Conscientiousness Construct Validation.....	26
Hypothesis 2b: Extraversion Construct Validation.....	26
Hypothesis 3a: Scale Medium Interact With Intention-Behavior Relationship for Conscientiousness .....	27
Hypothesis 3b: Scale Medium Interact With Intention-Behavior Relationship for Extraversion .....	27
Post-hoc Analyses .....	27
 Chapter Four: Discussion.....	 30
Theoretical Implications .....	32
Practical Implications.....	33
Limitations .....	34
Conclusion .....	35
 References .....	 37
 Appendices.....	 45
Appendix A: Demographic Information.....	45
Appendix B: Video Game Self-Efficacy .....	48
Appendix C: Personality .....	49
Appendix D: Prompts .....	51
Appendix E: IRB Approval .....	52

## LIST OF TABLES

<b>Table 1.</b> Influence of Faking on Selection Decisions (adapted from Holden, 2006).....	6
<b>Table 2.</b> Primary Meta-Analysis Results for Big-Five Personality Dimensions (adapted from Birkleand et al., 2006).....	10
<b>Table 3.</b> Experimental Conditions.....	22
<b>Table 4.</b> Correlation Matrix.....	24
<b>Table 5.</b> <i>t</i> -tests Justifying the Combination of C1 & C2, and C3 & C4.....	25
<b>Table 6.</b> Summary of Hypothesis Tests .....	29

## LIST OF FIGURES

<b>Figure 1.</b> Model of faking (adapted from McFarland & Ryan, 2000). .....	7
<b>Figure 2.</b> Modified model of faking (adapted from McFarland & Ryan, 2000).....	13
<b>Figure 3.</b> Screenshots of the game platform <i>The Road Ahead</i> .....	20
<b>Figure 4.</b> Visualization of hypotheses.....	23
<b>Figure 5.</b> Hypothesis 3a & 3b Interaction Plots .....	28

## **ABSTRACT**

One of the fundamental underlying assumptions of selection procedures is that the information gathered from applicants is accurate, and thus, will predict performance on the job (Donovan, Dwight, & Schneider, 2014; Schmitt & Sinha, 2011). As self-report instruments such as paper-and-pencil tests and unsupervised online surveys become more prevalent in organizational selection contexts (Truxillo & Bauer, 2011) due to ease of use and cost efficiency, the concern of applicants faking responses to inaccurately portray themselves as more highly desirable is increasingly critical (Hough, Oswald, & Ployhart, 2001). Depending on the exact magnitude of the particular selection event, this compromise of validity may cost an organization just as much as they stand to gain from an accurate self-report selection tool. The aim of this study is to test the viability of a video game platform designed to aid personnel selection by reducing faking. This thesis first outlines the most widely assessed predictors of job performance and briefly review the state-of-the-science of personality research in the context of employee selection. Then, a review of faking, drawing upon a model of faking proposed by McFarland & Ryan (2000), describes the impact it has on employee selection based on personality tests. Drawing upon Malone's (1981) theory of intrinsic motivation as well as Sweller's (1994) theory of cognitive load, I proposed the use of a video game platform as a counter-measure to faking selection assessments. Results suggest that participants are less able to fake personality assessments when assessed via video games as compared to online surveys.

## CHAPTER ONE: INTRODUCTION

One of the fundamental underlying assumptions of selection procedures is that the information gathered from applicants is accurate and thus, will predict performance on the job (Donovan et al., 2014; Schmitt & Sinha, 2011). As self-report instruments, such as paper-and-pencil tests and unsupervised online surveys, become more prevalent in organizational selection contexts (Truxillo & Bauer, 2011) due to ease of use and cost efficiency, the potential of applicants faking responses to inaccurately portray themselves as more highly desirable is increasingly problematic (Hough et al., 2001). When the assumption of honest responding is violated by faking, so too is the validity of the selection process (e.g., Christiansen, Goffin, Johnston, & Rothstein, 1994; Holden, 2006; Holden & Jackson, 1981; Mueller-Hanson, Heggestad, & Thornton III, 2003; Topping & O'Gorman, 1997; Worthington & Schlottmann, 1986). Depending on the exact magnitude of the particular selection event, this compromise of validity may cost an organization just as much as they stand to gain from an accurate self-report selection tool.

As defined by Paulhus (2002), faking is simply “the tendency to give overly positive self-descriptions” (p. 50). This definition—commonly referred to as “faking good”—highlights an issue with the implicit idea that when candidates respond to surveys designed to elicit self-descriptions, their responses are expected to accurately reflect their true score (Ziegler, MacCann, & Roberts, 2011a). The possibility of faking puts doubt in this assumption. This means that even the most internally-consistent, comprehensive, easy-to-use, and predictive



assessments are only as effective as the applicant allows them to be. From an organizational perspective, it is not only a waste of resources but an enormous misrepresentation of their work force if a selection instrument based on individual differences (that are known to predict job performance; e.g., Barrick & Mount, 1991, 1996; Hertz & Donovan, 2000) hinges almost entirely on the verisimilitude of applicants' accurate responses. Depending on the value given to these assessments and their susceptibility to faking attempts, the implication of this measurement shortcoming is that an applicant does not necessarily need to possess the desired characteristics required for success in a given field; but rather needs only to be proficient at guessing what is desirable in a given job and describing themselves in such a manner. From a practical perspective, it is therefore imperative that researchers strive to develop valid, fake-resistant assessment techniques to ensure the integrity of the selection process.

One promising area of research that is beginning to address this issue is the use of alternative selection instruments, such as video games. The underlying premise is that if applicants are cognitively engaged in another activity that can also assess the desired characteristics known to predict job performance in a specific field, they will be unable to fake. The aim of this study is to test the viability of such a platform designed to aid personnel selection by reducing faking. First, the amount that directed faking can inflate scores based on a prominent model of faking (i.e., McFarland & Ryan, 2000) will be assessed. Then, the reliability of the video game assessment tool will be determined and compared to the reliability of an equivalent online survey. Finally, the relative susceptibility to faking inherent in surveys presented via video games versus traditional online mediums will be measured and contrasted. Specifically, this study will investigate faking using the measurement of personality as it has been found to be one of the most powerful predictors of job performance (Barrick & Mount,

1991) and is, thus, a frequently used selection tool (McCarthy et al., 2013). The scope will be limited to the investigation of conscientiousness and extraversion because these two personality dimensions have been found to be the most influential across work domains (Barrick & Mount, 1991) and the most and least susceptible to faking attempts, respectively (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006).

### **Personality as a Predictor of Job Performance**

Job performance is perhaps the most widely researched employee outcome assessed by Industrial-Organizational (I-O) psychologists. Some of the most powerful predictors of job performance involve characteristics of the employee such as ability (e.g., cognitive ability, cognitive aptitudes, job knowledge; Hunter, 1986; Schmidt, 2002), personality (e.g., Barrick & Mount, 1991), and vocational interests (e.g., Van Iddekinge, Putka, & Campbell, 2011). Due to the recent resurgence in personality research (John, Naumann, & Soto, 2008) and its increased use in selection contexts, this paper will focus on measuring personality for selection purposes.

Personality is broadly understood to be the combination of characteristics or qualities that form an individual's distinctive character (John et al., 2008). Currently, the most widely used conceptualization of personality is the Five Factor Model or the Big Five factors of personality (Goldberg, 1990; McCrae & Costa, 1987). The Big Five posits that a person's personality can be conceptualized as having a certain proclivity towards being more or less open, conscientious, extraverted, agreeable, and neurotic which relate to that person's behavioral tendencies. Together, one's standing on these five latent traits or factors, often measured by the International Personality Item Pool (IPIP; Goldberg, 1999), reflects the range of behaviors that the individual will typically engage in.

To be useful for selection purposes, trait-level dispositional personality must (1) be able to be measured in a consistent and interpretable manner, and (2) provide meaningful results as a valid predictor performance. Thus, it is critical to understand how accurate individual response patterns to self-report personality measures generally are, particularly in selection contexts. Below, the current state of the science will be summarized with regards to these concerns.

**Can it be measured?** At the onset of personality research, it was unclear as to whether the thousands of descriptors that humans have come up with to communicate the idea of personality could be compiled into meaningful, interpretable chunks. Then, after the initial lexical approach to conceptualizing personality (Galton, 1884) spurred the field in the right direction, an explosion of personality scales were developed, the issue became one of scientific pluralism (John et al., 2008). The Big Five conceptualization of dispositional personality has aided personality researchers tremendously by enabling generalization of their findings through the use of a robust, reliable, and simple structure of personality (Goldberg, 1999; Tett & Christiansen, 2007). Since the widespread acceptance of the Big Five, personality research has seen an even greater surge and appears to only be escalating in popularity.

**Does it matter?** Personality as a predictor of job performance ratings has long been overshadowed by other traits such as general mental ability, physical ability, and impression management. This view shifted rapidly after two landmark events: the Equal Employment Opportunity Commission's (EEOC) mandates for the reduction of adverse impact, and Barrick and Mount's (1991) meta-analysis of job performance predictors. The EEOC's acts and guidelines in the 1960s and 1970s prompted researchers to find alternative predictors of job performance that were not differentially predictive for protected groups of people. This generated renewed interest in personality assessment as it is much more robust to racial and

gender group differences (Hough et al., 2001). Then, meta-analyses of the relationship between Big Five personality dimensions and job performance by Barrick and Mount (1991) and many others (e.g., Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Judge, Heller, & Mount, 2002; Salgado, 1997) suggested that personality was indeed a powerful predictor. Thanks largely to these influential events, the field of personality research has now regained much of the interest that it had in the early- and mid-90's because of how useful and generally applicable it is as a predictor of human behavior.

**Will applicants respond accurately?** Since personality assessments are inherently transparent in an effort to allow the assessee to accurately recall their own behavioral tendencies, they are also susceptible to inaccurate responses or faking (also referred to as aberrant responding, socially desirable responding, response bias, malingering, et cetera). This susceptibility is problematic in job selection contexts since scores on personality assessments may influence hiring decisions, and high-stakes situations, such as the job application process, may encourage applicants to misrepresent themselves on these tests (Smith & Ellingson, 2002). Although it is evident that some applicants engage in faking (estimates of the prevalence of faking range from 15% to 39%; Griffith & Converse, 2011) to various degrees, it appears that any amount of faking can drastically impact hiring decisions as a result of its potential influence on scale validities, scale means, and the interpretation of the scores of those who faked (see Table 1; Holden, 2006). In sum, the answer to the question of whether applicants will respond accurately is unclear. One thing that is certain is that test takers are capable of faking (Birkeland et al., 2006; Viswesvaran & Ones, 1999), so the degree of influence that faking has on hiring decisions must be minimized.

**Table 1.** Influence of Faking on Selection Decisions (adapted from Holden, 2006)

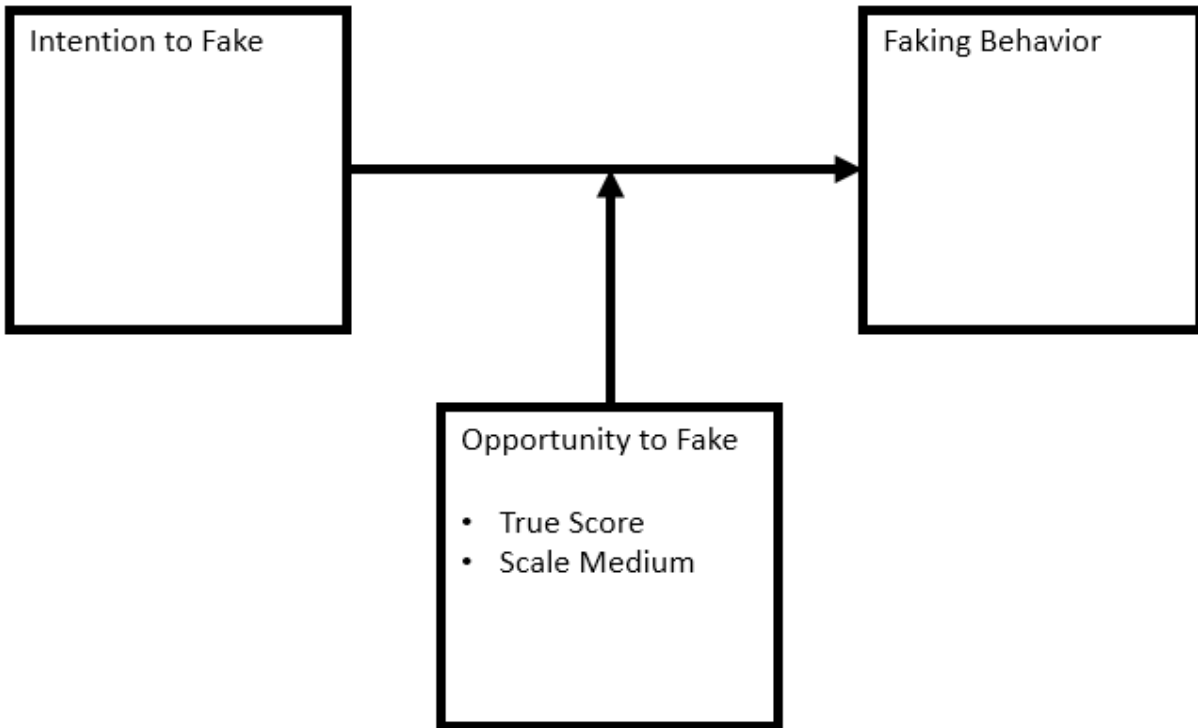
---

If 90% of respondents fake, will this influence:	
scale validities?	<i>Probably</i>
scale mean scores?	<i>Probably</i>
the valid interpretation of an individual protocol for someone who faked?	<i>Definitely</i>
If 20% of respondents fake, will this influence:	
scale validities?	<i>Unclear</i>
scale mean scores?	<i>Probably</i>
the valid interpretation of an individual protocol for someone who faked?	<i>Definitely</i>
If 2% of respondents fake, will this influence:	
scale validities?	<i>Probably not</i>
scale mean scores?	<i>Probably not</i>
the valid interpretation of an individual protocol for someone who faked?	<i>Definitely</i>

---

### **Faking in Selection Contexts**

One popular depiction of the factors influencing an assessee's test score is McFarland and Ryan's (2000) model of faking. This model attests that a person has implicit beliefs about faking that influences his or her intentions to fake which then in turn influence his or her faking behavior, ultimately producing a test score (see Figure 1). The factors that influence this path are the individual characteristics of the person such as values or morals, ability to fake, and situational components including the motivation and opportunity to fake. Although all of these factors are worthy of study and impact faking behaviors to varying degrees, this investigation will focus on the influence of intentions to fake on faking behavior (responding).



**Figure 1.** Model of faking (adapted from McFarland & Ryan, 2000).

The three most common approaches to studying the effects of faking intentions on faking behavior involve (1) directed faking studies, (2) applicant-incumbent comparisons, and (3) measuring faking with impression-management scales (Ziegler, MacCann, & Roberts, 2011b). Directed faking studies have been the most common approach and are typically conducted by directing participants to respond honestly or motivating them to fake (such as by incentivizing particular responses). This design is perhaps the most powerful tool for determining the extent to which people are able to distort responses, but it is lacking in its ability to observe the effects of assessee characteristics on faking intentions.

Applicant-incumbent comparisons are a more direct way of assessing the ways in which people choose to represent themselves in different contexts. The biggest strength of this

particular design is that it can better emulate the contexts in which faking may impact decisions; however, the fact that applicant groups are typically very different than non-applicant groups due to range-restriction, motivation, and several other factors suggests that this design's strength is also an inherent weakness, depending on the purpose of the study. Even with these drawbacks, studies utilizing an applicant-incumbent approach have gleaned many widely-applicable findings, such as the average amount by which incumbents' and applicants' responses to personality inventories differ (see Table 2; Birkeland et al., 2006).

**Table 2.** Primary Meta-Analysis Results for Big-Five Personality Dimensions (adapted from Birkleand et al., 2006)

Construct	<i>K</i>	Applicant ( <i>N</i> )	Non-applicant ( <i>N</i> )	Mean <i>d</i>	95% Confidence Interval	$\delta$ -hat	REVC	<i>Q</i>
Extraversion	29	53,745	18,096	.11*	.006-.217	.13	.0709	667.7**
Agreeableness	20	27,842	16,126	.16	-.010 to .324	.19	.1356	1057.5**
Conscientiousness	27	69,325	18,941	.45**	.303-.591	.52	.1332	1337.1**
Openness	20	46,037	14,224	.13*	.014-.243	.15	.0585	409.41**
Emotional Stability	25	21,219	13,991	.44*	.278-.593	.50	.1464	935.0**

\* $p < .05$ , \*\* $p < .01$ . REVC, random effects variance component; *Q*,  $\chi^2$  test of effect size heterogeneity,  $\delta$ -hat, mean *d* adjusted for reliability



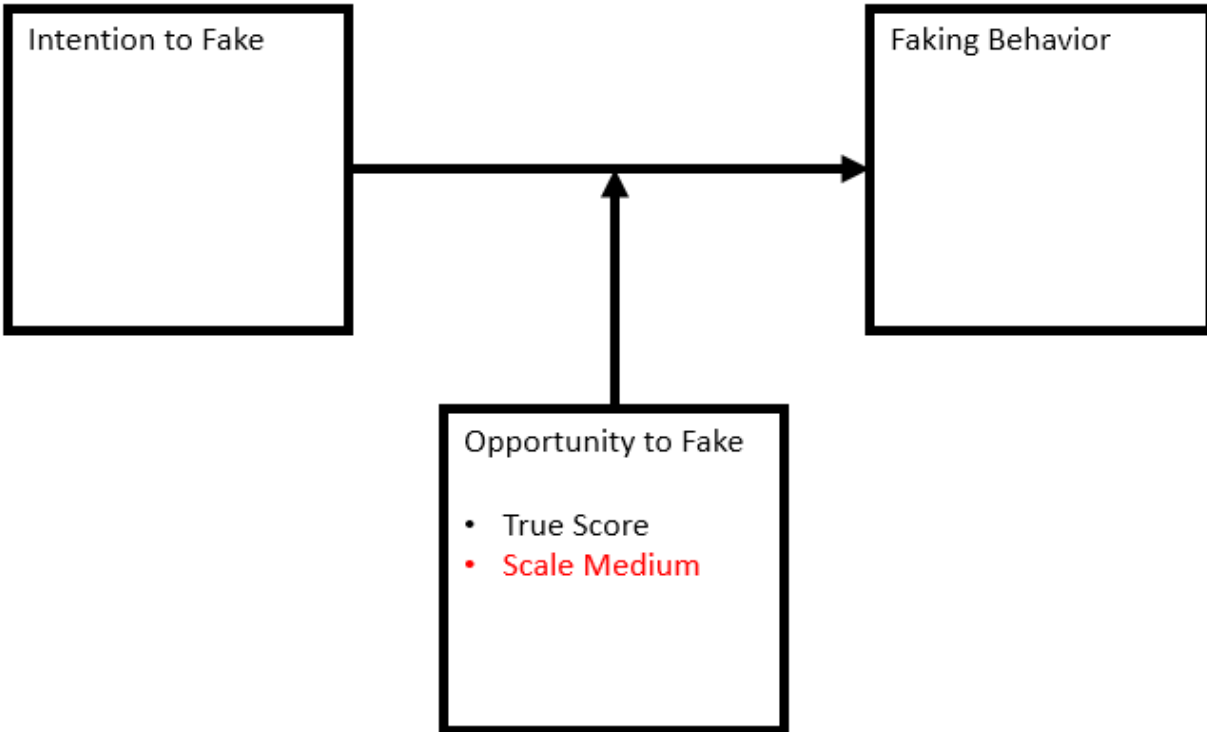
Impression-management scales are used to directly assess the degree to which a person is representing his or herself as better than expected. Ideally, this approach would reveal a person's inner faking tendencies and could then be used to better understand their responses; however, these scales have been shown to overlap significantly with Big Five personality characteristics (Smith & Ellingson, 2002). Therefore, they may be better suited as a measure of a person's perception of themselves rather than an indicator of intention to fake or faking behavior.

This study will utilize a directed faking approach in order to assess the amount that people are able to fake personality assessments. These approaches have typically been recognized as being the most powerful tool for detecting differences between motivated and unmotivated persons, and since it has been shown that Big Five personality assessments can be faked (Birkeland et al., 2006; Griffith, Chmielowski, & Yoshita, 2007; Smith & Ellingson, 2002; D. B. Smith & Robie, 2004; Van Iddekinge, Raymark, & Roth, 2005), McFarland and Ryan's (2000) model of faking proposes that a person's intention to fake will significantly predict their faking behavior. Given that conscientiousness and extraversion have been identified as the two most powerful predictors of job performance (Barrick & Mount, 1991) and the most and least susceptible to faking attempts respectively (meta-analytic Cohen's  $d$  for conscientiousness = 0.45 and extraversion = 0.11; Birkeland et al., 2006), this study will be restricted to looking at conscientiousness and extraversion. Although these personality constructs have been shown to be susceptible to faking attempts to varying degrees, both conscientiousness and extraversion display significantly inflated response patterns in faking versus non-faking contexts (Birkeland et al., 2006). Thus, I suggest the following:

*Hypothesis 1a: Intentions to fake will significantly influence faking behavior such that people given faking prompts will report higher conscientiousness scores.*

*Hypothesis 1b: Intentions to fake will significantly influence faking behavior such that people given faking prompts will report higher extraversion scores.*

Faking has the potential to dramatically influence selection decisions (e.g., Table 1; Holden, 2006) so it behooves us to mitigate the degree to which applicants are able to fake as much as possible. Some work has been done to adjust scores based on assessee responses to social desirability scales (e.g., Goffin & Christiansen, 2003), change the way that items are presented and scored such as through the use of forced choice assessments (e.g., Chernyshenko et al., 2009), or even by assessing constructs through entirely different formats such as structured interviews (e.g., Van Iddekinge et al., 2005) just to name a few (for a more comprehensive review see Dilchert & Ones, 2011). Van Iddekinge and colleagues' work (2005) using a different platform to assess the same constructs suggests an extension to McFarland and Ryan's faking theory (2000). Specifically, the scale medium has the potential to influence the opportunity to fake (see Figure 2). Thus, I argue that evaluating personality through the use of video games may reduce the degree to which assessees are able to distort their responses. Below, I outline how games, in particular, may increase the validity of personality assessment and ultimately improve selection procedures.



**Figure 2.** Modified model of faking (adapted from McFarland & Ryan, 2000).

*Note: Modification indicated by red text.*

### **Video Games in Selection Contexts**

Games such as chess have been used as training tools for centuries (Shenk, 2007; Smith, 2009). Beginning with the Roman Empire’s use of sand tables to represent soldiers and units in battle with abstract icons, games have been used to simulate high-risk scenarios such as battles or hiring decisions for as long as they’ve existed. These games have evolved as technologies allowed, producing such simulations as miniatures (e.g., sand tables), strategy board games (e.g., *Wei Hai*, *Chaturanga*, *War Chess*, *Kriegsspiel*, and *Tactics*), and computer games (e.g., *Camonette*, the *McClintic Theater Model*, *SimNet*, *Spearhead*, *Full Spectrum Warrior*, and *America’s Army*) which have been used for a variety of purposes ranging from battle preparation

(e.g., *War Chess*) to training (e.g., *Full Spectrum Warrior*), recruitment (e.g., *America's Army*), and many other applications (for a more exhaustive list see Smith, 2009). The continuing development and growing popularity of games is largely due to their adaptability as platforms for learning (Bedwell, Pavlas, Heyne, Lazzara, & Salas, 2012; Wenzler & Chartier, 1999), as well as the inherent intrinsic motivation that naturally occurs when having fun (Malone, 1981).

Modern mass-produced serious games—games that are developed for professional purposes—are working their way into industrial sectors at increasing rates due to reduction in development costs and increased popularity in the general public (Salas & Cannon-Bowers, 2001). Intuitively, the appeal of a flexible organizational training tool (Wenzler & Chartier, 1999) and the increasing interest in video games should also hold true for assessing the characteristics of job applicants. Likely due to how cost-effective mass surveys have become, the utility of games as self-report assessment tools is a relatively nascent area of research. However, it follows that utilizing video games as an assessment tool should provide investigators the ability to measure the same construct previously measured by any other means by manipulating relevant attributes of the game to fit specific needs (Bedwell et al., 2012). One method by which video games can be compared to other assessment mediums (e.g., online surveys) is by comparing the reliabilities of the assessments (e.g., Viswesvaran, Ones, & Schmidt, 1996). If one assessment is administered via multiple mediums, then the two measurement approaches should correlate with one another to the same degree as the assessment's internal consistency (e.g., its split-half reliability or correlation with itself). Therefore, the second set of hypotheses of this proposed study are that a measure of conscientiousness and extroversion will yield similar results as indicated by internal consistency metrics if measured by either video game or surveys.

*Hypothesis 2a: The split-half reliability of conscientiousness scales will meet or exceed the reliability of the original scales.*

*Hypothesis 2b: The split-half reliability of extraversion scales will meet or exceed the reliability of the original scales.*

## **Video Games and Surveys**

McFarland and Ryan's (2000) model encompasses a wide range of factors that influence an assessee's test score but appears to be missing the crucial contextual element. In today's modern age of assessing individual characteristics through the use of self-report surveys, interviews, situational judgment tests, online activity, or even video games, the medium deserves acknowledgement (Van Iddekinge et al., 2005). Applicants are frequently expected to present resumes, fill out assessments, and participate in interviews when applying for jobs or promotions. These different approaches to measuring the potential value of the candidate vary with respect to their inherent likeability, reliability, and potentially the degree to which they are susceptible to faking (e.g., McCarthy et al., 2013; Van Iddekinge et al., 2005).

As indicated previously, McFarland and Ryan's (2000) theory of faking fails to consider the effect of the delivery medium. Given the potential for various mediums to differ in the degree to which they are susceptible to faking, I argue that this is a necessary and important extension of McFarland and Ryan's (2000) work. The remainder of this paper will discuss the ways in which two scale mediums in particular (viz., online surveys and video games) may differ with respect to their susceptibility to faking responses to self-report assessments.

It would be ideal if assessees always responded honestly and accurately to assessments. However, situational demands, imperfect memory, and many other unavoidable factors often influence responses (Ziegler et al., 2011a). Moreover, if the assessee believes that s/he should

respond in a situationally desirable manner or fake (Ziegler et al., 2011a), intends to fake, has the ability and opportunity to fake (Ones & Viswesvaran, 1998), and/or the testing situation (e.g., job selection or clinical diagnosis) demands a particular response (McFarland & Ryan, 2000), then the assessee may choose to fake. In previous studies that have encouraged applicants to respond honestly versus directing the applicant to fake in a certain manner (thereby satisfying the previously outlined necessary conditions for faking to occur), applicants have repeatedly demonstrated that they are capable of misrepresenting themselves on evaluations (e.g., Griffith et al., 2007; Smith & Ellingson, 2002; D. B. Smith & Robie, 2004).

Video-game based assessments and survey based assessments may measure the same constructs in similar ways and be susceptible to faking attempts, but they differ along two important dimensions that should influence the degree to which an assessee is able to fake the assessment: (1) the amount of enjoyment inherent in the medium and (2) cognitive load. These two factors are critical drivers of assessee response patterns. Specifically, video games and surveys differ with respect to the amount of inherent fun as well as the cognitive load demands. These differences should influence applicant responses.

**Enjoyment.** Games are fun (Malone, 1980). The simple concept of inducing a state of enjoyment in the assessee may encourage them to become more invested in the assessment and limit their motivation to respond dishonestly. Malone's (1981) theory of intrinsically motivating instruction posits that the elements of challenge, fantasy, and curiosity draw the participant into the tool and encourage them to participate effectively. These elements encourage the development of a player's telepresence or psychological linkage (Ekman et al., 2012), which builds a player's investment in the outcome of the game and their avatar (Mallon & Lynch, 2014). This is the result of feeling empathetic towards the visualization of their commands via

the in-game avatar and the feedback that corresponds with the avatar's actions (Mallon & Lynch, 2014). Building off of this, an assessee's telepresence, investment, and/or empathetic connection with avatars or scenarios in a game should encourage the assessee to behave in ways that strengthen this connection through accurate and honest responses (McCarthy et al., 2013).

**Cognitive Load Theory.** Cognitive Load Theory (CLT) posits that the amount of simultaneous activity that a person is capable of performing is limited by the total cognitive load of the combined tasks or the sum of intrinsic, extraneous, and germane factors (Sweller, 1994). The intrinsic cognitive load (i.e., the amount of mental resources required to process the task; Paas, Renkl, & Sweller, 2003), and extraneous cognitive load (i.e., the method of communicating information; Chandler & Sweller, 1991) placed upon the assessee are a product of the task or task environment and dictate the amount of strain placed upon the assessee. If the task imposes a high amount of cognitive load, then the person will be less able to simultaneously handle other cognitive tasks.

Because participants will be having fun while playing a game, it follows that they will be less incentivized to fake (McCarthy et al., 2013). Additionally, since faking is hypothesized to be an active process that requires cognitive resources (McFarland & Ryan, 2000), and the video game medium demands more cognitive resources than the survey medium as a result of the game's demand for one's attention and a shift of presence (Takatalo, Häkkinen, Kaistinen, & Nyman, 2010), participants should be less able to fake assessments that are delivered via video games. I, therefore, propose that assessees will be less able to fake on assessments presented in a video game medium than a survey medium.

*Hypothesis 3a: Scale mediums will interact with the relationship between intentions to fake and faking behavior such that people will be less able to fake when conscientiousness is assessed via a video game than via a survey.*

*Hypothesis 3b: Scale mediums will interact with the relationship between intentions to fake and faking behavior such that people will be less able to fake when extraversion is assessed via a video game than via a survey.*



## CHAPTER TWO: METHODS

### Participants

A total of 304 students were sampled in this study. 77% of the participants were female, and the average age of participants was 20 years old. 52% of participants were white, 11% black, 18% Hispanic, 13% Asian, and 7% were other ethnicities. Participants over the age of 18 were recruited via e-mail at two large U.S. universities. Participants were compensated with course credit.

### Materials

The study was conducted online. The participants were required to use computers with internet connections. Surveys were delivered online via Qualtrics.

**Game platform.** The game platform being utilized is The Road Ahead, a proprietary program developed by Persona Labs for selection purposes. The game is designed as 2D side-scroller game similar to Mario (see Figure 3 for screenshots of The Road Ahead gameplay). The player controls an avatar which runs from left to right through a city street dodging various falling hazards. At occasional points, the game halts and the player is presented with a survey item. After responding to the item, the screen unfreezes and the player proceeds. The game repeats until the item pool is exhausted.



Figure 3. Screenshots of the game platform *The Road Ahead*.

## Measures

**Demographics.** In order to ensure that both samples are equivalent, demographic information was collected. Participants were asked to report their demographic information (e.g., gender, age, ethnicity, etc.) as well as video-game self-efficacy (Pavlas, Heyne, Bedwell, Lazzara, & Salas, 2010).

**Personality.** The International Personality Item Pool's Big Five 20-item Factors I and III scales were utilized. Response options for each item were represented in a Likert-type fashion ranging from 1 to 5 (*1 = very inaccurate; 5 = very accurate*). Conscientiousness: The 20-item scale has a reliability coefficient of  $\alpha = .88$ . This scale was split into two components, with the first half (the 10-item scale;  $\alpha = .79$ ) being presented in the game medium, and the second half (the remaining 10-items;  $\alpha \approx .79$ ) presented in an online-survey. An example item is "Please use the rating scale below to describe how accurately each statement describes you: [I] am always prepared." Extraversion: The 20-item scale has a reliability coefficient of  $\alpha = .91$ . This scale was split into two components, with the first half (the 10-item scale;  $\alpha = .87$ ) being presented in the game medium, and the second half (the remaining 10-items;  $\alpha \geq .87$ ) presented in an online-survey. An example item is "Please use the rating scale below to describe how accurately each statement describes you: [I] am the life of the party." For both personality scales, several items were reverse-coded (see Appendix C for the exact items).

## Procedure

Undergraduate students at a large eastern U.S. university and MBA students from a large western U.S. university who met the eligibility requirements were recruited. Participants were provided with links to participate in the study at two time points separated by one to three weeks in order to avoid any practice effects (Ployhart & Vandenberg, 2010).

**Time 1.** Participants were directed to one of four Qualtrics surveys that correspond to each of the four study conditions and serve to guide the participants through the experiment. Participants in the honest conditions (conditions 1 & 2; see Table 3) were given the prompt “The following surveys are for research purpose only, and as such, it is important that your responses be as honest as possible. Remember that your responses are confidential, and it is vital that you be as accurate and honest as possible.” Participants in the faking conditions (conditions 3 & 4; see Table 3) were prompted “The following surveys are typically used by organizations to hire new employees. Your task today is to respond to the following surveys as if you wanted to get the job.” Participants were then directed to complete the personality assessment (either via survey or video game as their condition dictates).

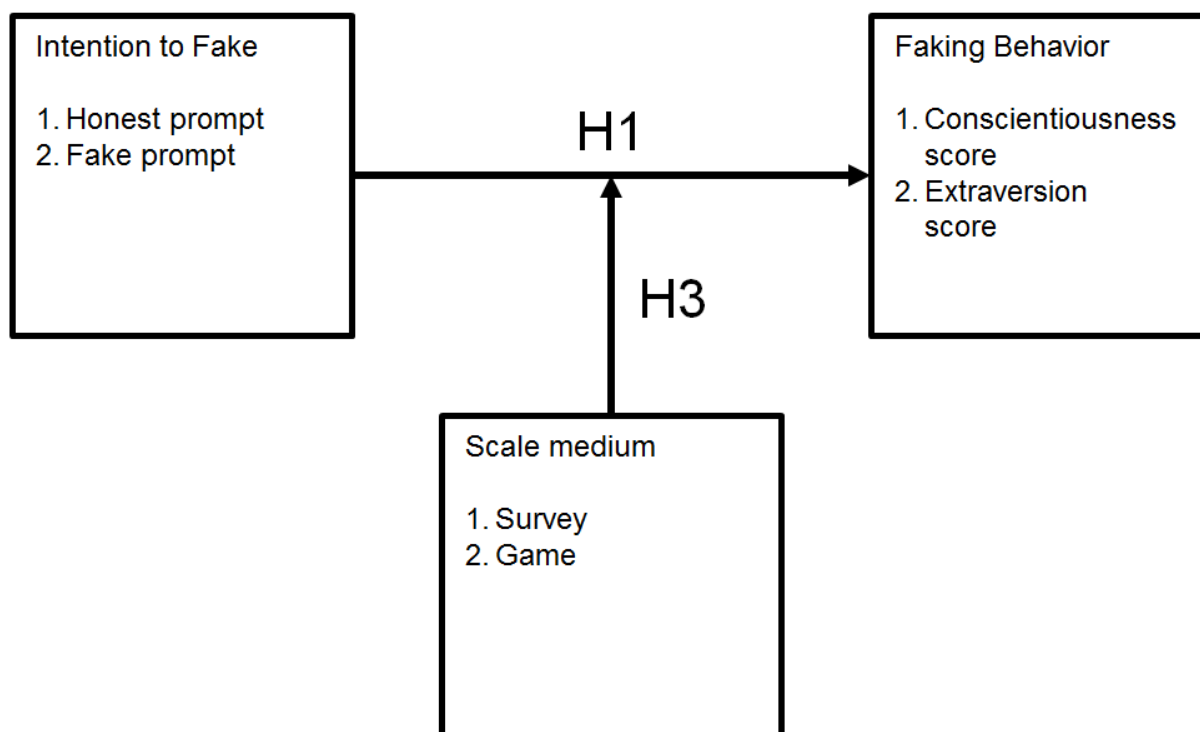
**Time 2.** Approximately 1-3 weeks later (Ployhart & Vandenberg, 2010), participants received a follow-up e-mail requesting their continued participation. If the participant was initially read the “honest” prompt at time 1, then they will be given the “honest” prompt at this time (time 2) and vice-versa. If the participant was initially given the assessment via survey at time 1, then they were given the video game at time 2 and vice-versa. After completing both surveys, participants in all conditions were debriefed and thanked for their participation.

**Table 3.** Experimental Conditions

	<b>Survey First</b>	<b>Game First</b>
<b>Honest</b>	Condition 1	Condition 2
<b>Fake</b>	Condition 3	Condition 4

### CHAPTER THREE: RESULTS

All analyses were conducted using SPSS Statistics 22. An illustration of Hypotheses 1 (a & b) and 3 (a & b) is depicted below in Figure 4. Bivariate correlations between all relevant variables are presented below in Table 4. A series of *t*-tests were conducted in order to test for order effects. All tests returned non-significant *p*-values, less than .05 (a summary of these tests is presented below in Table 5). Therefore, the combination of conditions 1 and 2, and 3 and 4 is supported and was conducted for hypothesis testing.



**Figure 4.** Visualization of hypotheses.

**Table 4.** Correlation Matrix

Variable	<i>M</i>	<i>SD</i>	<i>N</i>	V1	V2	V3	V4	V5	V6	V7
1 – Gender	1.77	0.42	279	-						
2 – Age	21.75	5.46	278	-.32**	-					
3 – VGSE <sup>a</sup>	33.23	9.76	276	-.30**	.07	(.97)				
4 – Conscientiousness (Honest)	3.67	0.64	88	.06	-.05	-.08	(.83)			
5 – Conscientiousness (Fake)	3.79	0.76	97	-.01	.17	.16	N/A	(.90)		
6 – Extraversion (Honest)	3.41	0.65	89	-.12	.09	-.02	.02	N/A	(.83)	
7 – Extraversion (Fake)	3.49	0.75	98	-.15	.11	-.01	N/A	.41**	N/A	(.86)

\*  $p < .05$ , \*\*  $p < .01$ , <sup>a</sup> VGSE = Video Game Self Efficacy

**Table 5.** *t*-tests Justifying the Combination of C1 & C2, and C3 & C4

Medium	Personality Construct	Condition	Condition	<i>t</i>	<i>df</i>	<i>p</i>
Survey	Conscientiousness	1	2	-0.54	75	.59
	Extraversion	1	2	0.68	76	.50
Game	Conscientiousness	1	2	-0.53	45	.56
	Extraversion	1	2	0.78	46	.44
Survey	Conscientiousness	3	4	0.17	85	.86
	Extraversion	3	4	0.35	86	.73
Game	Conscientiousness	3	4	0.49	39	.36
	Extraversion	3	4	1.87	40	.07

**Hypothesis 1a: Intentions to Fake Influence Faking Conscientiousness Scores**

To determine whether or not intentions to fake significantly predicted faking behavior on a conscientiousness test, a mixed model ANOVA was conducted where intentions to fake (manipulated by the “honest” and “fake” prompts coded 1 and 2 respectively) constituted the independent variable, and conscientiousness scores constitute the dependent variable.

Hypothesis 1a is not supported as no effect for intentions to fake was found for conscientiousness scores,  $F(1,52) = 1.89, p = .18$ .

**Hypothesis 1b: Intentions to Fake Influence Faking Extraversion Scores**

To assess whether or not intentions to fake significantly predicted faking behavior on an extraversion test, a mixed model ANOVA was conducted where intentions to fake (manipulated by the “honest” and “fake” prompts coded 1 and 2 respectively) constituted the independent

variable, and extraversion scores constitute the dependent variable. Hypothesis 1b is not supported as no effect for intentions to fake was found for extraversion scores  $F(1,53) = 3.01, p = .09$ .

### **Hypothesis 2a: Conscientiousness Construct Validation**

In order to test whether or not the game medium assesses the same conscientiousness construct that the traditional online survey captures, the two measures should result in equivalent reliability scores on the measures. The two measures of conscientiousness being utilized are split halves of the original 20-item IPIP conscientiousness scale ( $\alpha = .88$ ) with identical subscale reliabilities of  $\alpha = .79$ . Since the two mediums can be said to equivalently measure the same construct if the split-half reliability obtained by correlating the two measures of conscientiousness within-person in the honest conditions meets or exceeds the original test's reliability value ( $\alpha$ ), Hypothesis 2a is supported ( $r_{SB} = .85 \geq \alpha = .79$ ).

### **Hypothesis 2b: Extraversion Construct Validation**

To determine whether or not the game medium assesses the same extraversion construct that the traditional online survey captures, the two measures should result in equivalent reliability scores on the measures. The two measures of extraversion being utilized are split halves of the original 20-item IPIP extraversion scale ( $\alpha = .91$ ) with identical subscale reliabilities of  $\alpha = .87$ . Since the two mediums can be said to equivalently measure the same construct if the split-half reliability obtained by correlating the two measures of extraversion within-person in the honest conditions meets or exceeds the original test's reliability value ( $\alpha$ ), Hypothesis 2b was not supported ( $r_{SB} = .86 < \alpha = .87$ ).



### **Hypothesis 3a: Scale Medium Interact With Intention-Behavior Relationship for Conscientiousness**

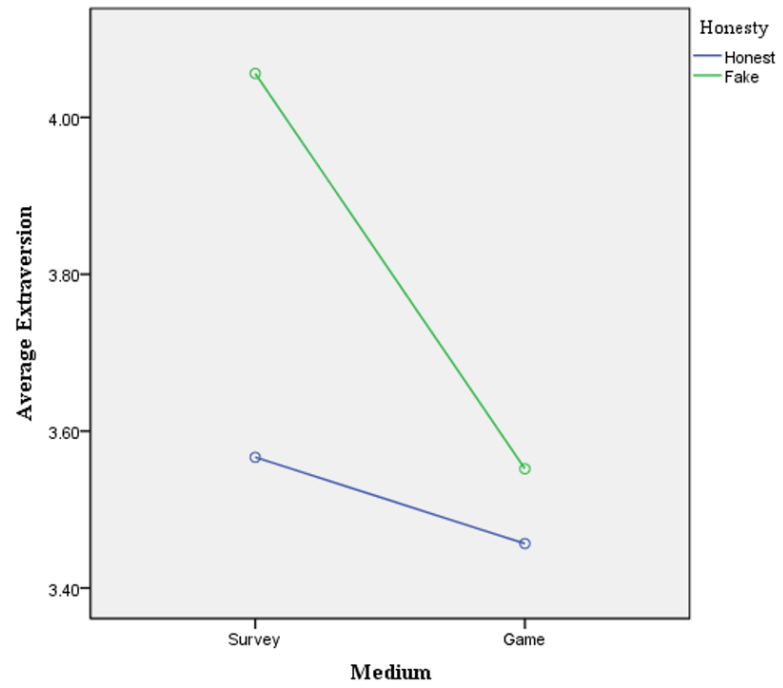
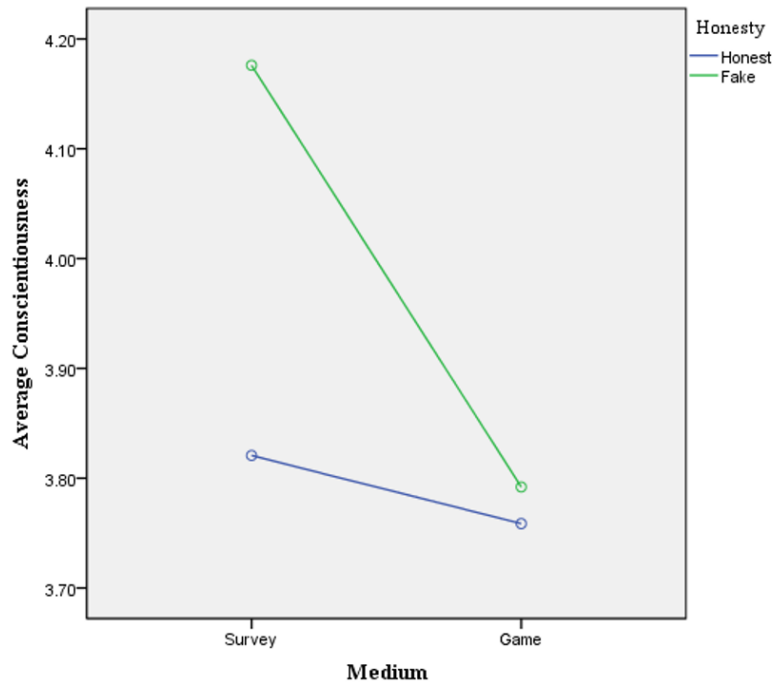
In order to test whether or not the scale mediums (surveys and games) interact with the relationship between intentions to fake and faking behaviors on a conscientiousness test, a mixed model ANOVA was conducted where an interaction term (intention x medium) was added to the ANOVA from Hypothesis 1a. The interaction term was significant and as a result Hypothesis 3a is supported,  $F(1,52) = 4.84, p = .03$ . Figure 5 depicts this relationship below.

### **Hypothesis 3b: Scale Medium Interact With Intention-Behavior Relationship for Extraversion**

In order to test whether or not the scale mediums (surveys and games) interact with the relationship between intentions to fake and faking behaviors on an extraversion test, a mixed model ANOVA was conducted where an interaction term (intention x medium) was added to the ANOVA from Hypothesis 1b. The interaction term was significant and as a result Hypothesis 3b is supported,  $F(1,53) = 5.21, p = .03$ . Figure 5 depicts this relationship below.

### **Post-hoc Analyses**

Following the initial data analyses, all hypotheses were rerun with restricted subsets of the data such that the undergraduate sample ( $N = 260$ ) and graduate sample ( $N = 28$ ) were compared against each other. Some differences with respect to hypotheses being supported or not supported were observed. For the undergraduate sample, Hypotheses 3a ( $F(1,37) = 0.71, p = .40$ ) and 3b ( $F(1,37) = 1.42, p = .24$ ) were not supported whereas the combined sample found sufficient evidence to support them. For the graduate sample, Hypotheses 1a ( $F(1,13) = 35.75, p < .01$ ) and 1b ( $F(1,13) = 5.89, p = .03$ ) were supported, and Hypothesis 2b ( $r_{SB} = .59$ ) was not supported whereas the combined sample drew opposite conclusions.



**Figure 5.** Hypothesis 3a & 3b Interaction Plots

**Table 6.** Summary of Hypothesis Tests

Hypothesis	Description	Test	Test Statistic	Significance	Conclusion
1a	Intentions to Fake Influences Conscientiousness Scores	Mixed Model ANOVA	$F = 1.89$	$p = .18$	Not Supported
1b	Intentions to Fake Influences Extraversion Scores	Mixed Model ANOVA	$F = 3.01$	$p = .09$	Not Supported
2a	Conscientiousness Construct Validation	Split-half Reliability	$r_{SB} = .85$	$.85 > .79$	Supported
2b	Extraversion Construct Validation	Split-half Reliability	$r_{SB} = .86$	$.86 < .87$	Not Supported
3a	Scale Medium Interacts With Intention-Behavior Relationship for Conscientiousness	Mixed Model ANOVA	$F = 4.84$	$p = .03$	Supported
3b	Scale Medium Interacts With Intention-Behavior Relationship for Extraversion	Mixed Model ANOVA	$F = 5.21$	$p = .03$	Supported

## CHAPTER FOUR: DISCUSSION

McFarland and Ryan's (2000) model of faking hinges on the impact of intentions to fake on faking behaviors such that if somebody intends to fake then they will try to fake. Considering Cognitive Load Theory (Sweller, 1994) coupled with Malone's (1980) theory of enjoyment, I posited that participants directed to fake would be less able to fake an assessment if it were presented via a video game platform. The evidence collected provide mixed support for these claims.

The data failed to support both Hypothesis 1a and Hypothesis 1b that intentions to fake would influence personality scores. Given that research has suggested that extroversion is difficult to fake, this particular personality trait may just be difficult to fake even when directed to try. However, conscientiousness, which has been identified as the easiest personality trait to fake (Birkeland et al., 2006; Griffith et al., 2007; Smith & Ellingson, 2002; D. B. Smith & Robie, 2004; Van Iddekinge et al., 2005), also did not demonstrate a significant main effect in faking versus honest conditions. This directed faking study relied on participants being properly directed to fake, therefore it is possible that participants in this study did not effectively fake due to an inadequate emphasis to fake in the faking prompt provided in the study's directions (see Appendix D for the prompts). The faking prompt was inspired by Van Iddeking and colleague's directed faking study where they compared response inflation by "manipulating the experimental instructions (i.e., respond honestly or like a job applicant)" (Van Iddekinge et al., 2005, p. 548),

but simple steps such as emphasizing key words with red text or adding audio cues may have strengthened the manipulation. However, since the study was conducted online and the participants were not supervised, it is unclear how much each participant actually attended to their prompt. That said, there *was* a significant interaction, which warrants caution in interpreting any main effect results. Specifically, both Hypothesis 3a and Hypothesis 3b were supported, such that the assessment medium moderated the relationship between one's intentions to fake and their response to a personality assessment. This moderation suggests that the inflation in scores observed between honest and faking conditions is less extreme in video game mediums than in online surveys (see Figure 5). This interaction will be discussed further below.

Hypothesis 2a and 2b received mixed support. Since the scales that are implemented in the survey and game are split halves of the original 20-item IPIP measure, the corrected correlation between the two split-halves represents a reliability of the combined scale (Gignac, 2009; Spearman, 1910). Further, the split-half reliability can be compared to an alpha since the alpha coefficient is mathematically equivalent to the average of every possible split-half reliability coefficient (Cronbach, 1951; Cronbach & Shavelson, 2004). When the corrected correlations between the 10-item online surveys and the 10-item game surveys are compared to the original 10-item alphas then the hypotheses are supported for the current conscientiousness personality inventory (Hypothesis 2a), but is not supported for the extraversion personality inventory (Hypothesis 2b). This suggests that honestly responding participants treat conscientiousness assessments the same regardless of the medium in which it is presented, but may not treat extraversion assessments identically. However, the observed reliability coefficient for extraversion was quite close to the original extraversion scale's alpha ( $r_{SB} = .86$  and  $\alpha = .87$  respectively), so future research should investigate this using a different sample to determine

whether the failure to support Hypothesis 2b was idiosyncratic to this particular study sample. Nevertheless, the mixed support for Hypothesis 2a and 2b suggests that some personality surveys may be treated slightly differently when presented in an online survey format versus a video game format.

As noted above, Hypothesis 3a and 3b were both supported. This evidence supports the assertion that video games reduce the amount that people fake on personality assessments when directed to fake. This effect may be attributed to the unique experiences involved with playing video games. Namely, video games are fun (Malone, 1980), encourage the player to empathize with the avatar they are controlling (Ekman et al., 2012; Mallon & Lynch, 2014; McCarthy et al., 2013), and demands the player's attention to both manipulate the avatar while engaging with the game environment (Takatalo et al., 2010). This results in the player being cognitively burdened by the internal and external cognitive load (Sweller, 1994) associated with interacting with the dynamic medium while simultaneously enjoying the experience, ultimately being less motivated (or able) to distort their responses than they would in a purely survey-based assessment.

### **Theoretical Implications**

The main theoretical contribution of this effort is the extension of McFarland and Ryan's (2000) faking theory to include the moderating influence of the assessment medium. This study found support for such a moderator, specifically that video games elicit different behavior in assesseees than traditional online surveys such that video games elicit reduced faking behavior, even when directed to fake. These findings, coupled with those of Van Iddekinge and colleagues' (2005) work involving the use of structured interviews in selection settings suggest that alternatives to the more traditional survey assessments have the potential to be superior

methods of assessing desired predictors of job performance. Future research will benefit from considering the nuances involved in the experience of the assessee. In addition to deepening our understanding on the influence of video games or interviews as assessment methods, other methods (e.g., situational judgement tests, assessment centers) should be critically examined for their susceptibility to faking attempts. We must continue to establish the validity of different assessment mediums with respect to faking susceptibility, user experience, and other critical aspects, as it is a certainty that each medium has inherent advantages and disadvantages which meaningfully influence the quality and amount of information that can be gleaned from each assessment.

### **Practical Implications**

Utilizing a directed faking approach, this study serves as both a cautionary tale of the importance of strong faking manipulations, as well as a welcome addition to the literature supporting gamified or video game based assessments as a valid measurement approach in selection contexts using personality. A faking manipulation must convince the participant that their objective is to intentionally distort their responses when appropriate in order to attain a particular score on the given assessment. In other words, the participant should be convinced that they are no longer representing themselves, and are instead essentially role-playing. When this is not done effectively, no differences between honest and faking assessees can be attributed to faking efforts. However, much research suggests that attempts to solicit honest responses in real world selection contexts are often met with increased attempts to fake (cite stuff from above). Thus, the findings from this study that game-based assessments demonstrate less faking are highly relevant to operational settings.

However, the significant interactions in this study suggests that video games may be a viable alternative to traditional assessment strategies. Since video games are fun (Malone, 1980), managers may seek to utilize video game based assessments in an attempt to improve their company's image by having potential (in the case of an assessment being used for selection purposes) or current (in the case of an assessment being used for evaluation purposes) employees enjoy the manner in which they are evaluated. Organizations are always looking for ways to gain a competitive advantage, and if an adjustment of assessment techniques can improve the perception or enjoyment of the assessment without sacrificing validity then it should be considered. Further, if that same assessment can help stave faking attempts, then organizations reap the benefits of an accurate assessment of desired traits that predict performance coupled with a more enjoyable selection process by applicants.

Finally, the return on investment for starting up an automated game platform in lieu of an interview process or other recurrently costly method should offer a financial incentive to consider the use of video games for selection purposes. Although the startup costs have the potential to be high in some industries, in general it has been estimated that "the cost of designing and validating a simulation may not be materially different from the cost of designing and validating more traditional measures" (Boyce, Corbet, & Adler, 2013, p. 20). However, the benefit of implementing a game or simulation is in the efficiency of the system over time, reducing the overall maintenance and/or manpower costs associated with maintaining alternative systems.

### **Limitations**

Although the results are intriguing, this study has some inescapable limitations. The exclusive use of student samples somewhat restricts the generalizability of the findings. The



inclusion of both undergraduate and graduate students was intended to broaden the variety of the sample as much as possible, but by not sampling non-student working adults this study is not able to perfectly capture the full gamut of possible working adults.

A second limitation of this study is the inability to manipulate the game platform. As a result of the platform's rigidity, it was impossible to disentangle the potential differences between responses to the items on the gamified survey and the traditional survey since the items were unable to be manipulated after data collection began. It is possible that any differences observed between the game and survey were actually due to the specific items contained within each scale. However, since the game's and survey's scales were comprised of equivalent items with previously established equivalent reliabilities, this may be unlikely to be the case.

Finally, the prompt used in this experiment may not be a powerful enough manipulation to statistically significantly alter participants' behaviors. Although the significant interactions observed temper the interpretation of direct or main effects, there is no way for this study to examine the potential impact of the manipulation beyond the effects tested by Hypotheses 1a and 1b since no manipulation checks were implemented. It may be that simply prompting a participant to respond honestly or dishonestly via a single statement displayed on a webpage may not be sufficient, and that it may be better or necessary to include more rich forms of communication such as video or audio prompts, or even external performance rewards.

## **Conclusion**

The present study extends McFarland and Ryan's (2000) theoretical model of faking by considering the potentially moderating effect that the assessment medium may play on the relationship between faking intentions and behavior. The results failed to fully support both the original and extended model of faking, but manages to lend some evidence to the notion that

video games may be an equivalent or even advantageous alternative to online surveys. Perhaps this means that our basic assumptions regarding both the nature of faking and the implementation of video games should be reconsidered. As concerns over faking and interest in video games both become more salient topics of consideration, it is finally time that we incorporate these factors into our understanding of how people respond to surveys, particularly in such high stakes settings as selection contexts.

## REFERENCES

- Barrick, Murray R, & Mount, Michael K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel psychology, 44*(1), 1-26.
- Barrick, Murray R, & Mount, Michael K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*(3), 261-272.
- Bedwell, Wendy L, Pavlas, Davin, Heyne, Kyle, Lazzara, Elizabeth H, & Salas, Eduardo. (2012). Toward a taxonomy linking game attributes to learning an empirical study. *Simulation & Gaming, 43*(6), 729-760.
- Birkeland, Scott A, Manson, Todd M, Kisamore, Jennifer L, Brannick, Michael T, & Smith, Mark A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14*(4), 317-335.
- Boyce, A. S., Corbet, C. E., & Adler, S. (2013). Simulations in the selection context: considerations, challenges, and opportunities. In *Simulations for Personnel Selection* (pp. 17-41). Springer New York.
- Chandler, Paul, & Sweller, John. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293-332.

- Chernyshenko, Oleksandr S, Stark, Stephen, Prewett, Matthew S, Gray, Ashley A, Stilson, Frederick R, & Tuttle, Matthew D. (2009). Normative scoring of multidimensional pairwise preference personality scales using IRT: Empirical comparisons with other formats. *Human Performance*, 22(2), 105-127.
- Christiansen, Neil D, Goffin, Richard D, Johnston, Norman G, & Rothstein, Mitchell G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology*, 47(4), 847-860.
- Cronbach, Lee J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3), 297-334.
- Cronbach, Lee J, & Shavelson, Richard J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and psychological measurement*, 64(3), 391-418.
- Dilchert, S, & Ones, DS. (2011). Application of preventive strategies. In M. Ziegler, C. MacCann & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 177-200). New York: Oxford University Press.
- Donovan, John J, Dwight, Stephen A, & Schneider, Dan. (2014). The impact of applicant faking on selection measures, hiring decisions, and employee performance. *Journal of Business and Psychology*, 29(3), 479-493.
- Ekman, Inger, Chanel, Guillaume, Järvelä, Simo, Kivikangas, J Matias, Salminen, Mikko, & Ravaja, Niklas. (2012). Social interaction in games measuring physiological linkage and social presence. *Simulation & Gaming*, 43(3), 321-338.
- Galton, F. (1884). Measurement of character. *Fortnightly Review*, 36, 179-185.
- Gignac, Gilles E. (2009). Psychometrics and the measurement of emotional intelligence *Assessing Emotional Intelligence* (pp. 9-40): Springer.

- Goffin, Richard D, & Christiansen, Neil D. (2003). Correcting personality tests for faking: A review of popular personality tests and an initial survey of researchers. *International Journal of Selection and Assessment*, 11(4), 340-344.
- Goldberg, Lewis R. (1990). An alternative" description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216.
- Goldberg, Lewis 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.
- Griffith, RL, Chmielowski, Tom, & Yoshita, Yukiko. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review*, 36(3), 341-355.
- Griffith, RL, & Converse, PD. (2011). The rules of evidence and the prevalence of applicant faking. In M. Ziegler, C. MacCann & R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 34-52). Oxford, UK: Oxford University Press.
- Holden, R. R. (2006). *Faking on noncognitive self-report: Seven primary questions*. Paper presented at the Educational Testing Services (ETS) Technical Advisory Group (TAG) Conference on Faking in Non-cognitive Assessment, Princeton, NJ.
- Holden, R. R., & Jackson, Douglas N. (1981). Subtlety, information, and faking effects in personality assessment. *Journal of Clinical Psychology*, 37, 379-386.
- Hough, Leatta M, Eaton, Newell K, Dunnette, Marvin D, Kamp, John D, & McCloy, Rodney A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75(5), 581-595.

- Hough, Leatta M, Oswald, Frederick L, & Ployhart, Robert E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9(1-2), 152-194.
- Hunter, John E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29(3), 340-362.
- Hurtz, Gregory M., & Donovan, John J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85(6), 869-879.
- John, Oliver P, Naumann, Laura P, & Soto, Christopher J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, 3, 114-158.
- Judge, Timothy A, Heller, Daniel, & Mount, Michael K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, 87(3), 530-541.
- Mallon, Bride, & Lynch, Ronan. (2014). Stimulating psychological attachments in narrative games engaging players with game characters. *Simulation & Gaming*, 1-20.
- Malone, Thomas W. (1980). *What makes things fun to learn? Heuristics for designing instructional computer games*. Paper presented at the Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems.
- Malone, Thomas W. (1981). Toward a theory of intrinsically motivating instruction\*. *Cognitive Science*, 5(4), 333-369.
- McCarthy, Julie M, Van Iddekinge, Chad H, Lievens, Filip, Kung, Mei-Chuan, Sinar, Evan F, & Campion, Michael A. (2013). Do candidate reactions relate to job performance or affect criterion-related validity? A multistudy investigation of relations among reactions,

- selection test scores, and job performance. *Journal of Applied Psychology*, 98(5), 701-719.
- McCrae, Robert R, & Costa, Paul T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81-90.
- McFarland, Lynn A, & Ryan, Ann Marie. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology*, 85(5), 812-821.
- Mueller-Hanson, Rose, Heggstad, Eric D, & Thornton III, George C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, 88(2), 348-355.
- Ones, Deniz S, & Viswesvaran, Chockalingam. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human performance*, 11(2-3), 245-269.
- Paas, Fred, Renkl, Alexander, & Sweller, John. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1-4.
- Paulhus, Delroy L. (2002). Socially desirable responding: The evolution of a construct. *The role of constructs in psychological and educational measurement*, 49-69.
- Pavlas, Davin, Heyne, Kyle, Bedwell, Wendy, Lazzara, Elizabeth, & Salas, Eduardo. (2010). *Game-based learning: The impact of flow state and videogame self-efficacy*. Paper presented at the Proceedings of the human factors and ergonomics society annual meeting.
- Ployhart, Robert E, & Vandenberg, Robert J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, 36(1), 94-120.

- Salas, Eduardo, & Cannon-Bowers, Janis A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52(1), 471-499.
- Salgado, Jesus F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82(1), 30-43.
- Schmidt, Frank L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15(1-2), 187-210.
- Schmitt, Neal, & Sinha, Ruchi. (2011). Validation support for selection procedures.
- Shenk, D. (2007). *The immortal game: A history of chess*. New York: Anchor Books.
- Smith. (2009). The long history of gaming in military training. *Simulation & Gaming*, 41, 6-19.
- Smith, & Ellingson. (2002). Substance versus style: a new look at social desirability in motivating contexts. *Journal of Applied Psychology*, 87(2), 211-219.
- Smith, D. Brent, & Robie, Chet. (2004). The implications of impression management for personality research in organizations. In B. Schneider & D. B. Smith (Eds.), *Personality and organizations* (pp. 111-138). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Spearman, Charles. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 1904-1920, 3(3), 271-295.
- Sweller, John. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295-312.
- Takatalo, Jari, Häkkinen, Jukka, Kaistinen, Jyrki, & Nyman, Göte. (2010). Presence, involvement, and flow in digital games *Evaluating user experience in games* (pp. 23-46): Springer.



- Tett, Robert P, & Christiansen, Neil D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology, 60*(4), 967-993.
- Topping, Geoffrey D, & O'Gorman, JG. (1997). Effects of faking set on validity of the NEO-FFI. *Personality and Individual Differences, 23*(1), 117-124.
- Truxillo, Donald M, & Bauer, Talya N. (2011). Applicant reactions to organizations and selection systems. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology* (Vol. 2: Selecting and developing members for the organization, pp. 379-397). Washington, DC: American Psychological Association.
- Van Iddekinge, Chad H, Putka, Dan J, & Campbell, John P. (2011). Reconsidering vocational interests for personnel selection: The validity of an interest-based selection test in relation to job knowledge, job performance, and continuance intentions. *Journal of Applied Psychology, 96*(1), 13-33.
- Van Iddekinge, Chad H, Raymark, Patrick H, & Roth, Philip L. (2005). Assessing personality with a structured employment interview: construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology, 90*(3), 536-552.
- Viswesvaran, Chockalingam, & Ones, Deniz S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 59*(2), 197-210.
- Viswesvaran, Chockalingam, Ones, Deniz S, & Schmidt, Frank L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*(5), 557-574.

- Wenzler, Ivo, & Chartier, Don. (1999). Why do we bother with games and simulations: An organizational learning perspective. *Simulation & Gaming*, 30(3), 375-384.
- Worthington, David L, & Schlottmann, Robert S. (1986). The predictive validity of subtle and obvious empirically derived psychological test items under faking conditions. *Journal of Personality Assessment*, 50(2), 171-181.
- Ziegler, M, MacCann, C, & Roberts, RD. (2011a). Faking: Knowns, unknowns, and points of contention. In M. Ziegler, C. MacCann & R. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 3-16). Oxford, UK: Oxford University Press.
- Ziegler, M, MacCann, C, & Roberts, RD (Eds.). (2011b). *New perspectives on faking in personality assessment*. Oxford, UK: Oxford University Press.

## APPENDICES

### Appendix A: Demographic Information

Please answer the questions about yourself and your parents/guardians to the best of your knowledge. If you do not know the answer to the question or the question does not apply to you, please write "N/A" to indicate it is not applicable.

1. What is your sex?

- Male  
 Female

2. What is your age?

\_\_\_\_\_

3. What is your race or ethnic background? (check all that apply):

- White/Caucasian  
 Black/African American  
 Hispanic or Latino  
 Asian  
 Pacific Islander or Native Hawaiian  
 American Indian  
 Alaskan Native  
 Middle Eastern  
 Other: Please Describe \_\_\_\_\_

4. If you chose more than one race or ethnic group in the previous question, which one do you most identify with?

- White/Caucasian  
 Black/African American  
 Hispanic or Latino  
 Asian  
 Pacific Islander or Native Hawaiian  
 American Indian  
 Alaskan Native  
 Middle Eastern  
 Other: Please Describe \_\_\_\_\_

5. Are you fluent in more than one language?

- Yes

No

If so, which languages, in order of most fluent to least fluent?

---

6. Marital Status:

- Single
- Married
- Separated
- Divorced
- Widowed
- Living with Another
- Domestic Partnership

7. Class:

- Freshman
- Sophomore
- Junior
- Senior

If Senior – please indicate your year (i.e., 4<sup>th</sup> year, 5<sup>th</sup> year, etc.)

---

8. How many credit hours are you enrolled in this semester?

---

9. Major: \_\_\_\_\_

10. Minor: \_\_\_\_\_

11. Do you have any other degrees?

- Yes
- No

If Yes, please list them here: \_\_\_\_\_

12. What is your employment status?

- Not Employed, Full-time Student
- Not Employed, Part-time Student
- Employed Part-Time
- Employed Full-Time
- Self-Employed

13. GPA: \_\_\_\_\_

14. SAT Score: \_\_\_\_\_

Verbal: \_\_\_\_\_

Math: \_\_\_\_\_

15. ACT Score: \_\_\_\_\_

16. GMAT Score (if applicable): \_\_\_\_\_

17. Are you the first one in your immediate family to attend college?

Yes

No

18. What is the highest education level of your mother?

High School

Some College

2-year College Degree

4-year College Degree

Some Graduate School

Master's Degree

Doctorate (including a Juris Doctorate – law degree)

19. What is the highest education level of your father?

High School

Some College

2-year College Degree

4-year College Degree

Some Graduate School

Master's Degree

Doctorate (including a JD)

## Appendix B: Video Game Self-Efficacy

Please answer the following questions about how you play videogames using the provided response scale.

1  
(Strongly Disagree)

2  
(Disagree)

3  
(Neutral)

4  
(Agree)

5  
(Strongly Disagree)

1. I can always manage to solve difficult problems within a videogame if I try hard enough.
2. In a videogame, if someone opposes me, I can find the means and ways to get what I want.
3. It is easy for me to stick to my plans and accomplish my goals in a videogame.
4. I am confident that I could deal efficiently with unexpected events in a videogame.
5. Thanks to my resourcefulness, I know how to handle unforeseen situations in a videogame.
6. I can solve most problems in a videogame if I invest the necessary effort.
7. I can remain calm when facing difficulties in a videogame because I can rely on my coping abilities.
8. When I am confronted with a problem in a videogame, I can usually find several solutions.
9. If I am in trouble in a videogame, I can usually think of a solution.
10. I can usually handle whatever comes my way in a videogame.

## Appendix C: Personality

Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Please use the rating scale below to describe how accurately each statement describes you.

1	2	3	4	5
Very Inaccurate	Moderately Inaccurate	Neither Inaccurate nor Accurate	Moderately Accurate	Very Accurate

### Conscientiousness

Item #	Medium	Item
1.	Game	Am always prepared.
2.	Game	Pay attention to details.
3.	Game	Get chores done right away.
4.	Game	Like order.
5.	Game	Follow a schedule.
6.	Game	Am exacting in my work.
7.	Survey	Do things according to a plan.
8.	Survey	Continue until everything is perfect.
9.	Survey	Make plans and stick to them.
10.	Survey	Love order and regularity.
11.	Survey	Like to tidy up.
12.	Game	* Leave my belongings around.
13.	Game	* Make a mess of things.
14.	Game	* Often forget to put things back in their proper place.
15.	Game	* Shirk my duties.
16.	Survey	* Neglect my duties.
17.	Survey	* Waste my time.
18.	Survey	* Do things in a half-way manner.
19.	Survey	* Find it difficult to get down to work.
20.	Survey	* Leave a mess in my room.

\* Indicates reverse-scored items

## Extraversion

Item #	Medium	Item
1.	Game	Am the life of the party.
2.	Game	Feel comfortable around people.
3.	Game	Start conversations.
4.	Game	Talk to a lot of different people at parties.
5.	Game	Don't mind being the center of attention.
6.	Survey	Make friends easily.
7.	Survey	Take charge.
8.	Survey	Know how to captivate people.
9.	Survey	Feel at ease with people.
10.	Survey	Am skilled in handling social situations.
11.	Game	* Don't talk a lot.
12.	Game	* Keep in the background.
13.	Game	* Have little to say.
14.	Game	* Don't like to draw attention to myself.
15.	Game	* Am quiet around strangers.
16.	Survey	* Find it difficult to approach others.
17.	Survey	* Often feel uncomfortable around others.
18.	Survey	* Bottle up my feelings.
19.	Survey	* Am a very private person.
20.	Survey	* Wait for others to lead the way.

\* Indicates reverse-scored items



## **Appendix D: Prompts**

### **Honest**

The following surveys are for research purpose only, and as such, it is important that your responses be as honest as possible. Remember that your responses are confidential, and it is vital that you be as accurate and honest as possible.

### **Fake**

The following surveys are typically used by organizations to hire new employees. Your task today is to respond to the following surveys as if you wanted to get the job at a department store. Recall that companies typically look for employees that are outgoing, talkative, energetic, hard working, self-disciplined, careful, and want to do their job well.

## Appendix E: IRB Approval



RESEARCH INTEGRITY AND COMPLIANCE  
Institutional Review Boards, FWA No. 00001669  
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799  
(813) 974-5638 • FAX (813) 974-7091

June 4, 2015

Philip Ramsay  
Psychology  
4202 E. Fowler Avenue, PCD4118G  
Tampa, FL 33620

RE: **Exempt Certification**  
IRB#: Pro00022383  
Title: The Road Ahead

Dear Mr. Ramsay:

On 6/4/2015, the Institutional Review Board (IRB) determined that your research meets criteria for exemption from the federal regulations as outlined by 45CFR46.101(b):

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:  
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Approved Items:

[The Road Ahead Protocol V#1\\_5\\_19\\_2015.docx](#)

[The Road Ahead Informed Consent V#1\\_6\\_02\\_2015.docx](#)

As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF IRB policies and procedures.

Please note, as per USF IRB Policy 303, "Once the Exempt determination is made, the application is closed in eIRB. Any proposed or anticipated changes to the study design that was

previously declared exempt from IRB review must be submitted to the IRB as a new study prior to initiation of the change."

If alterations are made to the study design that change the review category from Exempt (i.e., adding a focus group, access to identifying information, adding a vulnerable population, or an intervention), these changes require a new application. However, administrative changes, including changes in research personnel, do not warrant an amendment or new application.

Given the determination of exemption, this application is being closed in ARC. This does not limit your ability to conduct your research project. Again, your research may continue as planned; only a change in the study design that would affect the exempt determination requires a new submission to the IRB.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

A handwritten signature in cursive script, appearing to read 'Kristen Salomon', followed by a horizontal line.

Kristen Salomon, Ph.D., Vice Chairperson  
USF Institutional Review Board