Examination of the "Implicitness" of the Anchored Measure of Conscientiousness

by Megan N. Shaw

B.A. in Psychology, May 2003, University of Iowa
M.A. in Industrial/Organizational Psychology, May 2006, George Mason University

A Dissertation submitted to

The Faculty of
The Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

January 31, 2015

Dissertation Directed by

Nicholas L. Vasilopoulos
Chief of Personnel Assessment Research and Development
National Security Agency

David P. Costanza
Associate Professor of Psychology and of Organizational Sciences

The Columbian College of Arts and Sciences of The George Washington University certifies that Megan N. Shaw has passed the Final Examination for the degree of Doctor of Philosophy as of August 29th, 2014. This is the final and approved form of the dissertation.

Examination of the "Implicitness" of the Anchored Measure of Conscientiousness

Megan N. Shaw

Dissertation Research Committee:

Nicholas L. Vasilopoulos, Chief of Personnel Assessment Research and Development, National Security Agency, Dissertation Co-Director

David P. Costanza, Associate Professor of Psychology and of Organizational Sciences, Dissertation Co-Director

Tara S. Behrend, Assistant Professor of Organizational Sciences, Committee Member

Jeffrey M. Cucina, Personnel Research Psychologist, U.S. Customs and Border Protection, Committee Member

© Copyright 2014 by Megan N. Shaw
All rights reserved

Dedication

The author wishes to dedicate this work to her parents, Dr. Bob Shaw and Judy Shaw, for their unending and love and support. Further, I wish to thank Dr. Vasilopoulos for his encouragement, support, and effort in helping me throughout my graduate and professional journey.

Abstract of Dissertation

Examination of the "Implicitness" of the Anchored Measure of Conscientiousness

Over the past two decades, there has been a growing interest in using measures of implicit personality for selection purposes. The enthusiasm stems from theoretical expectations that measures of implicit personality can add to the prediction of job performance provided by traditional, self-report personality measures as well as the belief that they are more resistant to faking. While there have been several attempts to develop measures of implicit personality, few have demonstrated incremental prediction in the criterion space and provided evidence that these measures are truly implicit. This document describes the development and validation of a new conscientiousness measure – the Anchored Preferential Personality Measure (APPM) – which is specifically designed to harness implicit biases and mitigate response distortions. Each APPM item presents two equally desirable statements that anchor the low and high ends of a six-point response option scale. This format is expected to create greater ambiguity when determining the most desirable response, leading examinees to respond on the basis of self-referent cognitive biases that occur largely at an unconscious level. The APPM was validated on a sample of 212 college students in honest and faking conditions. Results showed that the APPM was resistant to faking, significantly correlated with the NEO-IPIP composite measure of conscientiousness, and provided incrementally to the prediction of first-year grade point average (GPA). Evidence that the APPM captures implicit elements of personality was provided in accordance with a double dissociation model (Asendorpf, Banse, & Mucke, 2002), with explicit (self-report) measures adding incrementally to the APPM in the prediction of a more deliberate criteria, and the APPM adding incrementally to the prediction over the explicit personality measure for the spontaneous criteria. The implications and future direction for this research are discussed.

Table of Contents

Dedication iv

Abstract of Dissertation v

List of Figures viii

List of Tables ix

Chapter 1: Introduction 1

Background 1

Problem Statement 5

Study Objectives 7

Chapter 2: Literature Review 8

Models of Implicit Personality 8

Implicit and Explicit Systems and Prediction of Job Performance 15

Measuring Implicit and Explicit Personality 24

Anchored Preferential Personality Measure (APPM) 43

Chapter 3: Methods 57

Participants 57

Measures 57

Procedure 62

Chapter 4: Data Analysis & Results 64

Chapter 5: Discussion 83

References 93

Appendices 118

Appendix A. Consent & Transcript Authorization 118

Appendix B. Participant Invite and Reminder Emails 121

Appendix C. Example List of APPM Items 123

Appendix D. Variable Intercorrelations in Honest and Faking Conditions 125

Appendix E. Analysis Replication Using NEO-IPIP Composite That Matches the APPM Facets (Excludes Discipline and Dutifulness Facets) 127

Appendix F. Analysis Replication Using study order as a Covariate 133

List of Figures

Figure 1. General Model Integrating Implicit with Explicit Social Cognitions. Adapted from “Integrating Implicit and Explicit Social Cognitions for Enhanced Personality Assessment: A General Framework for Choosing Measurement and Statistical Methods,” by Bing et al., 2007a, Organizational Research Methods, 10, p. 355. Copyright 2007 by Sage Publications. Adapted with permission..... 17

Figure 2. Scree Plot of original 90 APPM items. 68

List of Tables

Table 1. Similarities and Differences among Dual Process Models 14

Table 2. Comparison of Implicit and Explicit Measures 44

Table 3. Demographic Data for Participants 58

Table 4. NEO-IPIP and APPM Scale Scores Means (SDs) Overall and By Condition Order 65

Table 5. Number of items retained from the EFA set at 4, 5, and 6 Factors (Varimax Rotation). 69

Table 6. Parallel Analysis & Velicer's MAP test 72

Table 7. Intercorrelations of the APPM and NEO-IPIP 73

Table 8. Model Fit Summary Across Four CFA Models 76

Table 9. Correlations of the APPM and NEO-IPIP with Explicit and Implicit Criteria 77

Table 10. Summary of Hierarchical Regression Analysis for NEO-IPIP with the deliberate and implicit criteria 80

Table 11. Summary of Hierarchical Regression Analysis for APPM with the Deliberate and Implicit Criteria 81

Table 12. Paired Sample T-Tests between Faking and Honest Conditions for the NEO-IPIP and the APPM. 82

Table 13. APPM Item Writing Framework 89

Table 10. Paired Sample T-Tests between Faking and Honest Conditions for the NEO-IPIP and the APPM. 132

Chapter 1: Introduction

Background

There has been an ongoing debate regarding the value of using personality tests for the purpose of personnel selection (Murphy & Dzieweczynski, 2005; Morgenson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). On the one hand, there is evidence that personality tests predict a variety of job-relevant criteria (Barrick, Mount, & Judge, 2001; Huang, Ryan, Zabel, Palmer, 2013; Judge, Rodell, Klinger, & Simon, 2013), reduce adverse impact (Hough, Oswald, & Ployhart, 2001; Schmitt, Clause, & Pulakos, 1996), and, in some cases, add incrementally to the prediction of job performance beyond cognitive ability alone (Barrick, et al., 2001; Schmidt & Hunter, 2004). On the other hand, the operational validities observed for personality tests are often low (r = .00 to .31 with a median range of .13) (Barrick & Mount, 1991; Murphy et al., 2005; Morgenson et al., 2007). Thus, while there is evidence supporting the use of personality tests to make personnel selection decisions, the evidence also suggests there is room for improvement.

Attempts to increase the relationship between personality and job outcomes have focused on two issues. First, selection researchers have turned to models of personality that include multiple self-systems that should differentially relate to important job outcomes (Bing, LeBreton, Davison, Davison, Migetz, & James, 2007a; James & LeBreton, 2012). Most prominent are dual-process models of personality and social cognition that include two self-systems – an implicit system operating at the subconscious level and an explicit system operating at the conscious level – that have

both unique and interactive effects on behavior (Epstein, 1994; 2003; Fazio, 1990; Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellot, 2002; Mischel & Shoda, 1995). Conclusions about the utility of personality testing are based on the results of studies that included direct measures of the explicit system, such as self-report inventories, but gave little attention to indirect measures that tap into the implicit self-system that can drive cognitions and behavior[1].

Research suggests that dual-process models of personality offer a useful framework for specifying how indirect measures of personality can add to the prediction of performance typically observed when relying solely on self-report measures (Lang, Zettler, Ewen, & Hulsheger, 2012). For example, the modest correlations between indirect and direct measures of the same construct suggest that although the implicit and explicit systems share commonalities, they are distinct constructs (Dasgupta, 2010; Harms & Luthans, 2012; Hofmann, Gawronski, Gschwenderner, Le, & Schmitt, 2005). The evidence also suggests that both indirect and direct measures of personality relate to organizationally-relevant criteria such as interpersonal difficulties (Winter, John, Klohnen, & Duncan, 1998), social interactions (Asendorf, Banse, & Mücke, 2002), organizational citizenship behaviors (Johnson, Tolentino, Rodopman, & Cho, 2010), counterproductive behaviors (Bing, Stewart, Davison, Green, McIntyre, & James, 2007b; Burroughs & James, 2005; James, 1998; Johnson et al., 2010), risk-taking behaviors (Dislich, Zinkernagel, Ortner and Schmitt, 2009), task performance (Johnson et al., 2010), and entrepreneurial behavior (Collins, Hanges, & Locke, 2004). Overall, these findings suggest that dual-process models of personality offer a useful framework for

[1] Hereafter, "indirect" is used to describe a measure designed to tap the implicit personality system, and "direct" is used to describe a measure designed to tap the explicit personality system.

specifying how indirect measures of implicit personality can improve the prediction of performance typically observed when using only direct personality measures (Lang et al., 2012).

Secondly, researchers are particularly interested in the extent to which indirect personality measures are resistant to social desirable responding, which can reduce validity in high-stakes settings such as personnel selection (Anderson, Warner, & Spencer, 1984; Berry, Page, & Sacket, 2007; Caldwell-Andrews, Baer, & Berry, 2000; Griffith, Chmielowski, & Yoshita, 2005; McFarland & Ryan, 2000; Mueller-Hanson, Heggestad, & Thornton, 2003; Rosse, Stecher, Miller, & Levin, 1998). According to Paulhus (1984, 1991), social desirability is defined by two dimensions: the first dimension – self-deception – refers to the unconscious tendency to formulate favorably biased, but honestly held self-description; the second dimension – impression management – refers to the intentional distortion of responses in an effort to be viewed favorably by others. Both dimensions of social desirability can factor into how examinees respond to self-report measures of personality in a selection context.

There is ample evidence that direct, self-report personality measures are susceptible to self-deception bias. For example, several researchers have found that self-evaluations are often positively biased due to an inherent desire to preserve a positive self-image. Specifically, individuals tend to overestimate their standing on socially desirable constructs and underestimate their standing on socially undesirable constructs (Alicke, 1985; Dunning, Meyerowitz, and Holzberg, 1989; Kruger, 1999; Kuiper & McDonald, 1982; Zuckerman & Jost, 2001). Evidence also shows that individuals can, and often do, engage in impression management in an effort to present themselves as a

qualified applicants (e.g., Barrick & Mount, 1996; Bowler, Bowler, & Cope, 2013; Griffith et al., 2005; Stark, Chernyshenko, Chan, & Lee, 2001; Tett, Freund, Christiansen, Fox, & Coaster, 2012; Viswesvaran & Ones, 1999), and that faking can have a negative impact on the prediction of job performance (Anderson, Warner, & Spencer, 1984; Griffith et al., 2005; McFarland & Ryan, 2000; Mueller-Hanson et al., 2003; Rosse et al., 1998).

In an effort to mitigate faking, researchers have explored indirect measures of implicit personality. Research suggests that indirect measures may be more resistant to social desirable responding (faking) when compared to direct measures (McDaniel, Beier, Perkins, Goggin, & Frankel, 2009; Steffens, 2004; Vianello, Robusto, & Anselmi, 2010), For example, Vianello et al., (2010) gave participants both a direct and indirect measure of conscientiousness. The authors informed participants that the purpose of the study was to recruit students for a well-paid job at the university, in order to motivate them to respond favorably. They found a significant difference in assessment scores from the faking vs. non-faking condition for the direct measure ($t(70) = 3.13$, $p < .01$), and no significant differences between conditions for their indirect measure. Steffens (2004) found similar results where participants were able to fake the direct measure to a larger degree than the Implicit Association Test (IAT) when they were instructed to either modify their scores on a measure of conscientiousness to appear very conscientious or very unconscientious. Interestingly, Steffens also found that when candidates became more experienced with the indirect measure (IAT) and were provided more information in terms of how it operated (i.e., computing their reaction times), all respondents were able to modify their scores to some level but not to the same degree of the direct measure.

In sum, both theoretical models and empirical evidence suggest that indirect measures of implicit personality predict different work outcomes than do direct measures of explicit personality. Moreover, the evidence shows that indirect measures are less susceptible to the effects of social desirable responding than are direct measures. As a result, several researchers have expressed the need to develop and validate new measures of implicit personality that can be efficiently implemented in a high-stakes setting such as personnel selection (Bing, et al., 2007a).

Problem Statement

While research offers some support for including both indirect and direct measures of personality in a selection battery, for several reasons practitioners have ignored or avoided its application in a personnel-selection context. One reason is that indirect measures are largely associated with the Freudian concept of unconscious drives/needs, considered by most as unscientific and lacking substantiated evidence, and have thus been discarded (Latham, Stajkovic, & Locke, 2010). Second, while indirect measures such as the IAT have been proven as a valuable research tool, it has not been recommended as a selection tool (Greenwald, Uhlmann, Poehlman, & Banaji, 2009), and a universal or well researched alternative for use in personnel selection has not been identified. Third, the few indirect measures that have been used for selection purposes are difficult to interpret and score, often requiring an expert evaluator, which presents challenges to implement on a larger scale. Due to these limitations, there is no universally supported methodology for assessing implicit personality in applied settings.

In an effort to improve the measurement of implicit personality, Shaw and Vasilopoulos (2009) developed and validated a measure of conscientiousness – the Anchored Preferential Personality Measure (APPM) – specifically designed to capture aspects of the implicit self-system and mitigate socially desirable responding. Each APPM item presents two equally desirable statements that anchor the low and high ends of a six-point response option scale. The use of equally desirable anchors is intended to create ambiguity when determining the most desirable response, which leads examinees to respond on the basis of schematic components of the self that operate largely at an unconscious level and influence perceptions of the external world. Overall, the results reported by Shaw and Vasilopoulos find support for the construct validity of the APPM. In addition, while examinees inflated their responses to the APPM, the degree of inflation was significantly less compared to the inflation on the NEO-IPIP (Goldberg, 1993), a traditional self-report assessment.

While the results reported by Shaw and Vasilopoulos (2009) are promising, they did not find support for the hypotheses involving criterion-related validity. Specifically, the APPM did not predict the outcome of interest (GPA). One explanation for this finding is that the component of personality predictive of GPA is best assessed through direct, self-report measures that tap into the explicit self-system. For example, research has shown that indirect measures are likely to predict different spontaneous behaviors (Asendorpf et al., 2002). Since overall GPA likely involves deliberate behaviors on the part of the student, it is possible that GPA was not the appropriate criterion. Another limitation to the study was that there was no formal test that the APPM was truly measuring "implicit" conscientiousness. In absence of this information, it is impossible

to determine whether the APPM is a measure of implicit personality or simply a less fakeable measure of explicit personality.

Study Objectives

The purpose of this research is to examine the construct and criterion-related validity of the APPM using more theoretically-aligned criteria. De Houwer, Teige-Mocigemba, Spruyt, and Moors (2009) outlined a normative analysis of the criteria for assessing the implicitness of a measure. These criteria include verifying the attributes that cause variation in test scores (What Criterion), processes that attribute cause variation in scores (How Criterion), and conditions wherein the assessment measures the attributes (Implicitness Criterion).

In Chapter 2, the theoretical context for the validation study of the Anchored Preferential Personality Measure (APPM) and dual-process information models will first be reviewed. This discussion will include a description of how implicit and explicit processes drive behavior, the unique and interactive effects of indirect measures in predicting organizationally relevant outcomes, and a review of existing measures of implicit personality considered for use in organizational settings. After describing existing measures, an introduction to the APPM – including the development process, initial findings, and the previously unanswered questions to which the current research responds – will be provided. In Chapter 3, a study is proposed that seeks to answer these questions by further validating and testing the implicitness of the APPM. Lastly, the results, discussion, limitations and future research of this study are respectively discussed and provided in Chapters 4 and 5.

Chapter 2: Literature Review

Models of Implicit Personality

Several models have been offered to describe the implicit processing systems. This discussion focuses on the three models most often cited by organizational researchers: the Cognitive-Experiential Self-Theory (CEST; Epstein, 1994), the Reflective and Impulsive model (RIM; Strack & Deutsch, 2004), and the Cognitive-Affective Personality Systems model (CAPS; Mischel & Shoda, 1995). Each model is described in detail below.

Cognitive-Experiential Self-Theory (CEST)

CEST outlines two distinct information-processing modes - the rational self-system and the experiential self-system (Epstein, 1994). These two systems operate by a different set of rules, though they can work together to determine how individuals encode, process, and store information. The rational system is analytical, relies on logical rules of inference, is more susceptible to change, involves slower processing mechanisms, and is guided by more context-general principles. It is through the rational self-system that individuals conceive and formulate beliefs in respect to the environment and self through conscious and controlled information-processing mechanisms. In contrast, the experiential system is affect-driven, guided by heuristics and holistic judgments, involves efficient and rapid information-processing mechanisms, requires minimal resources, and is more resistant to change. The experimental system is where the self and environment are processed and conceptualized unconsciously and automatically through “schemas,” or generalizations that have been established through emotively charged experiences.

The general process by which emotions influence behavior involves the activation of schemata representing emotions tied to memories related to one or more aspects of the situation. These activated emotions subsequently drive behaviors that avoid, threat, or satisfy one or a combination of four needs: 1) maximizing pleasure and minimizing pain; 2) maintaining coherence and stability of one's conceptual system; 3) enhancing self-esteem; 4) relating with others. The manifestation of the experiential self-system may be observed as superstitious thinking, irrational fears, and religious beliefs that shape our understanding of our world (Epstein, 1994).

While the rational and the experiential self-systems are distinct, they work in parallel and interactively influence how we navigate and make sense of the world (Bonner, & Newell, 2010; Epstein, 1999; Epstein, 2003; Epstein & Pacini, 1999). For instance, the experiential system can bias conscious and rational thinking (Epstein, 2003), which can result in maladaptive behaviors. These urges and impulses, however, can be controlled within the rational system using reason, insight (Epstein & Pacini, 1999), self-control, and willpower (Mischel & Ayduk, 2002). The experiential system can also interact with the rational system by activating memories, beliefs, and experiences that serve as implicit information sources that feed into the experiential system (Epstein, 2003).

Reflexive and Impulsive Model

Proposed by Strack and Deutsch (2004), RIM shares many of the features emphasized in CEST. RIM includes two systems of reality: the reflective system and the impulsive system. The reflective system includes propositional representations that are connected through semantic relations. Processes in the reflective system are slow, deliberate, rational, goal-directed, and cognitively resource-intensive. The reflective

system activates behavioral schemata through intentions that guide and direct behavior in new situations. The impulsive system, on the other hand, links environmental cues to associative schemata that are acquired through experiences and stored in long-term memory. The emotional schemata within the impulsive system initiate behaviors directed at either maximizing pleasure (i.e., generating approach behaviors) or minimizing pain (i.e., generating avoidance behaviors) (Deutsch & Strack, 2006; Strack & Deutsch, 2004).

The reflective and impulsive systems are viewed as bidirectional in that activation of one system often provides informational cues that activate the other system. The nature of this interaction is very similar to that described in CEST. The reflective system provides the higher-order executive functions that regulate, overcome, and/or compensate for the impulses generated through the impulsive system. On the other hand, the impulsive system creates levels of arousal and overall positive or negative states that either facilitate or hinder the reasoning processes that occur within the reflective system (Deutsch & Strack, 2006; Russell, 2003).

While the impulsive and reflective systems often operate in parallel, the impulsive system plays a more central role, where it is proposed that it is always "on" and activates behavioral schemata while the reflective system is sometimes disengaged. The relative salience of the two systems, or when one system dominates the other, is said to be a function of arousal. The impulsive system has a greater influence on behavior when an individual experiences either low or high levels of arousal. Low levels of arousal are most likely to occur in situations where tasks are well learned, while high levels of arousal are most likely to occur in situations where tasks are extremely difficult and/or

important. Both systems influence behavior when an individual experiences moderate levels of arousal, such as in situations that require performing a moderately difficult task.

Cognitive-Affective Personality Systems (CAPS)

Proposed by Mischel and Shoda (1995, 1998), CAPS views the composition of personality as a function of "person-characteristics" and their interaction with the environment. According to the CAPS, behavioral variance is a result of individual differences in the availability, association patterns, and activation of cognitive-emotional mediating units in response to a situation (Mischel & Shoda, 1995). CAPS differ from more traditional perspectives of personality in that personality is a function of the CAPS network but also the interaction of that network with the environment. As such, under this theory, inconsistencies in behaviors across situations are viewed as systematic and predictive rather than as idiosyncratic error variance (Shoda, Mischel, & Wright, 1994).

In CAPS, the interplay between the structure of personality and the situation is mediated through five types of cognitive-affective units (CAUs) that explain how we categorize information (encoding), predict outcomes (expectancies and beliefs), value various outcomes (goals and values), establish goals (competencies and self-regulatory strategies), and situationally emote feelings (affective responses).

Unlike the CEST and RIM, which acknowledge the importance of the affective components and motivational orientations of implicit systems, CAPS emphasizes the role of situational cues in defining personality. That is, CAUs form complex associative networks that uniquely interact with the situation that frames and guides behavior. This phenomenon is a reflection of unique intra-individual constructions of if/then situation-behavior relationships that explain both inconsistencies and consistencies in behavior. For example, an individual may contribute as a team player when closely monitored by a

team leader or manager (if p then q_1), but act in an uncooperative and self-interested manner in the absence of being observed (if r then s_1).

Like CEST, CAPS acknowledges two information-processing systems: a hot, emotional "go" system that is impulsive and reflexive, and a cool "know" system, which is cognitive, emotionally neutral, contemplative, flexible, slow and strategic (Metcalfe & Mischel, 1999). The hot system is designed to react and engage in approach and avoidance behavior under the driving principles of reward and punishment. It is a system primarily driven by arousal, emotional reactions, and is specialized for quick emotional processing that generates largely unmediated reflexive approach-avoidance reactions to biologically-significant hot stimuli in the environment (Mischel & Shoda, 1995, 1998; Ayduk, Mischel, & Downey, 2002).

In contrast, the cool system is considered the seat of self-regulation and control and is designed to handle situations that require patience, reflection, flexibility, and formulation of strategic behaviors and complex processing functions (Metcalfe & Mischel, 1999). Mischel and Ayduk (2002) acknowledge that the hot system can be adaptive, but they place greater emphasis on the importance of controlling this system via the cool system under the logic that reacting prototypically and indiscriminately across all situations can lead to mal-adaptive outcomes (Mischel & Shoda, 1995; Mischel & Ayduk, 2002).

Dual-Process Model Comparisons

The process models covered in this section have many similarities. Yet they also differ in potentially meaningful ways. These commonalities and differences are highlighted in Table 1.

All models (CEST, CAPS, and RIM) outline two information-processing systems: one is more automatic and the other more deliberate, but all acknowledge affective and cognitive components influencing these processes. The CEST and RIM models emphasize the adaptive nature of the implicit system, given the limited cognitive resources required to generate an effective resource conserving behavioral response (Epstein, 2003), and emphasize affect, positive/negative valence, and preservation of self (self-esteem and identity) as key components in driving automatic responses. The CAPS model, on the other hand, emphasizes the role of situation in determining and guiding more deliberate and strategic functions. The CAPS model also provides a greater emphasis on the dynamic and purposeful nature of the personality system, taking into account characteristics of the situation and where self-regulatory mechanisms can override the "hot" system.

There are somewhat conflicting points of view in regards to the CAPS model in comparison to the CEST and RIM theory. In the CAPs model, the "hot" system is considered to be more primitive than the deliberate "cool" system, and drives mal-adaptive impulses that need to be controlled by the rational system.

Table 1. Similarities and Differences among Dual Process Models

Characteristic	CEST	RIM	CAPs
Includes cognitive and affective components in processing information.	X	X	X
Describes two operating systems that govern human action. One system that is automatic and fast acting while the in comparison is slower and deliberate.	X	X	X
Primary focuses on the role of the automatic fast action information processing system and emphasize its adaptive functions.	X	X	
Primary focuses on the role of the deliberate and purposeful rational system emphasizing the adaptive functions of this perspective.			X
Provides an interactive account between the implicit and explicit system.	X	X	X
Emphasizes the role of contextual drivers that are conducive to more deliberate and thoughtful action.			X
Emphasizes the role of affect and positive and negative valence in the automaticity of implicit responding.	X	X	

However under the theoretical framework of the CEST and RIM, emotions and affect are considered as viable and adaptive cues that promote behavior and have been shown to predict performance indicators of creativity, humor, aesthetic judgment and self-report measures of empathy and social popularity, while the rational system did not (Norris & Epstein. 2011).

Implicit and Explicit Systems and Prediction of Job Performance

Dual process models provide the theoretical foundation that supports the relevance of implicit and explicit systems for understanding and predicting behavior. Implicit and explicit personality systems are complex, and operate under independent and additive coaction, mediation, moderation, and interactive curvilinearity to predict various behavioral outcomes (Bing et al., 2007a). Which mechanism is activated can vary depending on the construct assessed, the nature of the situation, and the outcome of interest.

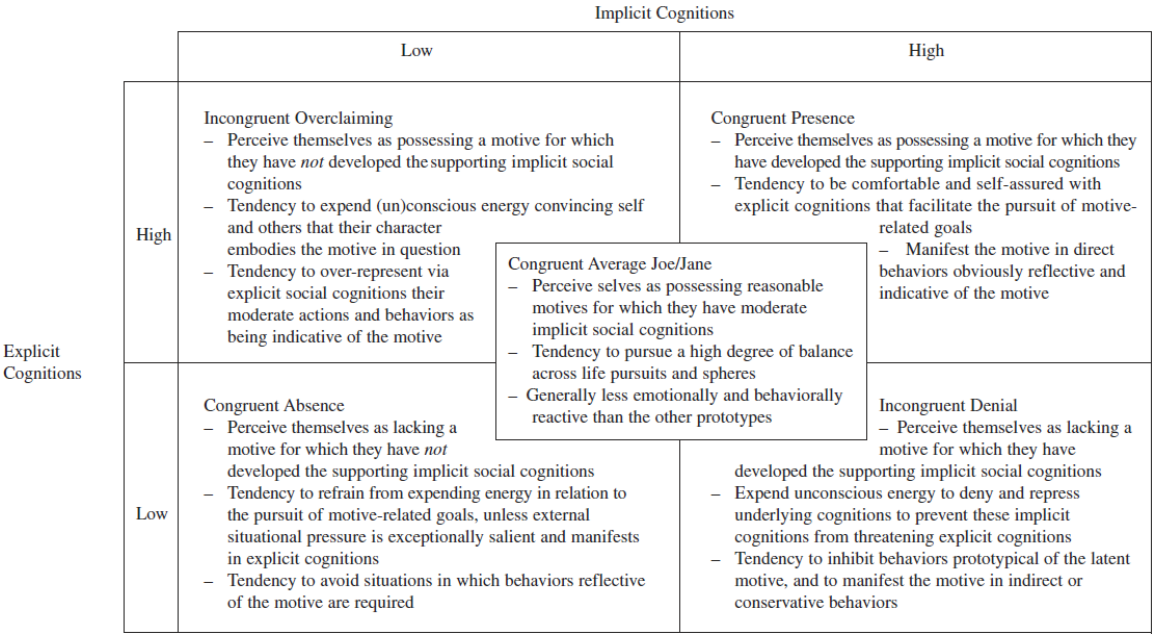
Integrative models have received growing interest and support in the organizational research literature, the Channeling Hypothesis in particular (Bing et al., 2007a; Frost, Ko, & James, 2007; Winter et al., 1988). This model describes the interaction between implicit and explicit systems, namely how implicit personality is channeled through and expressed via explicit personality that drives behavior toward a particular outcome. In addition to interactive effects, the implicit system can have direct effects on outcomes, depending on the nature of the outcome. The *double dissociation effect* illustrates how indirect and direct measures of personality have stronger effects on theoretically aligned outcomes (Asendorpf et al., 2002). The section below reviews this theory and research in more detail in respects to each of these models.

Channeling Hypothesis

Winter (1998) offered a model that integrated the concepts of motives and traits in response to the controversy (similar vs. distinct) within the personality literature. The model lays out how motives and traits interact to provide a more precise and complex interpretation of behavior. Motives refer to underlying internal drives (e.g., goals, wishes, desires) often experienced at the unconscious level, while traits are patterned individual differences, basic tendencies, or stylistic consistencies in cognitions, affect, and behavior (Emmons, 1989) that are accessible to our awareness. Winter argued that motives are expressed, channeled through, and shaped by traits. Thus, if an individual has a high motive, it is more likely expressed in an expected behavioral outcome if the individual's generalized trait predisposition also relates to that motive. To illustrate the interaction of implicit with explicit social cognitions, Bing et al., (2007a), laid out a generalized model as shown in Figure 1.

In the cases where implicit and explicit cognitions are congruent (i.e., Absence: low implicit – low explicit; Presence: high explicit – high explicit), self-perceptions of a trait align with the related implicit drive, resulting in the expected behavioral outcome associated with the explicit and implicit cognition in question. When implicit and explicit cognitions are incongruent (i.e., Overclaiming: low implicit – high explicit or Denial: high implicit – low explicit), however, individuals are less likely to exhibit behaviors that correspond in the expected outcome. In the case of overclaiming, individuals indicate a trait where the motive that drives that trait is lacking. In congruent denial, individuals perceive the absence of a trait where the implicit drive is actually present.

Figure 1. General Model Integrating Implicit with Explicit Social Cognitions. Adapted from "Integrating Implicit and Explicit Social Cognitions for Enhanced Personality Assessment: A General Framework for Choosing Measurement and Statistical Methods," by Bing et al., 2007a, Organizational Research Methods, 10, p. 355. Copyright 2007 by Sage Publications. Adapted with permission.



Lang et al., (2012) tested the channeling hypothesis by examining the interaction of explicit extroversion and achievement with implicit motives (power, affiliation, and achievement) on the prediction of task and contextual performance. The results showed that accounting for the interaction between motives and traits improved the prediction of both task performance and contextual performance. The strongest relationships with both task and contextual performance were observed when the implicit motive and explicit trait were congruently high or low. The weakest relationships were observed for individuals engaged in incongruent overclaiming (high explicit/low implicit) or incongruent denial (low explicit/high implicit). One explanation for the weaker prediction could be that incongruent over-claimers may attempt to engage in the behavioral pattern of the trait in question (e.g., taking on challenging work) but do not have the implicit drive to push through and sustain that behavior in order to achieve the desired end. On the other hand, individuals who demonstrate incongruent denial (low explicit but high implicit) may have reduced performance because they do not subscribe to social-extrinsic indicators to success (e.g., recognition, promotions) and are ultimately dissatisfied with the job because their implicit achievement drives are not being met.

In addition to achievement, support for the Channeling hypothesis has also been demonstrated with implicit affiliation motive and extraversion (Lang et al., 2012; Winter, 1998), implicit power and extraversion (Winter, 1998), and implicit/explicit assessments of aggression (Frost, Ko, & James, 2007; James & LeBreton, 2012). While the research on the Channeling hypothesis provides a promising theoretical framework for integrating indirect and direct measures of personality in a selection context, it fails to take into

account the direct effects of implicit personality on behavior and the conditions under which this occurs.

Double Disassociation

The double dissociation effect describes the predictive pattern of implicit and explicit measures, whereby implicit personality is the best predictor of behaviors that are more difficult to control (e.g., physiological responses to stress and nonverbal shyness behaviors), while deliberate behaviors are best predicted by explicit personality (Asendorpf et al., 2002). The double dissociation effect has gained support in a variety of contexts. For example, Asendorpf et al. (2002) demonstrated that deliberate behavioral indicators of shyness (i.e., speech and gestures) were better predicted by direct shyness measures, while spontaneous behavioral indicators of shyness (i.e., body tension, gaze aversion) were better predicted by indirect shyness measures.

Rudolph, Schroder-Abe, Riketta, and Schutz (2010) also found similar patterns in the context of self-esteem. In this research, the authors coded indicators of self-esteem via adaptors (unintentional spontaneous gestures such as nervous mouth movements, lip biting/licking/pressing, squeezing forearm, playing with fingers) and illustrators (deliberate gestures utilized to underlie spoken words such as hand movements to emphasize a phrase, controlled posture/speech) (Ekman & Friesen, 1969). Support has also been found in the domain of deliberate versus spontaneous helping behaviors (Perugini, Conner, & O'Gorman, 2011), explicit versus implicit romantic attraction (Eastwick, Eagly, Finkel, & Johnson, 2011), and deliberate versus spontaneous racial behaviors (Dovidio, Kawakami, & Gaertner, 2002; McConnel & Leibold 2000).

The hypothesized direct effects of indirect measures of implicit personality and direct measures of explicit personality can be applied to the context of the work

environment. Direct measures may be more predictive of work behaviors that can be more easily controlled and influenced by more purposeful and deliberate action, whereas indirect measures may be more predictive of work behaviors that require less attention to execute and are more automatic in nature. Typically, job performance is considered deliberate and purposeful. However, there are some types of performance criteria that are more spontaneous in nature and thus better predicted by indirect measures of implicit personality. It is also possible that some performance outcomes are a function of a combination of implicit and explicit personality, in which case both indirect and direct measures might provide an interactive effect of incremental validity. Three performance models, most commonly used by organizational psychologists can shed light on when work related behaviors will be best predicted by implicit or explicit personality will be reviewed. These models are: task versus contextual performance, (Borman & Motowidlo, 1993; Organ, 2006), organizational citizenship behaviors (Organ, 2006), and Adaptive Performance (Pulakos, Schmitt, Dorsey, Arad, Hedge, & Borman, 2000).

Task versus Contextual performance

Borman and Motowidlow (1993) differentiated two classes of performance: *task performance*, including behaviors that directly or indirectly contribute to the technical core of jobs that result in the services or goods of the organization's product, and *contextual performance*, including behaviors that contribute to the psychological and social environment in which the technical core resides. Research has shown that, while experience and job knowledge more closely relates with task performance, personality variables are more predictive of contextual performance (Motowidlo & Van Scotter, 1994).

Both implicit and explicit personality systems can drive contextual performance, though each may contribute to varying degrees depending on the particular behavior of interest. Borman and Motowidlo (1993) outlined five dimensions of contextual performance: volunteering, persisting with extra effort, helping and cooperating, following organizational rules and procedures, and endorsing organizational objectives. Three of these contextually-oriented behaviors – volunteering, following organizational rules and procedures, and endorsing organizational objectives – may be better predicted by explicit measures, as execution of such behaviors is likely to result in more immediate extrinsic recognition and reward. Given that the explicit personality system is deliberate and purposeful, self-report measures may be more predictive of these types of behaviors, as they are likely to be more immediately reinforced with praise and recognition.

On the other hand, persisting with extra effort and helping and cooperating are more intangible and may be more spontaneously activated, suggesting that measures of implicit personality may be more predictive of job performance. Lang et al. (2012) demonstrate that both task and contextual performance were predicted by the interaction of implicit and explicit systems, though the interactive effect-size of implicit and explicit personality were stronger for contextual as opposed to task performance.

Organizational Citizenship Behaviors

OCBs are similar in concept to Borman and Motowidlo's (1993) concept of contextual performance and have been defined as "individual behavior that is discretionary, not directly or explicitly recognized by the formal reward system, and that in aggregate, promotes the efficient and effective functioning of the organization" (Organ, 1988, pg.4). OCBs show similar predictive patterns as with contextual

performance, as they more strongly correlates with personality in comparison to task performance (r =.28 *vs* .19) (Chiaburu, Oh, Berry, Li, & Garnder, 2011).

Motowidlo and Boorman (2000) argued that the labels used to classify these discretionary behaviors, whether it is OCBs or contextual performance, are less important than the behavioral dimensions that are of research interest. For OCBs, the types of behaviors that have been investigated include sportsmanship, conscientiousness, civic virtue, altruism, helping, courtesy, cooperative behavior, and interpersonal facilitation. Some of these behaviors, such civic virtue, sportsmanship, and conscientiousness are more likely to be informally or formally recognized or rewarded than altruism, helping, courtesy, cooperative behavior, and interpersonal facilitation. Given the incentive for recognition and reward, direct measures may be more predictive than indirect measures, whereas the reverse may be true for those behaviors that are less likely to be noticed, recognized, or publicly noted.

Perugini et al. (2011) demonstrated that when examining the domain of helping behavior, explicit personality was a better predictor of self-report behaviors of volunteering in the past year and month, but that implicit personality better predicted the actual time volunteered by the participants during the experiment. Johnson and Soboe (2011) also demonstrated that indirect measures of interdependent self-concept (i.e., self-identity of worth strongly associated with high-quality relationships and appraisal from others) predicted task performance but more strongly predicted OCBIs and OCBOs in comparison to direct measures.

Adaptive Performance

Adaptive performance is defined as an effective response to a significant change in situation (Pulakos et al., 2000). The authors laid out a taxonomy of adaptive

performance and identified eight dimensions of adaptability: 1) solving problems creatively, 2) dealing with uncertain or unpredictable work situations, 3) learning new tasks, technologies, and procedures, 4) demonstrating interpersonal adaptability, 5) demonstrating cultural adaptability, 6) demonstrating physically-oriented adaptability, 7) handling work stress, and 8) handling emergencies or crisis situations.

Personality can drive our ability or inability to adjust and adapt to changing environments. In a meta-analysis conducted by Huang, Ryan, Zabel, and Palmer (2014), the authors examined seventy-one independent studies and found that while achievement and emotional stability significantly correlated with adaptive performance (ρ = .18 and .31 respectively $p < .05$). Emotional stability promotes an internal resilience to change and facilitates an approach (fight) rather than an avoidance (flight) response (Nettle, 2006) while achievement orientation can provides the drive to move forward and identify and engage in behaviors that can actually remedy the changing situation (Huang et al., 2014).

In addition to explicit personality measures, implicit personality traits are also expected to predict adaptive performance, and possibly to a larger extent, given the strong theoretical linkage of the predictor/criterion content. The first six adaptability dimensions listed above, describe domains in which the situation is ill-defined, ambiguous, and/or unfamiliar, while the remaining two described states of heightened arousal. Each of these adaptability dimensions serves as strong situational cues that can activate schematic experiences and implicit motives, and drive decisions and a specific course of action (Cervone, 2004; Lazarus, 1991; Teglasi, Nebbergall, & Newman, 2011).

Summary

The flow of work across and within jobs requires various types of behaviors that define the criterion domain. To determine the level of predictability of implicit or explicit measures, it is important to understand how they interact for a particular construct, as described via the channeling hypothesis. It is equally important to understand the degree of automaticity or deliberateness of the criteria and to investigate double dissociative effects (i.e., the expected differential predicted patterns with implicit and explicit measures).

Measuring Implicit and Explicit Personality

As previously discussed, organizations that use both indirect and direct measures of personality are likely to obtain a richer, more comprehensive assessment of personality (Bing et al., 2007a). This section introduces common approaches to measuring explicit and implicit personality, differences in their features, the underlying process that drives the implicit response, and the benefits and challenges in using each approach in applied settings.

Categories of Direct and Indirect Measures

Cook and Selltiz (1964) provided a qualitative review of direct and indirect measures of attitudes and classified these instruments into one of four categories. These categories include measures that draw dispositional inferences based on: 1) self-report, 2) overt behavior, 3) reactions or interpretations of partially structured material, and 4) performance on objective tasks. The first category - self-reports - includes direct measures that primarily assess explicit personality. Categories two through five include

indirect measures that primarily assess implicit personality and will be discussed in the following section.

Self-report measures. The most frequently utilized assessments are explicit, self-report-like scales where respondents are asked to rate the extent to which they agree with a statement representing a particular construct in question, such as "I listen to my conscience," or "I excel in what I do." A class of self-report questions includes biodata that ask the respondent more "factual" types of questions about themselves and their work experience. While there is some variance in the types of biodata item types (i.e., empirical, behavioral consistency, rational, and factorial) (Hough & Paullin, 1994), in general they are more behaviorally/experientially based than traditional self-report personality items. The Biodata theory relies on the premise that past behavior is the best predictor of future job performance, so biodata question types are expected to serve as more effective predictors of future behavior (Mael, 1991). Research for biodata assessments have been supportive, including a meta-analytic validity estimate of .35 (Schmit & Hunter,1998); it has also shown a demonstration of incremental prediction of manager-performance ratings beyond that accounted for tenure, general mental ability, and the big five personality traits (Mount, Witt, & Barrick, 2000).

The accuracy of self-report measures is contingent upon the assumption that individuals can provide accurate self-assessments. Unfortunately, the tenability of this assumption is questionable given the plentitude of research consistently demonstrating that individuals overestimate their standing on socially desirable constructs and underestimate their standing on socially undesirable constructs (Alicke, 1985; Dunning, Meyerowitz, and Holzberg, 1989; Kruger, 1999; Kuiper & McDonald, 1982; Zuckerman

& Jost, 2001). This general effect has been investigated under various names such as illusory superiority (Hoorens, 1995), above- and below- average affect (Windschitl, Conybeare & Krizan, 2008), nonselective superiority/inferiority biases (Giladi & Klar, 2002), and egocentrism and focalism (Kruger & Burrus, 2004). While there are slight variations among these theories, all generally support the premise that self-evaluations are inherently biased in a manner that preserves a positive self-image. Paulhus (1991) labeled this tendency to unconsciously formulate favorably biased but honestly held descriptions of oneself as self-deception.

While there may be a dispositional tendency to view oneself positively, contextual factors can influence the extent to which self-deceptive biases enhance self-reports. A particularly strong contextual factor is the extent to which self-report measure results are consequential for the respondent. For example, most individuals applying for a job are aware that being organized, timely, and a team-player are desirable qualities. When completing a self-report measure as part of the personnel selection process, applicants are more likely to access past experiences that confirm rather than disconfirm these desirable qualities. As a result of these biased information processing mechanisms (e.g., selective attention and memory), applicants provide inflated responses they genuinely believe to be true.

There is some evidence that self-deceptive biases can undermine the criterion-related validities of personality measures. For example, Berry, Page, and Sackett (2007) asked 277 managers to complete the big five-personality measure, an impression management scale, and a self-deceptive enhancement scale. Statistically controlling for self-deceptive enhancement significantly increased the relationship between emotional

stability and job performance. Others, however, have failed to show the benefits of controlling for self-deception when using personality measures to predict performance (Barrick & Mount, 1996, Ones, Viswesvaran, & Reiss, 1996).

The second assumption underlying the use of self-report measures is that respondents are willing to provide an accurate self-assessment. Finding that respondents can deliberately distort their responses in an attempt to be viewed favorably by others, however, threatens this assumption. The deliberate inflation of responses has been labeled impression management by Paulhus (1984). It has also been labeled as "faking" or "dissimulation," which have been similarly defined as the deliberate attempt to inflate responses in order to convey an overly positive impression (McFarland & Ryan, 2000; Zickar & Robie, 1999).

The extent to which impression management is likely to influence responses appears to depend on the context in which the self-report measure is completed. For example, research shows that respondents are able to successfully respond to a self-report measure in a manner that makes them appear highly qualified for a job when instructed to do so (Goffin & Christiansen, 2003; Morgeson, 2007; Murphy & Dzieweczynski, 2005; Viswesvara & Ones, 1999). Moreover, several studies have found that job applicants score higher than job incumbents who complete a self-report measure as part of a concurrent validation study (e.g., Barrick & Mount, 1996; Griffith, Chmielowski, & Yoshita, 2005; Stark, Chernyshenko, Chan, & Lee, 2001).

Impression management has shown to limit the extent to which the applicant's score reflects his or her true standing on the construct in question (Smith & Ellingson, 2002) as well as compromise predictive validity (Anderson, Warner, & Spencer, 1984;

Caldwell-Andrews, Baer, & Berry, 2000; Griffith, Chmielowski, & Yoshita, 2005; McFarland & Ryan, 2000; Mueller-Hanson et al., 2003; Rosse et al., 1998). However, there is also evidence to suggest that self-report measures predict performance regardless of the effects of faking (Barrick & Mount, 1996; Christiansen, Goffin, Johnston, & Rothstein, 1994; Hough, Eaton, Dunnette, & McCloy, 1990; Ones, Viswesvaran, & Reiss, 1996).

While researchers continue to debate the impact of self-deception and impression management on the validity of self-report measures, there is sufficient evidence to support their use in applied settings (Christiansen et al., 1994; Ones, Viswesvaran, & Reiss, 1996; Schmitt & Oswald, 2006). As a result, self-report measures are the most commonly used method for assessing constructs such as personality, biodata, work values, and work interests for the purposes of personnel selection. This is not surprising given the ease of administration and scoring, relatively lower costs, and the variety of constructs that can be assessed. That being said, it is important to acknowledge the factors that may influence the extent to which impression management is likely to be a concern, such as when the construct assessed is conscientiousness, the selection ratio, and when impression management weakly correlates with the job (Komar, Brown, & Komar, 2008), as these factors were found to have more severe negative impacts of faking on validities.

Measures of Overt Behavior. Measures of overt behavior use various behavioral incidences as indicators of the construct of interest. Procedures typically present a standardized scenario where the participant either believes it is staged or un-staged (i.e., role-playing). One example is the lost letter technique (Milgram, Marm, &

Hartner, 1965), where the authors dispersed 400 sealed, stamped, and self-addressed envelopes in public areas to see if they would be picked up and mailed. There have been a variety of manipulations of this study ranging from large to small towns and varying recipient type. This technique was designed to measure attitudes through actions in lieu of self-report.

Personnel psychologists conducting assessment centers for selection and promotional purposes often use observing behaviors to draw inferences about personal characteristics. While assessment centers provide high validities, ranging from .25 to .39 (Arthur, Day, McNelly, and Edens, 2003), construct validity is not as strongly supported. Research has shown that correlations between different measures of the same constructs are low, while the correlations between different constructs assessed using the same method are high (Kaufman, Jex, Love, and Libkuman, 1993). The lack of construct validity is likely due to the difficulty assessors have in differentiating among the constructs. For example, there are significant cognitive demands on the assessors that make it very challenging to process the information needed to provide ratings at the dimension level (Woehr & Arthur, 2003). Thus, while using observed behaviors to draw inferences about personality, measuring implicit personality would not be the appropriate methodology, as assessors would have difficulty making the distinction or appropriate inference of "implicit characteristics" based on the observed behavior(s).

Partially Structured Measures. This category includes an array of measures that require individuals to interpret, describe, or provide open-ended responses to ambiguous stimuli. Stimuli are not clearly structured, so responses are largely driven by the respondent's needs and dispositions. Responses are content coded by trained

psychologists who infer underlying dispositions, needs, and motivations. The most commonly known types of these assessments include: the Thematic Apperception Test (TAT; McClelland, 1985), the Miner Sentence Completion Test (MSCT; Miner, 1976), the Word Fragmentation Task (WFT) (Johnson & Steinman, 2009), and the Attributional Style Questionnaire (ASQ) (Seligman & Schulman, 1986).

Thematic Apperception Test. The TAT requires respondents to shape a story around ambiguous stimuli (e.g., who the people are, what they want, the nature of the situation, and causal implications of why and how the story unfolds and how it will be resolved) (McClelland, 1985). Presentation of the ambiguous stimuli force the respondent to make assumptions and to draw conclusions based on those assumptions, which in turn provides insight into their thought process.

The TAT is completely unstructured and somewhat voluntary, in that respondents are fully aware they are divulging self-reflecting information. There is little effort to hide the purpose of the assessment, however the ambiguity of the stimuli provides a disguising element that makes it challenging to determine what is being assessed and how responses are scored. That being said, this approach requires the full participation of the respondent. If the individual is unwilling or has alterior motives, the TAT may not always operate as intended.

Although the TAT is coded and scored by trained professionals, reliabilities have been mixed (Aiken, 1985; Layon & Goodstein, 1997). However, the TAT has demonstrated promising findings in organizational context because it was able to predict a wide range of different organizationally relevant constructs such as job performance, promotability, income, job level, and career choice (Dulsky & Krout, 1950; McClelland

& Boyatzis, 1982; Phelan, 1962; Spangler, 1992; Winter et al., 1998). It is unlikely, however, that the TAT will be incorporated in organizations given the resource requirements, rater expertise requirements, one-to-one administration setting, lack of face validity, and subjectivity in scoring (Bing, 2007a; Steiner & Gilliland, 1996).

Miner's Sentence Completion Scale. Another approach that takes advantage of the open-ended nature of responses is the sentence completion task. The most common sentence completion task used for organizational purposes is Miner's Sentence Completion Scale (MSCS). This measure presents a partial sentence the respondent must complete, such as "I like ___," "What annoys me ___," "My greatest fear ___." The responses are then content-analyzed into categories by experts and scored.

Research supports the use of the MSCS, with interrater reliabilities ranging from.73 and .95 (Butler, Lardent, and Minder, 1983); significant correlations observed between the MSCS with a variety of organizational outcomes such as turnover (Miner, Norman, Smith, & Bracker, 1994), increase on the number of employees in their organization (r = .35), growth in dollar volume sales (r =.31), and their income (r =.36) (Miner et al., 1994). Other criteria predicted by the MSCS include risk taking behaviors (Streufert, 1986), pilot proficiency (Lunneborg, 1970), psychiatric aid performance (Stotsky, 1956), and door-to-door salesman performance (Geisinger, 1960).

The benefit of the MSCS over the TAT is that the design is more parsimonious, demonstrates higher reliabilities, requires less administrative time for the respondent and the assessor, and allows assessment of multiple individuals in a given setting. However, the MSCS shares with the TAT the similar disadvantage respecting the resources and expertise required to produce respondent scores. In addition, relative to the TAT, the

MSCS is more susceptible to response distortions (Redmore, 1976) and correlates more highly with measures of social desirability and acquiescence bias (Mukherjee, 1968).

Word Fragment Test. The word fragment completion task has been a method to capture implicit processes for the past 30 years in psychology research (Tulving, Schacter, & Stark, 1982). Similar to the MSCS, a Word Fragmentation Test (WFT) requires that a respondent complete a list of correctly spelled fragmented words as fast as possible to ensure that words produce reflect the most accessible implicit content (Johnson & Steinman, 2009). For each fragmented word, one of the viable responses is tied to an underlying psychological construct. In the organizational research literature, Johnson, Tolentino, Rodopman, and Cho (2010) developed a WFT to assess affectivity, with ten word fragments assessing positive affectivity (PA) and ten word fragments assessing negative affectivity (NA). An example of a word fragment assessing positive affect is "_ O Y", where a point is given if the examinee inserted a "J" to spell the word "J O Y". An example of a word fragment assessing negative affect would be "_ _ N S E", where a point is given if the examinee inserted the letters "T" and "E" to spell the word "T E N S E." The proportion of target words generated by the respondent is thus used to infer the implicit construct of interest (Gilbert & Hixon, 1991).

Recent research has shown this method as a viable means to study implicit values, attitudes, and goals in organizational research. For example implicit self-concept as measured by a word fragment test was found to predict task performance, citizenship, and counterproductive behavior more than an explicit direct measure of the same construct (Johnson & Saboe, 2011), how none have specifically explored the use of WFT for personnel selection purposes and the extent to which this method is resistant to faking.

Attributional Style Questionnaire. Another type of measure, similar to the sentence and word completion measures, is the *Attributional Style Questionnaire* (ASQ), which capitalizes off projective biases. The ASQ presents twelve hypothetical situations - six that present a good outcome, and six that present a bad outcome. For each scenario, respondents are asked to generate and write down one major cause for each of the hypothetical situations. An example scenario would be: "You have been looking for a job unsuccessfully for some time," and a possible cause a respondent might provide would be "The poor economic environment." The individual is then asked to rate this cause in terms of internality (e.g., Is the cause attributed to you?), stability (e.g., Will the cause be present in the future?), and globality (e.g., Does it influence all other areas in your life?) (Seligman, Abramson, Semmel & Baeyer, 1979).

In comparison to the TAT and MSCS, the ASQ is more structured. While there is some evidence that the ASQ may be fakeable (Corr, 1995), there is far more evidence suggesting that the ASQ is resistant to the effects of faking (Schulman, Seligman, & Amsterdam, 1987). The ASQ has also shown moderate reliabilities and has predicted turnover, turnover intentions, and sales performance (Brewin & Shapiro, 1984; Corr & Gray, 1995, 1996; Proudfoot, Coor, Guest, & Gray, 2001; Seligman & Schulman, 1986;).

In light of these findings, the ASQ may prove to be a useful tool in personnel-selection situations. This procedure has a modest to high level of disguise, is administered quickly, demonstrates moderate reliability, and is predictive of various organizationally relevant outcomes. However, with exclusion of the measurement of "team orientation" (Shapcott & Carron, 2010), one setback is that this format is generally limited to the measurement of depressive attributional style, which is not particularly face

valid and could portray the organization in an inaccurate light to applicants. However, practitioners may become more receptive to utilization of the ASQ if it assessed more work-related constructs or skill sets.

Measures of Performance on Objective Tasks. The last classification of measures within Cook and Seltiz's (1964) framework involves performance on objective tasks. These measures involve assessing individual differences under the illusion of performing an objective task or an ability test. Unlike the measures discussed thus far, all assessments in this category are involuntary and structured. The underlying principle behind these types of assessments is that attributes of the task bias the "correctness" of endorsing one response option over another. Therefore, rather than requiring individuals to introspectively access and voluntarily report their characteristics, this method infers disposition through performance-based tasks that are designed to harness internal schemas.

Reaction Time. One classification of performance-based implicit assessments are reaction-time-based measures, such as the Evaluative Priming Task (Fazio, Jackson, Dunton & Williams, 1995), Semantic Priming Task (Wittenbrink, Judd & Park, 1997), Extrinsic Affective Simon Task (De Houwer et al., 2003), Go/No-Go Association Task (Nosek & Banaji, 2001), Approach-Avoidance tasks (Chen & Bargh, 1999), and Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998). While these procedures each have their unique qualities, all use respondents' reaction time during a categorization task to capture the strength of association between concepts. Given that the IAT is the most popular of these assessments, the following section will focus on this procedure.

The IAT presents the respondent with two sets of categories (e.g., white/black faces and positive words (+w)/negative words (-w)). Possible combinations presented to the respondent might include the following: White (+w), White (-w), and Black (+w), Black (-w). For the first half of the assessment, respondents are presented with one of two sets of the categories (e.g., White (+w) or Black (-w)); they press one key if they see the first pairing and press a different key for the second. The second half follows the same protocol but presents the other two category sets (i.e., White (-w) and Black (+w)).

In this example, faster reaction times to White (+w) and Black (-w) pairings, relative to White (-w) and Black (+w), indicate a stronger association of white with positive valence and black with negative valence (rather than white with negative and black with positive), which is interpreted as a racial bias. While strong associative patterns emerge from the reaction time of the categorization task, it is unclear what these associations actually are measuring. There is some support for the construct validity of the IAT (Nosek, Greenwald & Banaji, 2005; Gawronski, 2002), but some argue it should not be considered a pure measure of associative strength nor of a "true" attitude (Greenwald & Banaji, 2005).

The IAT has the potential for personnel selection purposes, as many applicants can be assessed simultaneously, scores can be calculated quickly, and this method provides greater flexibility in the types of constructs assessed such as the Big Five personality traits (Steffens, 2004; Steffens & Konig, 2006; Grumm, 2007). In addition, this assessment is highly disguised, as self-referent information is provided involuntarily. Even in those instances where an individual is aware of the purpose of the assessment and how his/her responses are scored, it is more challenging to distort responses in

comparison to self-report explicit measures. This being said, research has been mixed in regards to the fakeability of the IAT. There is some research that shows that the IAT is resistant to the effects of faking (Egloff & Schmukle, 2002; Nosek et al., 2007), while other research suggests that one can intentionally influence the IAT's effects (De Houwer, Beckers & Moors, 2007; McDaniel et al., 2009), although the degree of manipulation seems to be less than with explicit measures (Steffens, 2004).

The IAT also provides some flexibility in regard to the types of constructs assessed – though not without limitations, as "naturally contrasting" constructs (such as gender and race/ethnicity) are more easily assessed by the IAT (Greenwald, 2001). The reliability of the IAT has not been consistent across research. Cunningham (2001) demonstrated test/re-test reliabilities (.61) and stronger split-half internal consistency estimates (between .78 and .89), while Nosek, et al. (2007) found a test/re-test reliability of the IAT at a median value of .56 across nine reports. In addition, the IAT has demonstrated low and/or inconsistent prediction of job-relevant criteria. It also requires the use of computers to record reaction time.

Another challenge with the IAT in a personnel selection perspective is that it has low face validity. This provides implementation challenges, as it may be more difficult to secure stakeholder to buy into or use this test as a hiring tool, and/or results in negative applicant reactions. More importantly, if selection decisions are challenged in court, they may be more difficult to defend to a lay audience. For these reasons it is not surprising that many researchers do not presently advocate the use of the IAT in personnel selection situations (Schnabel, Asendorpf, & Greenwald, 2008).

Conditional Reasoning Measure. There are other alternatives to the IAT where purpose of assessment is disguised as a knowledge test or an ability test. The general design of these measures is to create distracter response options that are clearly wrong (and unlikely to be endorsed) and to target response options that express various levels of the attribute of interest. The assumption is that the response option endorsed is more consistent with the respondents' disposition than are the other response options (Jones & Kohler, 1958).

Tests disguised as syllogistic reasoning tests have also been developed under the same premise, and it was found that reasoning processes were influenced by attitudinal dispositions (Waston, 1925; Thistlethwaite, 1950). Thistlethwaite (1950) argued that attitudes and beliefs influenced reasoning processes and that the endorsed response was largely based on the content rather than the logical structure of the response options. Thistlethwaite described this tendency of dispositions filtering logical judgment as "a tendency for the normal and efficiently functioning organism to isolate certain interpretations of its environment from other intrapersonal wishes, needs, and motives which block correct perception and judgment" (p. 456).

The *Conditional Reasoning Measure* (CRM) is a more recent measure that follows suit with previous performance-based assessments (James, 1998). This method presents respondents with an assessment that on the surface appears as a logical reasoning test. Each conditional reasoning item presents the applicant with a vignette and four options. Two options are clearly irrelevant (i.e., "wrong") and are not scored. The other two options are both correct but express high or low levels of the trait in question.

A respondent's disposition is subsequently inferred based on endorsement of the option that is consistent with his/her standing on the construct of interest.

James (1998) argued that we see ourselves as rational beings – even if our behavior is less than desirable – we develop rationalizations to justify our position and preserve our self-worth. These justification mechanisms are largely self-deceptive biases that implicitly and subconsciously shape our reasoning processes. The conditional reasoning measure is specifically designed to capture these implicit biases by shaping the response options around these justification mechanisms.

CRMs have has acceptable reliabilities (α ranging from .74 - .87) (LeBreton, Barksdale, & Robin, 2007) and have shown to predict a variety of organizationally relevant outcomes, such as absences, theft, deception, patrol-officer performance, and attrition (Bing et al., 2007a; Frost et al., 2007; James, McIntyre, Glisson, Green, Patton, LeBreton, Frost, Russell, Ablynski, Mitchell, & Williams, 2005). There is also evidence that CRMs can mitigate faking when respondents are unaware of the purpose of the test (LeBreton et al., 2007). However, faking becomes a concern when respondents are made aware of the test purpose. As stated by LeBreton, et al., "by divulging the purpose of assessment, we took an indirect measure and turned it into a direct measure; moreover, as most direct measures have some degree of transparency, the results from this study reiterate the susceptibility of such measurement systems to faking" (p. 9). This susceptibility to response distortion once purpose of test is known can be problematic in a selection context because even when all precautions are taken, there is no guarantee that the security of the test can be preserved. Furthermore, given that there are high resource

requirements to develop a CRM, the loss in financial cost and subject matter expert time would be substantial if test security was compromised.

In summary, the CRM could be valuable for personnel selection because it demonstrates criterion-related validity, mitigates faking when the purpose of the measure is unknown, is flexible in administration modality, has high degree of disguise and automaticity, has some degree of reliability, does not require administrator expertise, and has moderate face validity and reliability. However, some of the weakness of the CRM include that it is very resource-intensive on the researcher or practitioner to develop, as well as for the respondent to take, it is easily fakeable when the purpose of the assessment is known, and there are limited types of constructs that have been assessed by the methodology (Frost et al., 2007). Lastly, there has been some indication that the CRM correlates with cognitive ability in faking conditions (r = .44) (Filipkowski, 2007). The argument to use personality assessments in work settings is that it provides incremental validity to cognitive ability tests, so if the assessment is highly correlated with cognitive ability, this relationship would remove or limit the incremental value of the assessment (Vasilopoulos & Cucina, 2006).

Situational Judgment Tests. The *Situational Judgment Test* (SJT) is another measure designed for organizational use that would fall under the performance-based measure category. Items on a situational judgment test present a work-related vignette describing a dilemma or problem and a set of response options for actions that can be taken to address the problem. SJTs are not typically designed to target construct-level personality traits, nor are they designed to be an implicit assessment – rather they are a measurement of generalized work skills (e.g., leadership, teamwork, and interpersonal

skills) that consist of a heterogeneous mix of KSAOs (Christian, Edwards, & Bradley, 2011). However, recent research has shown that personality constructs do correlate with SJT scores (r = .25 for agreeableness; r = .26 conscientiousness; r = .31 for emotional stability; r = .06 for extroversion; and r = .09 for openness) (McDaniel & Nguyen, 2001).

Motowidlo, Hooper, and Jackson (2006a) provided an explanatory framework of the process through which personality can be implicitly captured through an evaluation of effectiveness. The authors explain that SJTs indirectly assess personality via Implicit Trait Policy (ITP) - the implicit belief regarding the importance of personality traits in contributing to effective performance. Under the assumption that SJT response options express different trait levels, individuals demonstrate implicit biased responses of effectiveness depending on their standing in relation to the trait in question. Thus, when an SJT response option reflect a particular personality trait, individuals that possess high levels on that trait are more likely to overestimate the effectiveness of that response option and underestimate the effectiveness of response options that reflect low levels of that trait.

To test this model, Motowidlo et al., (2006a) operationalized the ITPs in two ways. The first was by calculating the correlation between respondents' ratings of effectiveness and the level to which the response option expressed a given level of the personality construct. The second was to subtract the sum of the effectiveness ratings given to response options that expressed high trait levels from the sum of those of the effectiveness ratings given to response options that expressed low level of the attribute. These two different methods correlated .95 with one another, and both significantly

correlated with extroversion and agreeableness, but not for conscientiousness (Motowidlo et al., 2006a).

This format takes advantage of naturally occurring implicit self-referent biases that manifest when making evaluations, which largely occur outside of the awareness of the individual. However, implicit biases are not the only factors that contribute to score variance. Experiences manifested through procedural knowledge also contribute to variance on SJT measures in addition to personality (Motowidlo et al., 2006a). Through experiences individuals learn that there is a time and a place for various behaviors – that in some situations high expressions of the trait are desirable, whereas in other situations they are less desirable. The vignette and the response options provide strong contextual cues, which can aid respondent recognition and application of work-related knowledge in addition to their personality. These types of responses involve a deliberate reasoning process, which means that explicit factors, as well as the proposed implicit factors, contribute to SJT score variance.

In addition to cueing experiences, the information in the vignettes and response options might also provide indicators to the respondent as to the true purpose of the assessment, what constructs are being assessed, and recognition of the most desirable response. Research on the fakeability of SJTs are mixed, with some finding that SJTs are susceptible to faking (Peeters & Lievens, 2005) and others finding SJTs are resistant to faking (Motowidlo, Hooper, & Jackson 2006b). While the SJT might capture some implicit elements within the individual, there has not be been sufficient evidence to demonstrate the degree to which the SJT is implicit. Research has only investigated explicit indicators of behavior (e.g., supervisor performance ratings) (Motowidlo, et al.,

2006a,b) and has not examined the extent to which the SJT predicts spontaneous outcomes. If the SJT is a strong measure of implicit personality, it should demonstrate double dissociative effects, where the SJT is more predictive of uncontrolled and spontaneous outcomes rather than deliberate and controlled outcomes.

In summary, the SJT demonstrates promising predictive validities and mitigates faking (Christian, Edwards, & Bradley, 2011). Other benefits are that, relative to other indirect measures, it can be administered under different types of formats to several individuals at once and requires very little administrator expertise. However, pitfalls with this approach are that it is not a pure implicit assessment, and the extent to which it is implicit is unknown. Thus, it is too soon to utilize the SJT for the purpose of capturing implicit personality.

Indirect Measure Contrast and Comparison. Table 2 summarizes the psychometric and practical considerations of indirect measures of implicit self-systems and the strengths and weakness for use in personnel selection. Currently, the CRM, SJT, and ASQ hold the most promise for use in a selection context. However, understanding the benefits and limitations of the different assessments can help facilitate application in work settings. For instance, the CRM could be utilized if there are considerable resources for test development and if steps can be taken to ensure test security. Given the vulnerability to applicant faking when the purpose of the test is known, CRM would be best used in a proctored setting to maximize test security, thereby increasing the difficulty for applicants to learn and/or be coached on the test. If it is desirable to have a more pure implicit assessment, the CRM or the ASQ might be more appropriate, though if construct

validity is less of a concern, the SJT might be the better choice because it yields higher predictive validities and is more face-valid.

All the assessments that have been discussed share common pitfalls. They all require considerable resources to develop and are limited in the constructs that can be assessed. In addition, these types of assessments require considerable time and cognitive effort by the respondent, which can monopolize test-taking time and limit the number of constructs that can be assessed.

To address some of the pitfalls in Table 2 inherent across and/or shared among these types of assessments, a new type of measure is proposed. This new measure address the following gaps on current implicit assessments: 1) reduces resources to develop, 2) reduces resources to take, 3) mitigates faking when the purpose of the measure is known, 4) offers greater flexibility in the constructs assessed, and 5) offers ease in scoring.

Anchored Preferential Personality Measure (APPM)

In an attempt to proactively reduce the extent of response distortions on self-report measures, Shaw and Vasilopoulos (2009) developed the *Anchored Preferential Personality Measure* (APPM), based on the premise that attributes possessed by the respondent can be unobtrusively accessed through judgments in response to a set of stimuli. Research has demonstrated that the self is infused into our perceptions and evaluations of our social world (Dunning and Hayes, 1996; Krueger & Acevedo, 2005).

Table 2. Comparison of Implicit and Explicit Measures

Test Characteristics	TAT	SC	ASQ	IAT	CRT	SJT	APPM	Self-Report
**Indirect Measure Features	Voluntary	Voluntary	Voluntary	Involuntary	Involuntary	Voluntary	Voluntary	Voluntary
**Degree of Disguise	(+) Moderate	(+) Mod - High	(+) Moderate	(+) High	(+) High	(-) Moderate	(+) Moderate	(-) Low
**Degree of Automaticity	(+) Moderate	(+) Moderate	(+) Moderate	(+) High	(+) High	(-) Low - Mod	(+) Moderate	(-) Low
Cognitions Assessed	Underlying needs	Underlying motivation	Attributional Style	Attitudinal preferences, biases and stereotypes	Motive-based biases in reasoning and inference	Implicit Trait Policies	Implicit Trait Policies	Explicit preferences and self-perceptions
Flexibility in Constructs Assessed	(+) Moderate	(+) High	(+) Low	(+) High	(-) Low	(+) High	(+) High	(+) High
Amount of peer-reviewed research	(+) Extensive	(+) Moderate	(+) Substantial	(+) Moderate	(-) Limited	(+) Substantial	(-) Limited	(+) Extensive
Reliability	(+/-) Mixed	(+) High	(-) Low – Mod	(+/-) Mixed	(+/-) Mixed	(+) High	(-) Limited	(+) High
Criterion related validity	(+) High	(+) High	(+) Moderate	(-) Low	(+) High	(+) High	(-) Limited	(-) Low to Mod
Face Validity	(-) Low	(-) Low	(+/-) Mixed	(-) Low	(+/-) Mixed	(+) High	(+) Moderate	(+/-) Mixed
Administrator expertise	(-) Extensive	(-) Extensive	(+) Minimal	(+) Minimal	(+) Minimal	(+) Minimal	(+) Minimal	(+) Minimal
Medium of Administration	(-) One-on-one interview	(+) Paper/Pencil/ Computer	(+) Paper/Pencil/ Computer	(-) Computer	(+) Paper/Pencil	(+) Paper/Pencil/ Computer	(+) Paper/Pencil/ Computer	(+) Paper/Pencil/ Computer
Administration and scoring time	(-) Extensive	(+) Moderate	(+) Minimal	(+) Minimal	(-) Moderate	(-) Moderate to Extensive	(+) Minimal	(+) Minimal
Administration setting	(-) Individual	(+) Individual or group	(+) Individual or group	(+) Individual or group	(+) Individual or group	(+) Individual or group	(+) Individual or group	(+) Individual or group
**Resources to develop (Cost, time, expertise)	(-) Moderate	(+) Minimal	(+) Minimal	(-) Moderate	(-) Substantial	(-) Substantial	(+) Minimal	(+) Minimal
**Applicant cognitive resource demands	(-) Extensive	(+) Minimal	(+) Minimal	(-) Moderate	(-) Mod - High	(-) Mod - High	(+) Minimal	(+) Minimal

For instance, individuals tend to overestimate the likelihood of others engaging in behaviors that align with their own (Burns & Knussen, 2005), inflate the extent to which others share similar thoughts and ideologies (Van Boven, Dunning, & Loewenstein, 2000), and make more confident and extreme inferences of others on traits they deem important within ourselves (Carpenter, 1988; Dunning & McElwee, 1995; Lambert & Wedell, 1991).

Why do these self-referent biases occur in our social judgments? The *Unified Theory of Implicit Cognitions* explains that the self serves as an anchor to which the structure of our social knowledge is linked (Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellot, 2002). Stimuli more closely associated with the self are more familiar and, as a result, lead to systematic biases in favor of that target (Zajonc, 1968). There is a natural drive to preserve the structure of the social-knowledge system and self-esteem by positively evaluating others who hold similar self-referent characteristics (Greewald & Banaji, 1995).

Self-related information has also been reflected in evaluations of political beliefs. Hovland and Sheriff (1952) found that judgments of the social position of African Americans were largely influenced by participants' attitudes (negative or positive) toward African Americans. Given the difficulty of collecting accurate reports by participants on politically controversial topics, the authors proposed using this approach to indirectly capture racial attitudes.

In line with this perspective, and in an effort to reduce the effects of self-deception and impression management, Donald T. Campbell designed an unpublished measure of attitudes towards freedom of speech (Cook & Selltiz, 1964). Each item

included two equally socially desirable statements, each worded either negatively or positively. These statements were placed on a scale capturing opposing ends of an attribute, whereby the respondent indicated which statement was closer to his/her position (Cook & Selltiz, 1964). For example, the statement, "Fascists and Communists are entitled to preach their beliefs in this country," reflects stronger support for freedom of speech and is paired with the statement, "Only those who are in agreement with this country's philosophy of government are entitled to preach their beliefs," which is designed to reflect lower support for freedom of speech. In this way, attitudes can be inferred based on which statement that they evaluate as more true, in lieu of directly asking respondents, "What is your attitude towards freedom of speech?"

Similar in design to Campbell's measure, Shaw and Vasilopoulos (2009) developed the APPM, where each assessment item presented two equally socially-desirable statements designed to express beliefs and opinions an individual with high and low levels of conscientiousness would be more or less likely to judge favorably. Conscientiousness was chosen as the construct to be assessed because it is the Big Five dimension of personality that has been shown to predict performance across occupations, and be highly susceptible to faking. Thus, if this measurement format were found to be effective for conscientiousness, it would also be applicable as a measurement tool for other, less-universally socially-desirable constructs.

APPM Format Criteria

In establishing a format for the APPM, three criteria were established to guide the development and selection of statements used to create items. Specifically, each item had to include statements that: 1) expressed high and low ends of the same construct, 2) were

equally socially desirable, and 3) were designed to enhance the target of evaluation outside the self.

The first criterion in the development of the APPM was that each item had to have a statement paired with a second equally socially-desirable opposing statement on the same construct. This format differs from the forced-choice assessment, which requires respondents to choose one of two statements reflecting different constructs. The comparative judgment of forced-choice measures has the potential to reduce response distortions (Christiansen, Burns & Montgomery, 2005), but ipsativity from this assessment methodology presents psychometric challenges that threaten score interpretation and construct- and criterion-related validity (Brown & Maydeu-Olivares, 2013, Hicks, 1970; Meade, 2004). While there has been some support that utilizing IRT can increase the interpretability of ipsative data and allow for comparisons between individuals (Brown et al., 2013), this approach requires substantially more resources (time, sample-size requirements, and expertise) to develop and validate the assessment. One of the intended goals in the design of the APPM is parsimony and ease in development, validation, and scoring, thus the APPM was developed as a mean to avoid having to address ipsative issues all together.

For the second criterion, statements were designed to be equally socially-desirable by either deemphasizing the appeal in the positive statement or spinning the negative statement in a more socially-desirable manner. This criterion was set to obfuscate the "correct response" as well as to create ambiguity which can enhance cognitive biases (Tversky & Kahneman, 1974). By pairing conceptually antithetical statements, it is expected that self-associative concept links will be activated whereby assimilation

(gravitating towards the self-associated stimuli) and contrasts (pushing away from the stimuli not associated with the self) will drive the implicit response to the question.

While it can be challenging to meet this criterion for a construct as socially desirable as conscientiousness, it is possible given that traits often hold a double-edge of positive and negative connotations. For instance, a highly agreeable individual could be perceived as cooperative, but he/she could also be perceived as a pushover, whereas someone with high levels of self-efficacy could be described as confident but also arrogant. The double-edge of conscientiousness is reflected in the research literature, given that conscientiousness has shown to be the most consistent and robust personality predictor of job performance (Schmidt & Hunter, 1998) yet also detrimental in situations that require adaptability, creativity, and learning acquisition (Feist, 1998; Judge & LePine, 2007; Le Pine, Colquitt, & Erez, 2000; Martocchio & Judge, 1997).

The purpose of the third design criteria was to enhance self-referent biases by removing the self as the target stimuli of evaluation. The APPM is designed to capture how respondents see and reason in their world and draw inferences on their characteristics through that lens, not how they think of themselves. The criterion was also set to reduce irrelevant biases associated with self-report (i.e., impression management and social desirability), while enhancing self-reflecting biases that naturally carry over in our evaluations outside ourselves. So rather than asking "What is your level of conscientiousness?" where self is the target of evaluation, the item would be crafted in a manner in which the statement itself (belief or opinion which expresses varying levels of the construct of interest), and not the individual, is the target of evaluation. For example, "It is better to be stable and follow a specific routine," would be an anchor that

is designed to express the high-end of orderliness, while the paired anchored that reflects the lower-end of orderliness would be, "It is better to be flexible and follow a general routine."

Unlike the SJT and CRM, the APPM does not include a vignette. As a result, it may be less resource intensive to develop, validate, and score. The omission of the vignette is also relevant in that it reduces contextual cues that may help guide individuals to the correct response. Omission of vignettes may also increase the potential of the measure over time and across situations, as the vignettes developed can often include content that can become quickly outdated or can only be applicable in a specific context. The APPM is also expected to require less time and cognitive resources from the applicant, which means there can be more time allotted to assessing additional constructs or incorporating different selection methodologies.

Development of the APPM

To address some of the issues associated with current implicit assessments, the objective of the APPM was to develop a measure that mitigates faking, is relatively easier to develop, has a broader scope of constructs that can be assessed, and is face-valid so that it can serve as a practical personnel-selection tool for use in an organizational setting. Shaw and Vasilopoulos (2009) conducted an initial development and validation study on the APPM. A total of sixty items were designed to measure the six facets that comprise the NEO-IPIP conscientiousness scale – self-efficacy, orderliness, dutifulness, achievement striving, self-discipline, and cautiousness. Before data collection, five subject matter experts (SMEs) reviewed these items (twelve items per facet) to ensure clarity in construct definition and equality in social desirability. Specifically, SMEs were

asked to link each item with one of the six sub-facets of conscientiousness listed above. Items were retained if at least four out of the five SMEs agreed that: 1) the items tapped into the intended construct, 2) each anchor statement was clearly worded, and 3) the anchor statements were equally socially-desirable. From this qualitative analysis twelve items were removed from the item pool resulting in a total of forty-eight items.

As a next step, data was collected on 202 undergraduate students who were instructed to take the same battery of assessment in both honest and faking conditions. Before conducting the item analysis, effect sizes from the mean differences of each item from honest and faking conditions were computed to evaluate whether the APPM could reduce response distortions in comparison to the NEO. For the NEO, over 72% of the items had effect size differences of .60 or greater, whereas only 33% of the APPM items met this criteria. This indicates that while the APPM is not immune to faking, it can reduce the magnitude of faking. For all sixty APPM items, the mean difference ($t(201) = 14.64$, $p < .00$, $d = 1.29$) was lower than that for the NEO ($t(201) = 19.33$, $p < .00$, $d = 1.44$).

Next, an item analysis was conducted to establish the final APPM. An item was excluded if removing the item increased the coefficient alpha, and the corrected item-total correlation was < .15. Through nine iterations fifteen items were removed from the measure, resulting in a total of nineteen items with a resulting reliability of ($\alpha = .80$). Items were also removed if the effect size of the difference between honest and faking conditions were equal to or exceed .65. It was deduced that this large effect size was evidence that the anchors were not equally socially-desirable or that that the "correct" response was too transparent. This final APPM resulted in a thirteen item measure (four

items measuring orderliness; five items measuring achievement; and four items measuring cautiousness ($\alpha = .62$)). None of the original items tapping into self-efficacy, dutifulness, and self-discipline met the inclusionary criteria. Confirmatory factory analysis did support the factor structure of the three-facet APPM conscientiousness scale in both hones and faking conditions.

Validation of the APPM

Shaw and Vasilopoulos (2009) hypothesized that the APPM correlations with GPA would be invariant across honest and faking conditions. This was not supported, as the APPM did not significantly predict GPA. However, self-reported achievement significantly predicted GPA in honest conditions ($r = .25$, $p < .05$), but not in the instructed to fake condition.

While this initial study demonstrated some promise for the APPM, more research is needed to create a more robust pool of items and improve reliability. Although 48 items were initially developed and tested, only 13 items proved useful, and only three constructs were assessed (orderliness, achievement, and cautiousness). In addition, further research needs to examine and test the "implicitness" of the construct and criterion-related validity of this assessment. If the APPM is an indirect measure of implicit personality, GPA may not be the most appropriate criteria because it is a function of deliberate, rather than spontaneous, behaviors. This could explain why GPA was predicted by the self-report assessment of achievement, but not the APPM sub-facet of achievement. This finding is in alignment with what the double dissociation effect would predict, where implicit and explicit measures have stronger effects on theoretically-

aligned outcomes –the former linked to spontaneous outcomes and the latter with more deliberate outcomes (Asendorpf et al., 2002).

A useful framework for establishing the construct validity of measures designed to assess implicit processes was offered by De Houwer et al., (2009). The authors defined an indirect measure of implicit personality as a "measurement outcome that is causally produced by the to-be-measured attribute in absence of certain goals, awareness, substantial cognitive resources, or substantial time" (p. 250). In their normative analysis of indirect measures of implicit constructs, De Houwer and colleagues identify three important considerations that are necessary in order to claim a measure is implicit: 1) identifying "what" attribute results in the measurement outcomes, 2) identifying "how" those attributes causally result in the measurement outcome, and 3) demonstrating that the attribute assessed causes the measurement outcome in an "automatic" fashion, that is, without awareness or outside the control of the respondent.

The implication of this framework is that a singular piece of information or study is insufficient to establish whether a measure is implicit. Rather, a collection of information addressing each of the criteria above through a body of research is necessary. To start the initial steps in the investigation of the APPM as a potential indirect measure of implicit conscientiousness, the goal of this research is to investigate and test the APPM using the first and third criteria by De Houwer et al, (2009).

The "What" Criterion: What Attributes Cause Variation in APPM Scores?

This criterion investigates whether variation in scores on the APPM can be directly attributable to the proposed attribute, or whether it is attributable to some other factor. Evidence can be collected through experimental or semi-experimental research (De Houwer et al., 2009). The most common methods include examining correlations

with existing measures of the same construct and/or the prediction of theoretically relevant criteria. While construct and content validity were established in the Shaw and Vasilopoulos study (2009), the APPM only assessed three of the six constructs the authors initially intended to measure, and the reliability of .62 was lower than desired in high stakes applied settings. This research expands upon the previous research by developing more items in an effort to assess the six facets of conscientiousness and improve the reliability and construct validity of the APPM.

Construct-validation evidence can be gathered by examining the relationship between the new indirect measure in question and existing indirect measures. Unfortunately, there is a dearth of indirect measures of implicit personality available to use as the standard for comparison. Though the relationship between implicit and explicit measures is often low and context-specific (Hofmann, Gawronski, Gschwenderner, Le, & Schmitt, 2005), explicit measures can be used to demonstrate convergent and divergent correlations with theoretically-related and unrelated constructs, respectively.

Hypothesis 1: The APPM facets will demonstrate convergent validity by significantly correlating with their self-report facet counterpart of explicit measures of conscientiousness.

Hypothesis 2: Confirmatory factor analysis will support a six-factor structure of the APPM.

The more persuasive evidence concerning the "what" criteria is the prediction of theoretically connected outcomes. Demonstrating that the measure in question predicts the theoretically-relevant outcome associated with the attribute being measured offers some evidence that the measurement outcome may be due to variance in the attribute. Thus, if the APPM is truly capturing "implicit" personality traits, it should connect with theoretically-relevant outcomes. In other words, the double-dissociative predictive patterns should emerge, through which the APPM is more predictive of spontaneous-performance criteria and less predictive of the deliberate-performance criteria (whereas the reverse is true for the explicit self-reports measures). The deliberate criterion used in this study is GPA, while the time-latency, following the deadline to take the assessment survey, is used as the spontaneous outcome. More information on each of these measures is provided in the research methodology. These findings can help determine whether the attributes (explicit and implicit conscientiousness) are truly the attributes causing variation in scores.

Hypothesis 3a: the NEO-IPIP will have stronger positive relationships with GPA than with the spontaneous implicit criteria.

Hypothesis 3b: The APPM will have stronger positive relationships with the spontaneous implicit criteria than with GPA.

Hypothesis 4a: The NEO-IPIP will demonstrate incremental prediction over the APPM for the GPA.

Hypothesis 4b: The APPM will demonstrate incremental prediction over the NEO-IPIP in predicting the implicit criteria.

The "Implicitness" Criterion: In What Way Does the APPM Provide an Indirect Measure of Attributes?

The "implicitness criterion" examines the conditions in which the attribute automatically causes variation in the measurement outcome (De Houwer et al., 2009). To meet this criterion, it must be clearly specified that the feature is causing the automaticity of the measurement outcome, and there must be empirical evidence showing that the measure actually possesses that feature (De Houwer et al., 2009). As a feature proposed to elicit the automaticity of response, the APPM presents two competing, socially-desirable target stimuli that are specifically designed to express high and low ends of a sub-facet of conscientiousness. As a rule, individuals see the world through a prism defined by their experienced and disposition-based perceptions and evaluations (Dunning and Hayes, 1996; Krueger & Acevedo, 2005; Hovland et al., 1952). Because this occurs at the subconscious level, individuals inherently believe they see the world as it is, without fully realizing that they largely see the world as they are.

One method to evaluate whether a measure possesses this automatic feature is to examine how it operates under proximal and distal goals (De Houwer et al., 2009). If it is the case that self attributes cause the measurement outcome in an automatic fashion, the measurement outcome should be invariant, even when the respondent seeks to control or manipulate the operation of the measurement outcome (e.g., faking). De Houwer et al. discussed this criterion with respect to examining of the implicitness of the IAT via faking studies. The authors argued that the strongest case for the implicitly of a measure

is when the measures is always unaffected by deliberate attempts to modify the measurement outcome.

Given this argument, the IAT was assessed as a measure that fails to robustly meet the criteria, as research has found that while there are some cases where the IAT is resistant to faking (Asendorpf et al., 2002), there have also been others where it is fakable (Steffens, 2004). While respondents can fake the IAT, considerable coaching and practice are required to fake effectively (Rohner, Schroder, & Michela, 2013). Thus, while the IAT isn't always independent of respondent control, it is much more difficult to fake compared to traditional self-report measures and more implicit than many existing indirect measures. This conclusion suggests that the implicitness of a measure lies on a continuum, and the degree to which the measurement can be faked can shed light on that continuum where the measure stands in respect to its level of "implicitness." Thus, if the APPM is truly implicit, it is expected that the measurement outcome should be invariant across honest and faking conditions.

Hypothesis 5: There will be significant difference between faking and honest scores for the explicit measure (NEO-IPIP) but not for the APPM.

Chapter 3: Methods

Participants

A power analysis was computed, and a minimum of two hundred participants was established for this study in order to find a small correlation (r = .25) and to run a confirmatory-factor analysis using the five-to-ten-cases rule-of-thumb per parameter (Bentler & Chou, 1987; Cohen, 1988). The sample consisted of 212 undergraduate students, from a mid-sized, private university on the eastern United States, who participated in the Psychology Department's Subject Pool. Gender, ethnicity, and year in school demographics are shown in Table 3.

Measures

Demographic Questionnaire

The demographic questionnaire included a self-report assessment of GPA, SAT, year in school, ethnicity, gender, and the location where they took the study. The average GPA for the sample was 3.15 (SD=.52) and average SAT score was 1251 (SD = 121). While the distribution was negatively skewed, there was sufficient variance in GPA with values ranging from 1.26 to 4.00. The Participants in this study completed demographics questions in the honest condition and the personality assessments (NEO-IPIP and APPM) in both honest and faking conditions.

Predictor Variables

NEO-IPIP. The established NEO-IPIP personality measure of conscientiousness was used in this study as a standard to compare the efficacy of the APPM, which was also designed to measure conscientiousness (Goldberg, 1993). The sixty item NEO-IPIP

Table 3. Demographic Data for Participants

	Frequency	Percent
Gender		
Male	65	31%
Female	134	63%
Unspecified	13	6%
Ethnicity		
Asian	34	16%
Black	18	8%
Caucasian	115	54%
Hispanic	10	5%
Middle East	12	6%
Unspecified	23	11%
Year in School		
Freshman	72	34%
Sophomore	62	29%
Junior	32	15%
Senior	32	15%
Non-Degree	1	1%
Unspecified	13	6%

measure used in this study included twelve items that each assessed the following facets of conscientiousness: self-efficacy, orderliness, dutifulness, achievement striving, self-discipline, and cautiousness. *Self-efficacy* items measure one's belief that he or she can effectively handle tasks and get things done. *Dutifulness* items tap one's tendency to feel a sense of duty/obligation. *Orderliness* items capture one's preference for routine schedules and planning things in advance. *Achievement-striving* items measure drive. Self-discipline items tap one's desire to be prepared for and start tasks sooner rather than later. Finally, *Cautiousness* items assess one's proclivity to think things through before acting.

Anchored Measure of Conscientiousness (APPM). Seventy-seven new items were developed for this study. These items and the 13 times on the existing APPM provided a pool of 90 items, with 15 items measuring each of the six facets of conscientiousness. In developing APPM items, phrases capturing the high and the low end of each facet were written. An effort was made to create anchors that minimized the favorability of the high end of the facet and to maximum the favorability of the low end of the facet, while simultaneously maintaining the integrity of the construct measured. For example, an item assessing the facet of dutifulness might include the statement, "It is better to follow the rules even if a better way to do the task was found," to anchor the high end of the response scale, and the statement, "It is better to go against the rules when a better way to do the task is found," to anchor the low end of the response scale.

It has been argued that four- and five-point Likert scales are often too coarse and often fail to provide and to adequately capture individual differences on the attribute assessed (Chan, 1994; Russell & Bobko, 1992). As a result, each APPM item presented anchors at the extreme ends of a six-point response scale. The six-point scale was used to

enhance variance of the assessment in order to maximize both validity and reliability of the tool. However, having the six-point scale rather than a five-point scale will force respondents to indicate a preference for one anchored statement over the other. The results section will outline the item-analysis process, decision criteria used to remove an item or participant from the study, and the number of final items and reliability of the APPM.

Cognitive Ability. Permission to access SAT scores was requested following the consent and authorization section of the survey. SAT scores were only released upon approval by the participant (See Appendix A for the consent and authorization form). Out of the 212 individuals who participated in the study, 168 (80%) provided their consent to access their SAT scores.

Criteria Variables

Deliberate Criterion. Cumulative grade point average (GPA) for each participant was gathered through a transcript obtained from the Office of the Registar after all experimental sessions were complete. Academic records were only released upon approval by the participant (See Appendix A for the consent and authorization form). Eighty percent of the participants (N=168) provided their consent. GPA was selected as the criteria of interest, as research has shown conscientiousness to be the strongest predictor of both high-school and college GPA (Noftle & Robins, 2007). GPA was also designated as the "deliberate" criterion, as attainment of a higher GPA is largely driven by controlled behaviors (Noftle & Robins, 2007) (e.g., studying regularly, completing assignments on time, attending class regularly, and seeking out additional help and feedback when needed). These behaviors are purposeful, logical, and deliberate in that

the individuals are fully aware of the implications of their behavior and how they can influence their academic performance.

Spontaneous Criteria. Two measures were used as spontaneous criteria. The first measure was the time lapse between when a participant was first sent the email inviting him/her to participate in the survey to the time when he/she completed the second survey. (See Appendix B for the email reminders.) All participants had signed up to take the survey before the invitation was sent. They were given instructions to complete the first survey within 24 hours, after which the second survey link would be sent to them within two days, wherein they would have another 24 hours to complete that survey.

The second criterion was the number of reminders sent to the participant before he/she finally completed both surveys. For both surveys, participants who did not complete the survey within 24 hours of receiving the initial email were sent a reminder via email to take the assessment. They were then sent another email reminder 24 hours after the last reminder. After a 24 hour period following the third reminder, a participant was dropped from the study. Thus, the range of reminder scores ranged from zero (from participants who completed both surveys within the 24 hour time frame) to six (reflecting candidates who took both surveys but required the maximum number of reminders for each to complete before they were removed from the study).

Participants were unaware that these variables were being used as performance criteria. They were also unaware of the specific hypotheses of this study to avoid deliberate modifications of behavior. These criteria were considered spontaneous, given that individuals had free reign within a 24 hour period to take the survey, at their leisure. Thus, it was expected that individuals who possessed high levels of conscientious would

react and take the surveys expeditiously to ensure its completion, while moderate-level conscientious participants would delay taking the test but would nevertheless complete it within the allotted 24 hour period. On the other hand, it was expected that low conscientious participants would procrastinate, fail to prioritize, and forget to take the survey altogether.

Procedure

Participants were solicited from an introductory psychology research pool and invited to take two thirty-minute online surveys at two different points in time. After signing up for the study, participants were informed they would be contacted in a week with the experiment instructions and survey link. After 100 participants had signed up, they were contacted on the same day to take the survey. For those individuals who followed instructions and took the survey within 24 hours, they were given the link to the second survey and given 24 hours to complete the survey. For any participant that did not complete the survey within the instructed 24 hour time frame, they were sent a reminder 24 hours following the deadline to take the survey. Up to 3 reminders were sent following the initial deadline, and any participant that failed to take the survey after the third reminder, were removed from the study. The survey included identical personality assessments across two different points in time, and participants were instructed to respond either honestly or as a motivated job applicant (i.e., respond in a manner that would make them appear as a highly qualified job applicant). The specific instructions we provided to the candidate included the following verbiage:

> *"Imagine you are applying for a job that is very interesting and pays very well. The application process requires that you take a personality test,*

where you will either be selected or rejected based on your score. Since you really want this job, please respond to the personality questionnaire below in a way that you think will give you the best chance of being hired."

The condition taken first was counterbalanced to control for order effects, which is common practice for repeated measures design. Participants were provided with a link informing them of the purpose of the study, gathering their informed consent to participate, and obtaining their permission to access academic transcripts.

It is important to note that the purpose of this research was not to build and to contribute to the applicant faking literature. Rather, faking was used as a manipulation in the study as a means to an end, in order to test whether the APPM was truly an indirect measure of implicit consciousness by determining whether the effect size differences between APPM scores in honest and faking conditions were smaller than the effect size difference for the self-report NEO-IPIP.

Chapter 4: Data Analysis & Results

Subsequent to the collection and cleaning of data, outlier analysis was conducted to identify cases more than two standard deviations from the average score on the personality assessment. No outlying cases were identified. In addition, response patterns of the assessment were also examined to identify and exclude responses that reflect lack of motivation (i.e., zero variance in responses options selected). Careless and inattentive responding was also examined by investigating the number of times participants consecutively responded with the same answer as recommended by Johnson (2005). The specific index used in this study followed the methodology used by Meade and Craig (2012), where the maximum number of times the same response option was endorsed on consecutive items on a single page was computed for each participant for the NEO-IPIP. For each participant, standardized scores were computed for the average, max, and sum LongString. Participants who had greater than two standard deviations above the mean for all three metrics were flagged as careless responders. Four participants were flagged and due to careless responding. Results were run with these participants included and excluded from the dataset. There were no differences in results or the conclusions drawn, so the four participants were retained in the analyses.

To examine whether the faking manipulation worked as intended, a t-test of the mean difference between faking and honest condition was computed. Results provided support of the faking manipulation, as NEO-IPIP scores were significantly higher for the faking conditions than for the honest condition ($t(176)= 6.06$, $p < .00$) and resulted in a moderate effect size difference ($d = .58$). (See Table 4 for scale means.)

Table 4. NEO-IPIP and APPM Scale Scores Means (SDs) Overall and By Condition Order

		NEO		APPM	
Condition Order	N	Honest	Fake	Honest	Fake
First	88	3.50 (.42)	3.93 (.48)	3.58 (.37)	3.61 (.33)
Second	89	3.76 (.48)	3.80 (.66)	3.60 (.40)	3.63 (.32)
Total	177	3.63 (.47)	3.87 (.58)	3.59 (.38)	3.62 (.32)

To examine whether order of administration in the within-subject study influenced faking or honest personality scale scores, a repeated measure ANOVA (i.e., mixed-design split plot) was run for the APPM and the NEO. Results were not significant for either the APPM ($F(1, 175) = 2.68$, $MSE = .37$, $p > .10$) or the NEO-IPIP ($F(1, 175) = .70$, $MSE = .04$, $p > .10$). Analyses were also run to examine if order of the study condition (faking first or honest first) influenced the magnitude of faking. The magnitude of the difference in scores between honest and faking honest conditions were not significantly different across order conditions (i.e., Fake-Honest vs Honest-Fake) ($\overline{X}_{FHdiff} = .18$, $SD_{FHdiff} = .44$; $\overline{X}_{HFdiff} = .30$, $SD_{HFdiff} = .60$) ($t(88) = -1.5$, $p = .13$). Given these results, order was not used as a covariate for the remaining analyses[2].

The next set of analyses where conducted to investigate the structure and validity of the six facet scales of the APPM. These analyses included an exploratory factor analysis of all items, examining the internal consistency of the conscientiousness facets, and examining the convergent validity with the NEO-IPIP. Once the final APPM was established, t-tests, correlations, and hierarchical regression were computed to examine the "implicitness" of the APPM and test Hypotheses 3, 4 and 5.

Validation of the APPM

Item analyses were conducted on the initial pool of 90 items that included 15 items to assess each of the six facets of conscientiousness the APPM was designed to assess: achievement, efficacy, discipline, orderliness, cautiousness, and dutifulness. While it is assumed that the items measure a given construct, it is possible that some of

[2] Although the order effects were not significant for either the NEO-IPIP or APPM, given the p-value of .104 for the F test of the within (Fake vs Honest) by order (FH vs HF) interaction for the NEO-IPIP, an additional analysis using experiment order as a covariate was run. No differences were found in overall results in comparison to when no covariate was used. Results are provided in Appendix F.

the new items inadvertently measure a different construct or fail to measure any of the conscientiousness facets at all. To test whether the items captured the intended constructs and to reduce the APPM to those working items an exploratory factor analysis was initially conducted, followed by a confirmatory factor analysis. This approach is common practice when evaluating a new measure (Brown, 2006; Gerbing & Hamilton, 1996; Schmitt, 2011).

First, the intercorrelation matrices were evaluated using Bartlett's Test of Sphereicity (Bartlett, 1954) and the Kaiser-Meyer-Olkin (KMO; Kaiser, 1974) statistic to ensure that the matrices were suitable for factor analysis. Results from Battlett's Test of Sphericity (Bartlett, 1950) indicated the correlations matrices were not random $\chi^2 = 7730$, *df* (4005), $p < .000$. While larger sample sizes is desirable, the sample size of 189 of useable data in this study is minimally sufficient given that the KMO (Kaiser, 1974) statistic was .60, which reaches the minimum value of .60 in conducting a factor analysis suggested by Kline (1994). Thus, the correlation matrices were appropriate for factor analysis.

The initial exploratory factor analysis was run with the original 90 items. The visual scree plot was examined indicated evidence for four strong factors with the possibility of two additional factors (See Figure 2). An EFA set at four, five, and six factors using orthogonal (i.e., Varimax) rotations were used. Using a cutoff value of .35 for factor loadings, 59 items retained in the six factor structure, 55 items in the five factor structure, and 55 in the 5 factor structure (see Table 5).

Figure 2. Scree Plot of original 90 APPM items.

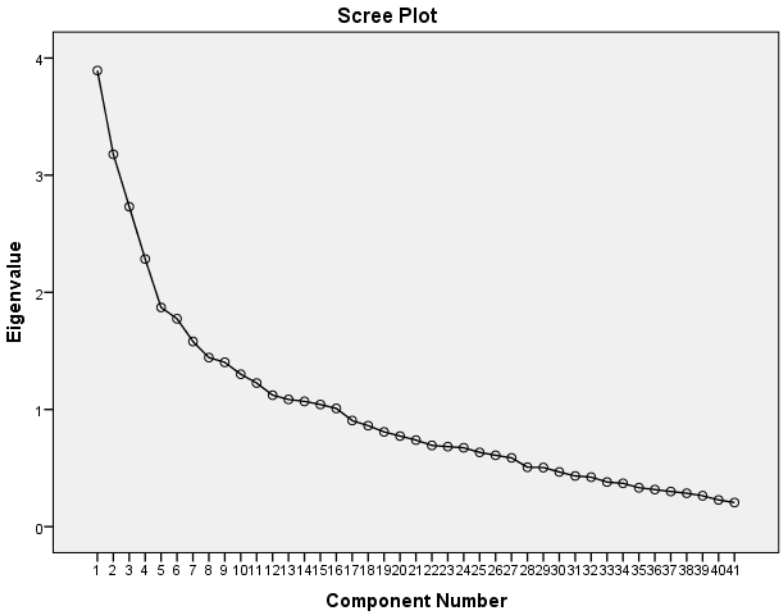


Table 5. Number of items retained from the EFA set at 4, 5, and 6 Factors (Varimax Rotation).

	Total # of Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
4 Factor	55	21	15	9	10	-	-
5 Factor	55	21	13	9	9	3	-
6 Factor	59	15	13	14	8	3	6

Using the combined criteria of eigen value of one there were 16 factors that emerged. Upon a deeper visual analysis of the scree plot and using Cattell's criteria of dropping components after the "elbow" of the plot, a four factor structure emerged. This approach seemed the most viable as the reliability of the five and six factors components were less than $\alpha = .43$.

A hybrid empirical and rational approach was then conducted for the remaining 55 items to remove items operating in theoretically counterintuitive directions. Items that demonstrated significant negative correlations with GPA or with NEO-IPIP composite measure of conscientiousness, and items that conceptually failed to link with the other items in its respective scale were removed from the measure. Based on this analysis, 27 items were removed and the resulting measure consisted of 28 items ($\alpha = .64$).

Examining a scree plot and eigen–values are common approaches for determining the number of underlying factors in a dataset; however both approaches involve a variety of decision rules that, although supported, are somewhat arbitrary in nature and can provide over or underestimates of the number of factors (O'Connor, 2000). An empirical approach to identifying the number of factors can be computed via Parallel Analysis (PA) or Velicer's minimum average partial (MAP) tests, which have been argued to yield more optimal solutions to the number of components problem as they are procedures that are statistically based, rather than rules of thumb (O'Connor, 2000; Wood, Tataryn, & Gorsuch, 1996; Zwick, & Velicer, 1986). For the PA, the number of components extracted is compared to the number that would be due to chance from a random dataset. The For the MAP test, the systematic and unsystematic variance that remains after extractions of increasing number of components is examined (O'Connor, 2000). Both of

these tests were run using O'Connor's 2000 SPSS syntax on the remaining 28 items, and the output was very similar for both analyses (differences were at the hundredths) and resulted in five underlying factors. (See Table 6.)

The results of the earlier item analysis also revealed a five factor structure, however, there were only three items loaded on the fifth factor, and the reliability was low $\alpha < .43$. Given this, a four factor structure was adopted that included the following scales: Achievement Striving (7 items $\alpha = .52$), Cautiousness (6 items, $\alpha = .52$), Orderliness (5 items, $\alpha = .45$), and Dutifulness (10 items $\alpha = .50$). See Appendix C for a sample APPM items.

To test Hypothesis 1, correlations between the APPM scales and NEO-IPIP scales were examined to examine convergent validity. Results provide partial support for Hypothesis 1 where the APPM Achieve, Orderliness, and Cautiousness scales correlated significantly with their respective scales on the NEO-IPIP while the APPM Dutifulness scale did not. Specifically, the highest correlations for the APPM Achievement and Orderliness scales were with the corresponding NEO-IPIP scales. For the APPM Cautiousness scale, there were significant correlations with the corresponding NEO-IPIP scale but, it also correlated to the same degree with the NEO-IPIP Orderliness scale. Lastly, while the APPM Dutifulness failed to significantly correlate with the NEO-IPIP Dutifulness scale, it did significantly correlate with the NEO-IPIP Orderliness scale. (See Table 7.)

The initial item analysis reflected a four-factor structure, however the analysis on the correlation patterns between among the APPM and NEO-IPIP facets suggest a three-factor structure. In the original study conducted by Shaw and Vasilopoulos (2009) the

Table 6. Parallel Analysis & Velicer's MAP test

	Parallel Analysis			Velicer's MAP Test		
Root	Raw Data	Means	Prcntyle	Raw Data	Means	Prcntyle
1	3.35	1.80	1.91	3.35	1.80	1.91
2	2.83	1.68	1.76	2.83	1.68	1.76
3	2.15	1.58	1.66	2.15	1.58	1.66
4	1.85	1.51	1.57	1.85	1.51	1.57
5	1.59	1.44	1.50	1.59	1.44	1.50
6	1.39	1.38	1.43	1.39	1.38	1.43
7	1.28	1.32	1.37	1.28	1.32	1.37
8	1.12	1.26	1.31	1.12	1.26	1.31
9	1.03	1.21	1.26	1.03	1.21	1.26
10	1.01	1.16	1.20	1.01	1.16	1.20
11	0.93	1.11	1.15	0.93	1.11	1.15
12	0.86	1.06	1.10	0.86	1.06	1.10
13	0.83	1.02	1.06	0.83	1.02	1.06
14	0.79	0.98	1.02	0.79	0.98	1.02
15	0.75	0.93	0.98	0.75	0.93	0.98
16	0.70	0.89	0.93	0.70	0.89	0.93
17	0.68	0.85	0.89	0.68	0.85	0.89
18	0.62	0.81	0.85	0.62	0.81	0.85
19	0.59	0.77	0.81	0.59	0.77	0.81
20	0.58	0.74	0.77	0.58	0.74	0.77
21	0.52	0.70	0.73	0.52	0.70	0.73
22	0.46	0.66	0.70	0.46	0.66	0.70
23	0.42	0.62	0.66	0.42	0.62	0.66
24	0.38	0.59	0.62	0.38	0.59	0.62
25	0.37	0.55	0.59	0.37	0.55	0.59
26	0.33	0.51	0.54	0.33	0.51	0.54
27	0.30	0.47	0.50	0.30	0.47	0.50
28	0.30	0.41	0.46	0.30	0.41	0.46

Table 7. Intercorrelations of the APPM and NEO-IPIP

	APPM Composite	APPM Achieve	APPM Orderliness	APPM Cautious	APPM Dutifulness
NEO - Composite	0.37***	0.27***	0.25***	0.27***	0.15**
NEO - Achieve	0.29***	0.29***	0.12*	0.16**	0.14*
NEO - Cautious	0.33***	0.18**	0.22**	0.24***	0.17*
NEO - Orderliness	0.40***	0.13*	0.46***	0.25***	0.22**
NEO- Discipline	0.26***	0.13*	0.19***	0.18**	0.18**
NEO- Dutifulness	0.14*	0.22***	0.00	0.21**	-0.11
NEO-Efficacy	0.12	0.25***	-0.05	0.11	-0.02
NEO – Extraversion	.033	.13	-.04	.06	-.04
NEO - Opennes	-.09	.16*	-.27***	.07	-.19*
NEO – Emotional Stability	.11	.02	.05	.01	.16*
NEO - Agreeableness	-.07	-.10	-.14	.21**	-.18**

*** Significant $p < .001$; ** Significant $p < .05$; *Significant p .10

model that demonstrated fit through a confirmatory factor analysis (CFA) was also a three factor structure (i.e., Achievement, Cautiousness, and Orderliness), which seems to be replicated in this study.

In the Shaw & Vasilopoulos (2009) study, item parcels were created and used as indicators of the facets in lieu of single items. Parcels were created by randomly assigning two items that measured the same construct together (Williams & Anderson, 1994). Parcels were used on account of the perspective that when CFAs are run on single personality items, this can often result in poor fit, as items have secondary and tertiary loadings due to unreliability in measurement (Landis, Beal, & Tesluk, 2000).

CFAs were run in this study using both individual items, parcels, and item level poly-choric unweighted least squares (ULS). The poly-choric ULS were run as research has shown they help to overcome grouping and transformation errors that can result when using Pearson correlations (Brown, 2006; Morata-Ramirez & Holgado-Tello, 2013). These three CFA approaches were examined on four and three-factor models, and also a one factor structure, thus a total of six CFA models were examined to test Hypothesis 2. Model one through three tested a four factor structure (Achievement, Cautiousness, Order, and Dutifulness) while models four through six tested a three factor structure (i.e., Achievement, Cautiousness, and Order).

Chi-squared fit test was examined to examine model fit. If there is model fit, the chi-squared difference of the expected and observed model should be near zero and not significant. Given the sensitivity of the chi-squared statistic to larger sample sizes, the guidelines outlined by Bollen and Long (1993) and Kline (2010) were also explored to examine Model fit. The specific indices examined were the Normed Fit Index (NFI),

Root Mean Square Error of Approximation (RMSEA), Goodness-of-fit index (GFI), Adjusted goodness-of-fit index (AGFI), and Comparative Fit Index (CFI). Each of these metrics examined and the criteria that determined good fit is provided in Table 8 below.

While both models that used item level data failed to demonstrate CFA model fit, the models that used parcels demonstrated marginal fit in supporting the four and three factor structure. The factor loadings for each of these models were examined and while there were a few cases of strong factor loadings, most were poor and weak at best. Thus, hypothesis 2 was not supported.

In summary there was moderate support for the convergent validity of the APPM facets (hypothesis 1) and weak to poor support for the factor structure of the APPM (hypothesis 2), which provides some indication that the APPM is measuring conscientiousness however the facet level measurement is less clear. As a result, the composite APPM conscientiousness scale was computed and used in the analyses conducted to examine Hypotheses 3 through 5[3].

Examining the Implicitness of the APPM

Four separate tests were conducted to investigate the emergence of the double dissociative pattern of the NEO-IPIP and the APPM (see Table 9). To test the hypothesis that the correlation for the NEO-IPIP and the deliberate criteria (GPA) was stronger than for the implicit criteria (*Hypothesis 3a),* a Fisher r-to-z transformation was computed.

[3] Given the equivocal nature of the constructs assessed by the APPM the analysis in the subsequent section was run using a NEO-Composite with all facets but also a NEO-composite with only the facets that the APPM most highly correlated (i.e., achievement, cautiousness, orderliness, and self-efficacy included in the composite with dutifulness and discipline excluded). Results were the same with both NEO-Composites and are presented in Appendix E.

Table 8. Model Fit Summary Across Four CFA Models

Fit Index	Fit Criteria	**Model 1:** 4-Factor (Items)	**Model 2:** 4-Factor (Parcels)	**Model 3:** 4-Factor (ULS)	**Model 4:** 3-Factor (Items)	**Model 5:** 3-Factor (Parcels)	**Model 6:** 3-Factor (ULS)	**Model 7:** 1-Factor (Items)
χ^2	-	665**	-	1002**	679	-	1023**	525**
df	-	344	-	344	347	-	347	350
RMSEA	< .08	.07	.08	.10	.07	.08	.10	.05
NFI	> .90	.42	.63	.40	.41	.61	.39	.12
GFI	> .90	.79	.89	.74	.79	.89	.73	.80
CFI	> .90	.58	.75	.50	.57	.74	.48	.19
TLI	> .90	.54	.69	.44	.53	.70	.43	.12
AGFI	> .90	.75	.84	.69	.74	.84	.69	.77
Reg Wts*	≥ .70	50%	82%	54%	54%	92%	54%	54%

Note: Reg wts (Regression Weights). Each column represents the percent of factor loadings that showed strong factor loadings with values ≥ .70.
** $p < .001$

Table 9. Correlations of the APPM and NEO-IPIP with Explicit and Implicit Criteria

	GPA	DVTime	DVRemind
NEO - Composite	.30**	-.12	-.11
APPM – Composite	.26**	-.15*	-.17*

Note: **DVTime:** Implicit Criteria 1 (Difference in time); **DVRemind:** Implicit Criteria 2 (# of reminders)
** Significant $p < .01$; *Significant $p < .05$
N = 155-175

There was a stronger relationship between the NEO-IPIP and the deliberate criteria (r = .30) than the first implicit criteria (duration of time between completion of the second study and the first study) (r = -.12) (z = 3.82, p < .00), and also for the second implicit criteria (number of reminders) (r = -.11) (z = 3.74, p < .00). These results support Hypothesis 3a.

Fisher r-to-z calculation was also used to test whether the APPM demonstrates a stronger correlation with the spontaneous implicit criteria than the deliberate criteria (*Hypothesis 3b*). Results showed there was a stronger relationship between the APPM and the deliberate criteria (r = .26) than the first implicit criteria (duration of time between completion of the second study and the first study) (r = -.15) (z = 3.73, p < .00), and also for the second implicit criteria (number of reminders) (r = -.17) (z = 3.92, p < .00). These results were opposite to what would be expected in accordance with the double dissociation effect, hence Hypothesis 3b was not supported.

Hierarchical regressions were computed to test whether the double dissociative pattern emerges from the APPM and the NEO-IPIP scales on implicit and deliberate criteria. First, incremental prediction of the NEO-IPIP over the APPM and cognitive ability (SAT score) for GPA was first examined, with cognitive ability and SAT entered in step 1 and NEO-IPIP scores were entered in Step 2 with GPA as the criteria. All variables were significantly predictive of GPA (SAT (β = .24, p < .01); APPM (β = .30, p< .01); NEO-IPIP (β = .20, p< .05). Addition of NEO after cognitive ability and the APPM was entered in the model, did significantly improve prediction (ΔR^2 = .03, $\Delta F(1, 116)$ = 4.80, p = .03), which supports hypothesis 4a. (See Table 10.)

To test Hypothesis 4b, SAT and NEO-IPIP was entered in step 1and APPM was entered in Step 2 for both of the spontaneous implicit criteria in two separate models (see Table 11). The APPM provided incremental prediction beyond what SAT and the NEO-IPIP for number of reminders as the criteria (ΔR^2 = .04, ΔF *(1, 115)* = 4.89, p = .03) but not number of reminders (ΔR^2 = .01, ΔF *(1, 114)* = 1.86, p = .06). Thus, Hypothesis 4b was supported. Consistent with expectations, the NEO did not provide incremental prediction once SAT and APPM was taken into account, however the APPM did provide incremental prediction for the deliberate criteria (ΔR^2 = .05, $\Delta F(1, 116)$ = 6.37, p = .01).

To test whether the APPM was more resistant than the NEO-IPIP to faking conditions, a paired samples t-test was run for each overall composite between faking and honest conditions (see Table 12). Results showed that there was a significant difference in scores between faking (M = 3.87, SD = .58) and honest (M = 3.63, SD = .47) conditions for the NEO-IPIP with an effect size difference of d = .45 (t(176) = 6.06, p < .00) (See Table 12). However, no mean differences were found for the APPM between the faking (M = 3.62, SD = .32) and honest (M = 3.59, SD = .38) conditions; t(176) = 1.18, p = .24[4]. These results provide support for Hypothesis 5.

[4] The NEO-IPIP was on a 5 point while the APPM was on a 6 point scale.

Table 10. Summary of Hierarchical Regression Analysis for NEO-IPIP with the deliberate and implicit criteria

	Step 1			Step 2			
Variable	Model 1	SAT	APPM	Model 2	SAT	APPM	NEO
DV: GPA							
β	-	**.26***	**.29****	-	**.26****	**.22****	**.21***
ΔF	**9.64****	-	-	**5.53****	-	-	-
ΔR^2	.14	-	-	.03	-	-	-
R	.14	-	-	.17	-	-	-
DV: Time Difference							
β	-	.00	-.16	-	.00	-.19	10
ΔF	1.46	-	-	1.00	-	-	-
ΔR^2	.03	-	-	.01	-	-	-
R	.03	-	-	.03	-	-	-
DV: Reminder Count							
β	-	.02	-.17	-	.02	-.21*	.12
ΔF	1.86	-	-	1.63	-	-	-
ΔR^2	.03	-	-	.01	-	-	-
R	.03	-	-	.05	-	-	-

Note: **SAT:** Scholastic Aptitude Test; **APPM:** Anchored Preferential Personality Measure; **NEO:** Neuroticism-Extraversion-Openness Inventory (Includes Conscientiousness & Agreeableness; **GPA:** Cumulative Grade Point Average; **Time Difference:** Time difference from starting first survey to the completion of the second survey; **Reminder Count:** Number of total reminders sent before second survey was complete.

*$p < .05$. **$p < .01$

Table 11. Summary of Hierarchical Regression Analysis for APPM with the Deliberate and Implicit Criteria

	Step 1			Step 2			
Variable	Model 1	SAT	NEO	Model 2	SAT	NEO	APPM
DV: GPA							
β	-	**.24***	**.28***	-	**.26****	**.21****	**.22***
ΔF	**9.26***	-	-	**6.22***	-	-	-
ΔR^2	.13	-	-	.04	-	-	-
R	.36	-	-	.42	-	-	-
DV: Time Difference							
β	-	.03	.04	-	.01	.10	-.19[5]
ΔF	.16	-	-	3.58	-	-	-
ΔR^2	.00	-	-	.03	-	-	-
R	.05	-	-	.18	-	-	-
DV: Reminder Count							
β	-	.04	.06	-	.02	.12	**-.21***
ΔF	.26	-	-	**4.84***	-	-	-
ΔR^2	.00	-	-	**.04***	-	-	-
R	.00	-	-	.05	-	-	-

Note: **SAT:** Scholastic Aptitude Test; **APPM:** Anchored Preferential Personality Measure; **NEO:** Neuroticism-Extraversion-Openness Inventory (Includes Conscientiousness & Agreeableness; **GPA:** Cumulative Grade Point Average; **Time Difference:** Time difference from starting first survey to the completion of the second survey; **Reminder Count:** Number of total reminders sent before second survey was complete.

*$p < .05$. **$p < .01$

[5] This item was close but non-significant, p = .06

Table 12. Paired Sample T-Tests between Faking and Honest Conditions for the NEO-IPIP and the APPM.

	Fake		Honest		95% CI					
Variable	Mean	SD	Mean	SD	Lower	Upper	t	df	Sig. (2-tailed)	Cohen's *d*
NEO Faking – Honest	3.87	.58	3.63	.47	0.16	0.32	6.06	176	.00	.45
APPM Faking – Honest	3.62	.32	3.59	.38	-0.02	0.08	1.18	176	.24	.09

Chapter 5: Discussion

Summary of Main Findings

The primary goals of this research study were to examine the construct validity of the Anchored Preferential Personality Measure (APPM) of conscientiousness, and to examine relationships with theoretically linked constructs as measured by the NEO-IPIP. In addition, analyses were conducted to explore prediction of the APPM with a deliberate criterion (GPA) and two spontaneous implicit criteria – the duration in time from when the first survey was taken to when the second survey was complete (TimeDiff), and the number of reminders that the participants were given before they completed the study (Reminders). These analyses were conducted to examine the predictive utility of the assessment as a selection tool as well as the predictive patterns that help evaluate the "implicitness" of the measure. It was hypothesized that the APPM should demonstrate the double dissociation pattern as outlined by Asendorpf et al., (2009), where the APPM should be more predictive of the spontaneous implicit criteria (TimeDiff and Reminders) while the NEO-IPIP should be more predictive of the deliberate outcome (GPA). Lastly, the extent to which the APPM was resistant to proximal and distal goals in comparison to NEO-IPIP was also examined. In this study this goal was the motivation to fake or to respond honestly.

Construct Validity. There were a total of 90 items developed in the original APPM, with 15 items specifically targeted to capture one of six facets that comprise the NEO-IPIP conscientiousness. However, 60 were dropped due to insufficient loadings in the item analysis resulting in the following scales: Achievement (7 items $\alpha = .52$), Cautiousness (6 items, $\alpha = .52$), Orderliness (5 items, $\alpha = .45$), and Dutifulness (10 items

$\alpha = .50$). There was partial support for Hypothesis 1, in that the Achievement and Orderliness APPM scales demonstrated a pattern of larger correlations with their respective NEO-IPIP facets. However, the APPM Dutifulness and Cautiousness scales demonstrated less refined and interpretable relationships with the NEO-IPIP. Hypothesis 2 was not supported, as the confirmatory factor analysis failed to support the factor structure of the APPM. This is not particularly surprising given the low reliabilities of each factor. Thus, while the facet level of what the APPM is measuring is unclear, there is some indication that the measurement as a whole is assessing some aspects of conscientiousness given the composite measure of the APPM significantly correlated with the NEO-IPIP composite measure ($r = .37$, $p < .001$).

Normative Analysis of Indirect Measures of Implicit Personality

In light of the double dissociation model as proposed by Asendorpf et al., (2002), it was hypothesized that the NEO would demonstrate stronger correlations with the deliberate criteria (GPA) than with the spontaneous implicit criteria (time difference and number of reminders) (hypothesis 3a) while the reverse would be true for the APPM (hypothesis 3b). While there was support for hypothesis 3a, hypothesis 3b was not supported as the APPM demonstrated stronger correlations with GPA than with the implicit criteria.

In addition to comparing bivariate correlations, the double dissociation model was also examined through hierarchical regressions. Full support for the double dissociation effect was found where the NEO-IPIP incrementally predicted beyond the APPM and cognitive ability for the deliberate criteria (GPA) (Hypothesis 4a), and the APPM

provided incremental validity over the NEO-IPIP and cognitive ability for number of reminders (Hypothesis 4b).

A finding that was not predicted was the incremental validity the APPM beyond that of the NEO-IPIP and cognitive ability when predicting GPA. GPA was set as the deliberate criteria, as it is understood that it is largely determined by controlled behaviors such as studying regularly, completing assignments on time, attending class regularly, and seeking out additional help and feedback when needed (Noftle et al., 2007). One possible explanation for this finding is that implicit processes may be playing a bigger role in GPA than expected.

Vianello, Robusto, and Anselmi (2010) found that indirect measures of implicit conscientiousness incrementally predicted beyond a self-report measure on a self-report of average exam grade and also the number of exams successfully passed in the previous year. What the findings of this research and those by Vianello et al. (2010) suggest is that while academic performance requires deliberate behaviors, poor academic performance may also be a result of unintended spontaneous behaviors. In most cases, students do not set out to fail a test, sleep through class, forget homework assignments, or miss deadlines. Thus, it is possible that low GPA students have the sincere deliberate intentions to perform well academically, but misaligned implicit drives to achieve a desirable end. This perspective is consistent with Bing et al.'s (2007a) integrative model of implicit and explicit cognitions, whereby prediction is maximized when the explicit and implicit self are congruent when implicit motives are channeled through aligned explicit traits.

The deliberate/spontaneous heterogeneity of GPA may explain the equivocal support for the double dissociation effect in this study where the APPM failed to demonstrate higher correlations with the spontaneous than the deliberate criteria (hypothesis 3b not supported) but had incrementally predicted beyond that of the NEO for the number of reminders (Hypothesis 4b supported). If GPA is an outcome influenced by a combination of deliberate and spontaneous behaviors, the APPM may have had stronger relationships with GPA because it could be directly and indirectly (as channeled through explicit cognitions) capturing features of the criterion space. In addition to the deliberate/spontaneous heterogeneity of GPA, another explanation to these findings is that the APPM may not be a pure indirect measure of implicit Conscientiousness, but rather a hybrid measure of implicit and explicit traits. The implicit/explicit nature of the APPM alignment in combination with the deliberate/spontaneous aspect of GPA could explain why the correlations between the APPM and GPA were so much higher than expected and why there was unexpectedly incremental prediction of the APPM for GPA over the NEO.

In sum, support for Hypothesis 3a, 3b, 4a, and 4b would have provided robust support in favor of the double dissociation effect. Results however were mixed in nature. Given the incremental prediction of the spontaneous outcome over the NEO, there is some evidence of some implicit aspect of the APPM. However, the lack of support for hypothesis 3b, and the unexpected findings of the APPM incrementally predicting GPA over NEO, opens up questions on the purity of the APPM in capturing implicit personality, and the degree to which GPA is deliberate or spontaneous.

The "what" criterion is one of the three major considerations in the examination of measure implicitness as outlined by De Houwer et al., (2009). A second consideration investigated in the study is the "implicitness" criteria which "*examines the conditions in which the attribute automatically causes variation in the measurement outcome.*" A test of this was accomplished by examining if variation in scores could be manipulated by instructions to answer honestly or to respond in a manner that they think would get them a job. While there were significant mean differences in scores from the faking and honest conditions for the NEO-IPIP, there were no significant differences found for the APPM, which provides support for Hypothesis 5.

As a practitioner these results are promising as there is evidence that the APPM has the potential of reducing the effects of faking and provides incremental prediction in both the deliberate and spontaneous outcome after explicit self-report measures and cognitive ability is taken into account. Currently, the most common non-cognitive measure used for personnel selection purposes designed to mitigate faking is the forced-choice format. A drawback to using the forced-choice measure is the challenges in making inter-individual comparisons, given the ipsative design of this measure. Research has shown that IRT is a viable means to circumvent the ipsativity challenges of forced-choice assessments (Brown & Maydeu-Olivares, 2013), however sample size requirements are demanding to run these analyses with a recommended sample size of 500 for accurate parameter estimates (Reeve & Fayers, 2005). For some organizations and occupations this sample size requirement can be met, but for others, this isn't a possibility, and the APPM has the potential of providing an alternative option in the case of faking concerns.

Limitations and Directions for Future Research

The most notable limitation of this research is the lack of reliability of the APPM. The reliability of the composite APPM measure ($\alpha = .64$) was far below the general rule of thumb for selection measured used in high stakes applied settings ($\alpha >= .80$). Future research needs to be conducted to cross validate these results on another sample to ensure generalizability of these findings, given such low reliabilities. One of the issues in creating the APPM items is that often times the anchors developed were over indexed in terms of social desirability or undesirability, where the high conscientiousness anchor was too negative and the low conscientiousness anchor was too positive. There were several items that had to be removed due to significant but opposite prediction. These items were not kept in the pool intentionally as it was not clear what exactly these items were measuring.

When items were initially developed the items were constructed so that the positive anchor was spun in a more negative manner and the negative anchor in a more positive way. The intention was that both anchors would possess some grain of truth where the respondent, depending on their level of the trait in question, would gravitate to one anchor more so than the other. However, there was no specific framework or formula in constructing each anchor for each item. Using the results from this study, the content of the items that were retained vs. those that were dropped was examined. Trends were identified that could serve as future test specifications are shown in Table 13.

Table 13. APPM Item Writing Framework

Test Specification	APPM Item Example	Effective/Ineffective
Avoid using absolutes	**High:** "Managers should punish all infractions." **Low:** "Managers should use discretion in punishing infractions	Ineffective
Simplify verbiage	**High:** "People often hurt themselves by taking on responsibilities that are not their own" **Low:** "People do what is necessary by taking on responsibilities that are not their own."	Ineffective
Avoid statements in the negative form	**High:** "It is not that difficult to stick to a regular workout routine" **Low:** "It is challenging to stick to a regular workout routine"	Ineffective
Use equally socially desirable antonyms	**High:** "Intense people are higher performers" **Low:** "Laid back people are higher performers"	Effective
Use paired statements in the affirmative form	**High:** "It's is important to accept and learn from mistakes" **Low:** "It is important to avoid and prevent making mistakes"	Effective
Use commonly used justifications for behavior	**High:** "One can rejuvenate themselves if they take time off from their tasks" **Low:** "One can lose track of their progress if they take time off from their tasks	Effective

In content analyzing the effective vs. ineffective items, items that were both worded in the affirmative and present opposing concepts (e.g., It is good to X vs It is good to Y) were more effective than having one anchor in the affirmative and the other in the negative form (It is good to do X vs It is not good to do X). In addition, items that were shorter, less complex in structure also surfaced as more effective or more likely to be retained. These nuanced differences in anchored statements may have resulted in a measure that captured both implicit and explicit personality, with items having the affirmative/negative form as more explicit and the affirmative/opposing concepts may have been more implicit. This may have contributed to the earlier results of higher correlations with GPA. With these specifications in mind, future research should develop more items in the effort of enhancing the implicitness of this measure, improve reliability, and improve construct validity of this method.

The second limitation of this research is the nature of the sample. Sixty-three percent (n = 126) of the participants in the study were female, with 31% (n= 60) male and 6% (n = 13) unspecified. Research has shown that males tend to underreport their level of conscientiousness on explicit measures but not implicit measures (Vianello, Schnabel, Sriram, & Nosek, 2013). For this sample, while the average NEO conscientiousness measure appeared larger for females than males ($\bar{X}$ = 3.65 vs. 3.55) there were no significant differences in gender for the NEO ($t(184) = 1.26$, $p > .05$) or for the APPM ($t(176) = 6.06$, $p = .00$). Although there were no significant differences between males and females, future research should further explore the incremental validity of the APPM over the NEO for various outcomes, and whether gender moderates that relationship.

A third limitation is the performance criteria used in this study. Contrary to expectations, the APPM demonstrated a stronger correlation with GPA than the spontaneous criteria. The reason for this could be due to the implicit/explicit nature of the APPM, but also that GPA has more implicit spontaneous elements than what was initially thought. Future research should explore and utilize a variety of criteria more deliberate in nature to provide greater support of the double dissociation effect and to investigate situational moderators of this effect. For many occupations implicit personality may not provide incremental prediction of job performance as many jobs are highly structured and under the deliberate control of the individual. However, as conditions that elicit schematically driven behavior (e.g., law enforcement positions, high stress environments) increase on the job, the more indirect measures of implicit personality would be expected to predict job performance. Further research should be conducted to examine which industries, job families, or specific tasks/major duties align with what would be classified as an "implicit criteria." This would not only help provide a framework of what criteria should be used in future research validating implicit measures, but also shed light on the types of job in which implicit assessments would be of most value for personnel selection purposes.

A third limitation is that while the APPM seemed to be resistant to faking, the criterion related validities of the APPM for both the deliberate and implicit criteria were attenuated in the faking condition. Faking was primarily used in this study as a tool to test the hypothesis that scores would be invariant across goals to modify scores. Future research should be conducted to replicate these findings and to further investigate the reasons behind the decrements of validity but lack of significant mean difference in

scores across honest and faking conditions. Additionally, further research should explore the extent to which test familiarity with the APPM moderates the extent to which the APPM is fakable. Research has shown that the IAT is easier to fake with increasing familiarity (Steffens, 2004), so it is possible similar results would surface for the APPM. If the APPM is truly capturing implicit biases, examinees may be able to inflate scores with practice or coaching, but the degree of inflation should be less relatively to explicit self-report measures.

Conclusion

The current study developed and validated a new measure of implicit personality and demonstrated some evidence of the "implicitness" of the APPM. While this measure shows promise, additional research needs to be conducted to further refine the APPM and improve the reliability and construct validity of this measure. Once these additional research efforts are made, this measure has the potential of not only improving the prediction of job performance, but to be of practical value for applied personnel selection psychologists given the minimal time requirements for applicants to take, fewer resources to deliver, ease in scoring, and flexibility in delivery method.

References

Alicke. M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49*, 1621–1630.

Anderson, C. D., Warner, J. L., & Spencer, C. C. (1984). Inflation bias in self-assessment examinations: Implications for valid employee selection. *Journal of Applied Psychology, 69*, 574 –580.

Arthur, W. r., Day, E., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 56(1), 125-154.

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal Of Personality And Social Psychology*, *83*(2), 380-393.

Ayduk, O., Mischel, W., & Downey, G. (2002). Attentional mechanisms linking rejection to hostile reactivity: The role of 'hot' versus 'cool' focus. *Psychological Science, 13*(5), 443-448.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology, 3*, 77– 85.

Bartlett, M. S. (1954). A further note on the multiplying factors for various χ^2 approximations in factor analysis. *Journal of the Royal Statistical Society, 16,* 296–298.

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.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next?. *International Journal Of Selection And Assessment, 9*(1-2), 9-30.

Berry, C. M., Page, R. C., & Sackett, P. R. (2007). Effects of Self-Deceptive Enhancement on Personality-Job Performance Relationships. International *Journal of Selection And Assessment, 15*(1), 94-109.

Bing, M.N., LeBreton, J.M., Davison, H.K., Davison, H.K., Migetz, D.Z., James, L.R. (2007a). Integrating implicit and explicit social cognitions for enhanced personality assessment. *Organizational Research Methods, 10* (1), 136- 179.

Bing, M., Stewart, S., Davison, H., Green, P., McIntyre, M., & James, L. (2007b). An integrative typology of personality assessment for aggression: Implications for predicting counterproductive workplace behavior. *Journal of Applied Psychology, 92*(3), 722-744.

Bollen, K.A. & Long, J.S. (eds). (1993). *Testing Structural Equation Models*. Newbury Park, CA: Sage.

Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition, 38*(2), 186-196.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel Selection in Organizations* (pp. 71-98). San Francisco: Jossey-Bass.

Bowler, J. L., Bowler, M. C., & Cope, J. G. (2013). Measurement issues associated with conditional reasoning tests: An examination of faking. *Personality and Individual Differences, 55*(5), 459-464.

Brown, A., & Maydeu-Olivares, A. (2013). How IRT can solve problems of ipsative data in forced-choice questionnaires. *Psychological Methods, 18*(1), 36-52.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York, NY: Guilford Press.

Burns, L. J. & Knussen, C. (2005). False consensus and accuracy of perceptions of nurses regarding universal precaution practices. *Psychology, Health & Medicine, 10(4)*, 344-354.

Burroughs, S.M., James, L.R., (2005). Advancing the Assesssment of Dispositional Aggressiveness through conditional reasoning. In Fox, S., & Spector, P.E. (Eds.), *Counterproductive work behavior: Investigations of actors and targets* (pp. 127-150). Washington DC: American Psychological Association.

Butler, R. P., Lardent, C. L., & Miner, J. B. (1983). A motivational basis for turnover in military officer education and training. Journal of Applied Psychology, 68(3), 496-506.

Caldwell-Andrews, A., Baer, R. A., & Berry, D. R. (2000). Effects of response sets on NEO-PI-R scores and their relations to external criteria. *Journal of Personality Assessment, 74*(3), 472-488.

Carpenter, S. L. (1988). Self-relevance and goal-directed processing in the recall and weighting of information about others. Journal of Experimental Social Psychology, 24, 310-332.

Cervone, D. (2004). The Architecture of Personality. *Psychological Review, 111*(1), 183-204.

Chan, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement, 18*, 205-215.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224.

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*(1), 83-117.

Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance, 18*, 267–307.

Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47,* 847– 860.

Chiaburu, D. S., Oh, I., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal Of Applied Psychology*, 96(6), 1140-1166.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Collins, C. J., Hanges, P. J., & Locke, E. A. (2004). The Relationship of Achievement Motivation to Entrepreneurial Behavior: A Meta-Analysis. *Human Performance, 17*(1), 95-117.

Cook, S. W., & Selltiz, C. (1964). A multiple-indicator approach to attitude measurement. *Psychological Bulletin, 62*(1), 36-55.

Corr, P. J., & Gray, J. A. (1995). Relationship between attributional style and Lie scores in an occupational sample motivated to fake good. Personality and Individual Differences, 18(3), 433-435.

Corr, P. J., & Gray, J. A. (1996). Attributional style as a personality factor in insurance sales performance in the UK. *Journal of Occupational and Organizational Psychology, 69*(1), 83-87.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*(2), 163-170.

Dasgupta, N. (2010). Implicit measures of social cognition: Common themes and unresolved questions. *Journal of Psychology, 218*(1), 54-57.

De Houwer, J., Beckers, T., & Moors, A. (2007). Novel attitudes can be faked on the Implicit Association Test. *Journal of Experimental Social Psychology, 43*(6), 972-978.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit

measures: A normative analysis and review. *Psychological Bulletin, 135*(3), 347-368.

Deutsch, R., & Strack, F. (2006). Duality Models in Social Psychology: Response to Commentaries. *Psychological Inquiry, 17*(3), 265-268.

Dislich, F. R., Zinkernagel, A., Ortner, T. M., & Schmitt, M. (2010). Convergence of direct, indirect, and objective risk-taking measures in gambling: The moderating role of impulsiveness and self-control. *Journal of Psychology, 218*(1), 20-27.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82*, 62–68.

Dulsky, S. G., & Krout, M. H. (1950). Predicting promotion potential on the basis of psychological tests. *Personnel Psychology*, 3345-351.

Dunning, D. & Hayes, A. F. (1994). Evidence for egocentric comparison in social judgment. *Journal of Personality and Social Psychology, 71*, 213-229.

Dunning, D., Meyerowitz, J., & Holzberg, A. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, 57(6), 1082-1090.

Eastwick, P. W., Eagly, A. H., Finkel, E. J., & Johnson, S. E. (2011). Implicit and explicit preferences for physical attractiveness in a romantic partner: A double dissociation in predictive validity. *Journal of Personality And Social Psychology, 101*(5), 993-1011.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83*(6), 1441-1455.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage and coding. *Semiotica, 1*, 49–98.

Emmons, R.A. (1989). Exploring the relations between motives and traits: The case of narcissism. In D. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (p. 32-44). New York, NY: Springer.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*, 709-724.

Epstein, S. (2003). Cognitive-experiential self-theory of personality. In T. Millon, M. J. Lerner (Eds.) , *Handbook of psychology: Personality and social psychology, Vol. 5* (pp. 159-184). Hoboken, NJ, US: John Wiley & Sons Inc.

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive–experiential self-theory. In S. Chaiken, Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 462-482). New York, NY, US: Guilford Press.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013-1027.

Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In C. Hendrick, M. S. Clark, C. Hendrick, M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 74-97). Thousand Oaks, CA US: Sage Publications, Inc.

Feist, G.K. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2* (4), 290-309.

Filipkowski, J.N. (2010). Measuring Conscientiousness with Explicit and Implicit Measures. *Master's Thesis.*

Frost, B. C., Ko, C., & James, L. R. (2007). Implicit and explicit personality: A test of a channeling hypothesis for aggressive behavior. *Journal of Applied Psychology, 92*(5), 1299-1319.

Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology*, 49(3), 171-180.

Geisinger, K. W. (1960). The use of the sentence completion technique for the selection of salesmen: A methodological and validation study. *Dissertation Abstracts, 21*, 1630.

Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling, 3*, 62–72.

Giladi, E. E., & Klar, Y. (2002). When standards are wide of the mark: Nonselective superiority and inferiority biases in comparative judgments of objects and concepts. *Journal of Experimental Psychology: General, 131*, 538–551.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*, 509–517.

Goldberg, D. (1993). The structure of phenotypic personality traits. *American Psychologist* (48), 26-34.

Goffin, R.D. & Christiansen, N.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,* 340-344.

Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4-27. Greenwald, Anthony, G. "Revised top ten list of things wrong with the IAT". SPSP Autstin, Texas, January 29,2004. Attitudes Conference.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality And Social Psychology, 74*(6), 1464-1480.

Greenwald, A.G. (2001. October). *What's wrong with the IAT?* Paper presented at the meeting of the Society of Experimental Social Psychology, spokae, WA.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*(1), 3-25.

Greenwald, A. G., Poehlman, T., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality And Social Psychology, 97*(1), 17-41.

Griffith, R.L.; Chmielowski, T, Yukiko, Y., (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review, 36*(3), 341-355.

Grumm, M., & von Collani, G. (2007). Measuring Big-Five personality dimensions with the implicit association test--Implicit personality traits or self-esteem?. *Personality and Individual Differences, 43*(8), 2205-2217.

Harms, P.D., Luthans, F. (2012). Measuring implicit psychological construct sin organizational bheaivor: An example using psychological capital. *Journal of Organizational Behavior 33*(4), 589-594.

Hicks, L. E. (1970). Some properties of ipsative, normative, and forced-choice normative measures. *Psychological Bulletin, 74*(3), 167-184.

Hoorens, V. (1995). Self-favoring biases, self-presentation, and the self-other asymmetry in social comparison. *Journal of Personality, 63*(4), 793-817.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Personality and Social Psychology Bulletin, 31*(10), 1369-1385.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581–595.

Hough, L. M., Oswald, F. L., & Ployhart, R. 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.

Hough, L., & Paullin, C. (1994). Construct-oriented scale construction: The rational approach. In G. S. Stokes, M. D. Mumford, W. A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 109-145). Palo Alto, CA, US: CPP Books.

Hovland, C.I., & Sherif, M. (1952). Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology, 47,* 822-832.

Huang, J. L., Ryan, A., Zabel, K. L., & Palmer, A. (2014). Personality and adaptive performance at work: A meta-analytic investigation. *Journal if Applied Psychology, 99*(1), 162-179.

James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods, 1*(2), 131-163.

James, L. R., & LeBreton, J. M. (2012). Conditional reasoning and the implicit personality: Concepts and theoretical foundations. In , Assessing the implicit personality through conditional reasoning (pp. 17-68). Washington, DC US: American Psychological Association.

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., & ... Williams, L. J. (2005). A Conditional Reasoning Measure for Aggression. *Organizational Research Methods, 8*(1), 69-99.

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from Web-based personality inventories. *Journal of Research in Personality, 39*(1), 103-129.

Johnson, R. E., & Saboe, K. N. (2011). Measuring implicit traits in organizational research: Development of an indirect measure of employee implicit self-concept. *Organizational Research Methods, 14*(3), 530-547.

Johnson, R. E., & Steinman, L. (2009). Use of implicit measures for organizational research: An empirical example. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 41*(4), 202-212.

Johnson, R.E., Tolentino, A.L., Rodopman, O.B., Cho, E. (2010). We (sometimes) know not how we fell: predicting job performance with an implicit measure of trait affectivity. *Personnel Psychology, 63,* 197-219.

Jones, E. E., & Kohler, R. (1958). The effects of plausibility on the learning of controversial statements. *The Journal of Abnormal and Social Psychology, 57*(3), 315-320.

Judge, T. A., & LePine, J. A. (2007). The bright and dark sides of personality: Implications for personnel selection in individual and team contexts. In J. Langan-Fox, C. L. Cooper, R. J. Klimoski, J. Langan-Fox, C. L. Cooper, R. J. Klimoski (Eds.), *Research companion to the dysfunctional workplace: Management challenges and symptoms* (pp. 332-355).

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology, 98*(6), 875-925.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31–36.

Kauffman, J.R., Jex, S.M., Love, K.G. and Libkuman, T.M. (1993). The construct validity of assessment center performance dimensions. *International Journal of Selection and Assessment, 1,* 213-223.

Kline, P. (1994). An easy guide to factor analysis. New York, NY: Routledge.

Komar, S., Brown, D.J., Komar, J.A. (2008). Faking and the validity of conscientiousness: A Monte Carlo investigation. *Journal of Applied Psychology, 93* (1), 140-154.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77,* 221–232.

Kruger, J.I. & Acevedo, M. (2005). Social projection and the psychology of choice. In M.D. Alicke, D.A. Dunning, & J. I. Krueger (Eds.), *The Self in Social Judgment* (pp. 17 – 41). New York, NY: Taylor & Francis Group.

Kruger, J., & Burrus, J. (2004). Egocentrism and focalism in unrealistic optimism (and pessimism). *Journal of Experimental Social Psychology, 40,* 332–340.

Kuiper, N. A.. & MeDonald, M. R. (1982). Self and other perception in mild depressives. *Social Cognition.* 1, 223-239.

Lambert, A. J., & Wedell, D. H. (1991). The self and social judgment: Effects of affective reaction and "own position" on judgments of unambiguous and ambiguous information about others. *Journal of Personality and Social Psychology, 61,* 884-898.

Landis, R.S., Beal, D., & Tesluk, P.E. (2000). A comparison of approaches to forming composite measures in structural equation models. *Organizational Research Methods, 3,* 186-207.

Latham, G. P., Stajkovic, A. D., & Locke, E. A. (2010). The relevance and viability of subconscious goals in the workplace. *Journal of Management, 36*(1), 234-255.

Lang, J. B., Zettler, I., Ewen, C., & Hülsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology, 97*(6), 1201-1217.

Lanyon, R. I., & Goodstein, L. (1997). Personality assessment (3rd ed.). Oxford, England: John Wiley & Sons.

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 46*(8), 819-834.

Lebreton, J.M., Barksdale, C.D., & Robin, J. (2007). Measurement issues associated with conditional reasoning tests: Indirect measurement and test faking. *Journal of Applied Psychology, 92* (1), 1-16.

LePine, J.A., Colquitt, J.A., Erez, A. (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology, 53*, 563 – 593.

Lunneborg, P. W., & Olch, D. (1970). Sentence completion correlates of airline pilot attitude and proficiency. *Journal of Projective Techniques & Personality Assessment, 34*(6), 497-502.

Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items". *Personnel Psychology, 44*, 763-927.

Marata-Ramierz, M., & Holgado-Tello, F.P. (2013). Construct Validity of Likert Scales through Confirmatory Factor Analysis: A Simulation Study Comparing Different Methods of Estimation Based on Pearson and Polychoric Correlations. *International Journal of Social Science Studies, 1*(1), 54-61.

Martocchio, J.J., & Judge, T.A. (1997). Relationship between conscientiousness and learning in employee training: Mediating influences of self-deception and self-efficacy. *Journal of Applied Psychology, 82*, 764 – 773.

McClelland, D. C., & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied Psychology*, *67*(6), 737-743.

McConnel, A.R., & Leibold, J.M. (2000). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435-442.

McDaniel, M. J., Beier, M. E., Perkins, A. W., Goggin, S., & Frankel, B. (2009). An assessment of the fakeability of self-report and implicit personality measures. *Journal Of Research In Personality*, *43*(4), 682-685.

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, *9*(1-2), 103-113.

McFarland, L.A., & Ryan, A.M. (2000).Variance in faking across non-cognitive measures. *Journal of Applied Psychology*, *85,* 812–821.

McFarland, L. A., Ryan, A., & Ellis, A. (2002). Item placement on a personality measure: Effects on faking behavior and test measurement properties. *Journal of Personality Assessment, 78*(2), 348-369.

McClelland, D.C. (1985). How motives, skills, and values determine what people do. *American Psychologist, 40,* 812-825.

Meade, A. W. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational & Organizational Psychology, 77*(4), 531-551.

Milgram, S., Mann, L., & Harter, S. (1965).The lost-letter technique: A tool of social research. *The Public Opinion Quarterly, 29*(3), 437-438.

Miner, J. B., Smith, N. R., & Bracker, J. S. (1994). Role of entrepreneurial task motivation in the growth of technologically innovative firms: Interpretations from follow-up data. Journal of Applied Psychology, 79(4), 627-630.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*(1), 3-19.

Mischel, W., & Ayduk, O. (2002). Self-regulation in a cognitive-affective personality system: Attentional control in the service of the self. *Self and Identity, 1*(2), 113-120.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*(2), 246-268.

Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review Of Psychology, 49,*229-258.

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

Motowidlo, S. J., & Borman, W. C. (2000). Contextual Performance and Organizational Citizenship Behavior in Human Management. *Human Resource Management Review, 10*(1), 1-2.

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006a). Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology, 91*(4), 749-761.

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006b). A theoretical basis for situational judgment tests. In J.A. Weekley & R.E. Ployhart (Eds.), *Situational judgment tests.* Mahwah, NJ: Erlbaum.

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology, 79*(4), 475-480.

Mount, M. K., Witt, L. A., & Barrick, M. R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology, 53*(2), 299-323.

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

Mukherjee, B. N. (1968). Achievement values, social desirability, and endorsement of trait names on the Berdie Check list. Journal of Applied Psychology, 52(2), 127-132.

Murphy, K. R., & Dzieweczynski, J. L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18,* 343–358.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist, 61,* 622– 631.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality And Social Psychology, 93*(1), 116-130.

Norris, P., & Epstein, S. (2011). An experiential thinking style: Its facets and relations with objective and subjective criterion measures. *Journal of Personality, 79*(5), 1043-1079.

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go Association Task. *Social Cognition, 19*(6), 625-666.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York: Psychology Press

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments & Computers, 32(3)*, 396-402.

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

Organ, D. W. (1988). A restatement of the satisfaction-performance hypothesis. *Journal of Management,* 14, 547–557.

Paulhus, D.L. (1984) Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*(2), 307-317.

Peeters, H., & Lievens, F. (2005). Situational Judgment Tests and Their Predictiveness of College Students' Success: The Influence of Faking. *Educational and Psychological Measurement, 65*(1), 70-89.

Perugini, M., Conner, M., & O'Gorman, R. (2011). Automatic activation of individual differences: A test of the gatekeeper model in the domain of spontaneous helping. *European Journal of Personality, 25*(6), 465-476.

Phelan, J. G. (1962). Projective techniques in the selection of management personnel. *Journal of Projective Techniques, 26*(1), 102-104.

Proudfoot, J. G., Corr, P. J., Guest, D. E., & Gray, J. A. (2001). The development and evaluation of a scale to measure occupational attributional style in the financial services sector. *Personality and Individual Differences*, 30(2), 259-270.

Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Hedge, J. W., & Borman, W. C. (2002). Predicting adaptive performance: Further tests of a model of adaptability. *Human Performance*, *15*(4), 299-324.

Redmore, C. (1976). Susceptibility to faking of a sentence completion test of ego development. *Journal of Personality Assessment, 40*(6), 607-616.

Reeve, B. B., & Fayers, P. (2005). Applying item response theory modeling for

evaluating questionnaire item and scale properties. In P. Fayers & R. D. Hays (Eds.), Assessing quality of life in clinical trials: *Methods of practice* (2nd ed). New York: Oxford University Press, 2005. p. 55–73.

Röhner, J., Schröder-Abé, M., & Schütz, A. (2013). What do fakers actually do to fake the IAT? An investigation of faking strategies under different faking conditions. *Journal of Research In Personality, 47*(4), 330-338.

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

Rudolph, A., Schröder-Abé, M., Riketta, M., & Schütz, A. (2010). Easier when done than said!: Implicit self-esteem predicts observed or spontaneous behavior, but not self-reported or controlled behavior. Journal of Psychology, 218(1), 12-19.

Russell, C.J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too coarse for comfort. *Journal of Applied Psychology, 77,* 336-342.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172.

Schmitt, N., Clause, C. S., &Pulakos, E. D. (1996). Subgroup differences associated with different measures of some common job relevant constructs. In C.L. Cooper & I.T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology, 11*, pp.115–139). New York: Wiley.

Schmidt FL, Hunter JE. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262–274.

Schmidt, F. L., & Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86,* 162–173.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29,* 304–321.

Schnabel, K., Asendorpf, J. B., & Greenwald, A. G. (2008). Assessment of individual differences in implicit cognition: A review of IAT measures. *European Journal of Psychological Assessment*, 24(4), 210-217.

Schulman, P., Seligman, M. E., & Amsterdam, D. (1987). The Attributional Style Questionnaire is not transparent. *Behaviour Research and Therapy,* 25(5), 391-395.

Seligman, M. E., Abramson, L. Y., Semmel, A., & von Baeyer, C. (1984). Depressive attributional style. *Southern Psychologist, 2*(1), 18-22.

Seligman, M. E., & Schulman, P. (1986). Explanatory style as a predictor of productivity and quitting among life insurance sales agents. *Journal of Personality and Social Psychology*, 50(4), 832-838.

Shapcott, K. M., & Carron, A. V. (2010). Development and validation of a team attributional style questionnaire. *Group Dynamics: Theory, Research, and Practice,* 14(2), 93-113.

Shaw, M.N., Vasilopoulos, N.L. (2009). Anchored Preferential Personality Measure: Reducing the Susceptibility to Response Distortions. Paper presented to the 24th Annual Meeting of the Society for Industrial and Organizational Psychology, New Orleans.

Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality And Social Psychology, 67*(4), 674-687.

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.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*(1), 140-154.

Stark, S., Chernyshenko, O. S., Chan, K. Y., & Lee, W. C. (2001). Effects of the testing situation on item responding: Cause for concern. *Journal of Applied Psychology, 86,* 943–953.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology, 81*(2), 134-141.

Steffens, M. C. (2004). Is the Implicit Association Test Immune to Faking? *Experimental Psychology, 51*(3), 165-179.

Steffens, M. C., & König, S. (2006). Predicting spontaneous big five behavior with implicit association tests. *European Journal of Psychological Assessment, 22*(1), 13-20.

Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review, 8*(3), 220-247.

Streufert, S. (1986). Individual differences in risk taking. *Journal of Applied Social Psychology, 16*(6), 482-497.

Stotsky, B. A., Sacks, J. M., & Daston, P. G. (1956). Predicting the work performance of psychiatric aides by psychological tests. *Journal of Counseling Psychology*, 3(3), 193-199.

Teglasi, H., Nebbergall, A., & Newman, D. (2012). Construct validity and case validity in assessment. *Psychological Assessment, 24*(2), 464-475.

Tett, R. P., Freund, K. A., Christiansen, N. D., Fox, K. E., & Coaster, J. (2012). Faking on self-report emotional intelligence and personality tests: Effects of faking opportunity, cognitive ability, and job type. *Personality and Individual Differences, 52*(2), 195-201.

Thistlethwaite, D. (1950). Attitude and structure as factors in the distortion of reasoning. *The Journal of Abnormal and Social Psychology, 45*(3), 442-458.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology—Learning Memory and Cognition, 8,* 336–342.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 4157, 1124 – 1131.

Van Boven, L., Dunning, D., & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology, 79,* 66-67.

Vasilopoulos, N. L. & Cucina, J. M. (2006). Faking on non-cognitive measures: The interaction of cognitive ability and test characteristics. In R. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp.303-329). Greenwich, CT: Information Age Publishing, Inc.

Vianello, M., Robusto, E., & Anselmi, P. (2010). Implicit conscientiousness predicts academic performance. *Personality and Individual Differences, 48*(4), 452-457.

Vianello, M., Schnabel, K., Sriram, N. N., & Nosek, B. (2013). Gender differences in implicit and explicit personality traits. *Personality and Individual Differences, 55*(8), 994-999.

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

Watson, G. B. (1925). *The measurement of fair-mindedness.* Columbia Univ. Contrib. to Educ.: New York.

Windschitl, P., Conybeare, D., & Krizan, Z. (2008). Direct-comparison judgments: When and why above- and below-average effects reverse. Journal of Experimental Psychology: General, 137(1), 182-200.

Williams, L.J., & Anderson, S.E. (1994). An alternative approach to method effects by using latent-variable models: Applications in organizational behavior research. *Journal of Applied Psychology, 79*, 323 -331.

Winter, D.G., John, O.P., Stewart, A.J., Klohnen, E.C., & Duncan, L.E. (1998). Traits and motives: Towards an integration of two traditions in personality research. *Psychological Review, 105*, 230-250.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*(2), 262-274.

Woehr, D. J., & Arthur, W. r. (2003). The construct-related validity of assessment center ratings: A review and meta-analysis of the role of methodological factors. *Journal Of Management*, 29(2), 231-258.

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods, 1*(4), 354-365.

Yovel, I., & Friedman, A. (2013). Bridging the gap between explicit and implicit measurement of personality: The questionnaire-based implicit association test. *Personality and Individual Differences, 54*(1), 76-80.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal Of Personality And Social Psychology, 9*(2, Pt.2), 1-27.

Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. *Journal of Applied Psychology, 84*, 551– 563.

Zuckerman, E., & Jost, J. (2001). What makes you think you're so popular? Self-evaluation maintenance and the subjective side of the friendship paradox. *Social Psychology Quarterly, 64*(3), 207-223.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432-442.

Appendices

Appendix A. Consent & Transcript Authorization

Thank you for participating in this survey!! Your responses are valuable to this research. Today you will fill out various questionnaires which will take on average 30 minutes to complete. Before we start the study we will provide information regarding the study and proceed through the informed consent process on the next page. However, I would like to note, that since this is an online survey, and it allows you the flexibility to take it wherever you want, we just ask that you please:

1. Have reliable access to the internet
2. Take the survey alone in a quiet location without distractions (e.g., library, computer labs, etc.)
3. Please take the survey from beginning to end without breaks

Research Consent Form
Anchored Preferential Personality Measure

GW IRB number: 020982
Principal Investigator: David P. Costanza (202) 994 -1875
Sub-Investigator: Megan N. Shaw (202) 994-8960
Sponsor: (N/A)

1) Introduction
You are invited to participate in a research study under the direction of Dr. Costanza, chair of the Department of Organizational Sciences and Communications, George Washington University (GWU). Before you decide to be a part of this study, you need to understand the risks and benefits.
This consent form provides information about the research study. A staff member of the research study will be available to answer your questions and provide further explanations. If you agree to take part in the research study, you will be asked to sign this consent form. This process is known as informed consent. Taking part in this research is entirely voluntary. You are free to choose whether or not you will take part in the study. Your academic standing will not, in any way, be affected should you choose not to participate or if you decide to withdraw from the study at any time. If you decide to withdraw from this study, you will still receive the two credits that are due to you.

2) Why is this study being done?
The purpose of this study is to develop, validate, and test the effectiveness of a new employment personality measure [i.e., Anchored Preferential Personality Measure (APPM)]. The investigator in charge of this research is David P. Costanza, Ph.D. Participants at 1 institution will be asked to take part in this study. You will be one of approximately 200 participants to be asked to take part at this location. Given the study is conducted on-line, the location of the research will be conducted at your discretion.

3) What is involved in this study?
If you choose to take part in this study, you will take several personality measures, and provide a self-report of your SAT scores, GPA, and demographics information (e.g., gender, ethnicity, etc.). This two part research study has two separate conditions. Once you have signed up for study, you will be mailed a link to take the survey in a location and a time at your discretion. The first condition will not take more than 42 minutes. Once this survey is complete, you will participate in the second condition which will take no more than 32 minutes. You will receive research credit for the completion of each survey link. If you take both surveys, you will receive total of 2 research credits. Also for this study, we need to collect information to confirm the GPA and SAT scores you will be reporting in the Demographics section. This information will be valuable, as we will be using this information to investigate relationships with your responses on the questionnaire. Your GPA and SAT scores will be kept 100% confidential. Only the researcher will have access to the password protected data file. At the end of the data collection (April 21st) any identifying information will be deleted from the database and there will be no way in which your GPA and SAT scores can be directly attributed to you.

4) What are the risks of participating in this study?
Participating in this study poses no risks that are not ordinarily encountered in daily life.

5) Are there benefits to taking part in this study?
You will not benefit directly from your participation in the study. The benefits to science and humankind that might result from this study are to better measure personality which will potentially advance our understanding of human behavior in the working environment.

6) What are my options?
You do not have to participate in this study if you do not want to. Should you decide to participate and later

change your mind, you can do so at any time.

7) Will I receive payment for being in this study?
You will not be paid for taking part in this study; however you will receive two credits for your participation.

8) Can I be taken off the study?
The investigator can decide to withdraw you from the study at any time. You could be taken off the study for reasons related solely to you (for example, not following study-related directions from the Investigator) or because the entire study is stopped.

9) How will my privacy be protected?
The initial dataset will have your identifying information. However your privacy will be protected, as only one person has access to the password protected data files. Furthermore, once all the data has been collected and research credits distributed, we will assign a random code to your data and then permanently delete your name in the dataset. In addition, if the results of this research study are reported in journals or at scientific meetings, the people who participated in this study will not be named or identified. GW will not release any information about your research involvement without your written permission, unless required by law.

10) Problems or Questions
The Office of Human Research of George Washington University, at telephone number (202-994-2715), can provide further information about your rights as a research participant. If you think you have been harmed in this study, you may report this to the Principal Investigator of this study. Also, if you have felt any level of psychological stress and or discomfort regarding this study please contact the university counseling center (202-994-5300).

Further general information regarding this study may be obtained by contacting Dr. David P. Costanza in the Department Organizational Sciences and Communications by e-mail dcostanz@gwu.edu or by phone (202-994-1875). For specific questions regarding this study, you may e-mail the leader researcher, Megan Shaw (megshaw@gwu.edu).

*Please print a copy of this screen in case you want to read it again.

Name of Persons Obtaining Consent: Megan Shaw
Principal Investigator's Electronic Signature: Dr. David P. Costanza

11) Electronic Documentation of Information Consent
I understand the information on this screen. Any questions that I may have had have been answered by contacting the researchers. My electronic signature below indicates my willingness to participate in this study and my understanding that I can withdraw at any time. (If you decide to provide your consent please type in your actual name and not a pseudonym, so that I can verify that you have take the first survey and then be able to give you a link to the second survey).
First Name
Last Name
Date (Mo/Da/Year)

Do we have your permission to access your transcript? Again, information regarding your GPA and SAT scores are valuable to this research. As such we will need your transcripts in order to gather this information. These scores will be kept **100%** confidential. ONLY the researcher will have access to this information.
Yes
No

Appendix B. Participant Invite and Reminder Emails

Hello,

[Initial invite]

Below is the link to the **first (second) survey** for the study "Development of the AP Measure." The survey on average will take anywhere from 15 - 35 minutes to complete. Please fill this survey today. We allow you the flexibility to take the survey wherever you like, however we just ask for the following:

[First Reminder]

I sent an e-mail with a link for the study "Development of the AP Measure." **(If you have already taken the survey, please disregard)** For those of you who have not taken this survey, please take it today. The survey on average will take anywhere from 15 - 30 minutes to complete. Please fill this survey out in the next 24 hours. We allow you the flexibility to take the survey wherever you like, however we just ask for the following:

[Final Reminder]

This is your final reminder. I sent an e-mail with a link for the study "Development of the AP Measure." **(If you have already taken the survey, please disregard)** For those of you who have not taken this survey, you will need to take the survey today in order to receive your research credits. The survey on average will take anywhere from 15 - 30 minutes to complete. Please fill this survey out in the next 24 hours. We allow you the flexibility to take the survey wherever you like, however we just ask for the following:

[Email Instructions Provided in all correspondence]

1. Pick a quiet location where you won't be bothered for at least 35 minutes
2. Be in a location that has a reliable internet access
3. Take the survey from beginning to end without stopping

Thank you for your participation!!

- Megan

P.S. Be very careful to respond to each question. If you missed a question you will receive the warning "You must provide an answer for QXX." However, the question in the warning, does not indicate the question that actually has the missing response. You will have to search on the page for the question that has the missing response and respond to it. This will make more sense when taking the survey. If you have any questions, please feel free to contact me.

P.P.S. When you are finished with the survey make sure you click on the "submit" button or your responses will not be recorded!

Appendix C. Example List of APPM Items

Item		Original Facet	Anchor A	Anchor B	Decision	Final Facet
1	-	Ach	Competition can stimulate individual performance.	Competition can stimulate conflict between coworkers.	Drop - S2	-
2	+	Ach	Laid back people are higher performers.	Intense people are higher performers.	Keep	ACH
3	-	Caut	All possible scenarios need to be thought through before acting.	Only the most important scenarios need to be thought through before acting.	Keep	CAUT
4	-	Caut	Being reserved is stuffy.	Being spontaneous is overrated.	Drop - S1	-
5	-	Disci	Most people don't stick to diets because they don't have self-control.	Most people don't stick to diets because life gets very busy.	Keep	ORDER
6	+	Disci	It is better to give your mind a few breaks to complete a task effectively.	It is better to work without interruptions to complete a task effectively.	Drop - S1	-
7	+	Duti	It is good to challenge the system and be a revolutionary.	It is good to follow the system and be predictable.	Keep	ORDER
8	-	Duti	Right and wrong is typically black and white.	Right and wrong is typically gray.	Drop - S2	-
9	-	Effic	Success is largely influenced by the person.	Failure is largely influenced by factors outside the person.	Keep	ACH
10	+	Effic	It is good to be aware of weaknesses.	It is good to be aware of strengths.	Drop - S1	-
11	-	Order	When given a task, it is instinctual to start planning.	When given a task, it is instinctual to immediately get things done.	Keep	CAUT
12	+	Order	It is normal to turn the other cheek when people are messy.	It is normal to get upset when people are messy.	Drop - S1	-
90	-	Order	Working with people that must have things "just so" can be helpful.	Working with people that have to have things "just so" are controlling.	Drop - S2	-

Appendix D. Variable Intercorrelations in Honest and Faking Conditions

	N	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. A-TOTAL	186	3.58 (*3.62*)	.38	-	*.49***	*.30***	*.69***	*.62***	*.43***	*.35***	*.39***	*.40***	*.28***	*.24***	*.38***	*-.12*	*.01*	*-.07*	*-.09*
2. A-ACH	186	4.01 (*3.96*)	.64	.56**	-	*-.04*	*.09*	*.10*	*.34***	*.30***	*.26***	*.34***	*.20***	*.33***	*.21***	*.04*	*-.04*	*.02*	*.04*
3. A-CAUT	186	4.09 (*4.07*)	.74	.41**	-.05	-	*-.24***	*-.16**	*-.04*	*.00*	*-.02*	*-.11*	*.05*	*.00*	*-.05*	*-.13*	*-.08*	*-.03*	*-.02*
4. A-DUTI	186	3.03 (*3.19*)	.54	.74**	.23**	-.05	-	*.49***	*.33***	*.23***	*.33***	*.37***	*.19**	*.12*	*.30***	*.01*	*.16**	*-.06*	*-.10*
5. A-ORD	186	3.35 (*3.46*)	.65	.65**	.18*	.07	.50**	-	*.29***	*.25***	*.25***	*.25***	*.15**	*.08*	*.38***	*-.22**	*-.08*	*-.08*	*-.10*
6. N-TOTAL	186	3.62 (*3.87*)	.47	.37**	.27**	.27**	.15*	.25**	-	*.80***	*.81***	*.86***	*.83***	*.85***	*.72***	*.10*	*.26***	*-.11*	*-.10*
7. N-ACH	186	3.89 (*4.09*)	.66	.29**	.29**	.16*	.14	.12	.77**	-	*.54***	*.57***	*.64***	*.63***	*.47***	*.05*	*.13*	*-.10*	*-.11*
8. N-CAUT	186	3.39 (*3.63*)	.71	.33**	.18*	.24**	.17*	.22**	.74**	.42**	-	*.64***	*.59***	*.58***	*.50***	*.07*	*.17**	*-.12*	*-.13*
9. N-DISCI	186	3.03 (*3.55*)	.72	.26**	.13	.18*	.18*	.19**	.77**	.60**	.42**	-	*.61***	*.70***	*.57***	*.17*	*.21***	*-.05*	*-.02*
10. N-DUTI	186	4.11 (*4.21*)	.53	.14	.22**	.21**	-.11	.00	.68**	.48**	.50**	.32**	-	*.78***	*.47***	*.09*	*.27***	*-.07*	*-.06*
11. N-EFFI	186	3.93 (*4.11*)	.53	.12	.25**	.11	-.02	-.05	.71**	.56**	.40**	.49**	.59**	-	*.49***	*.21**	*.24***	*-.05*	*-.02*
12. N-ORD	186	3.34 (*3.63*)	.74	.40**	.13	.25**	.22**	.46**	.63**	.30**	.40**	.43**	.23**	.21**	-	*-.10*	*.24***	*-.14*	*-.13*
13. SAT	126	1251	12.84	-.08	.15	-.09	-.09	-.16	-.03	.05	.00	-.05	.02	.22*	-.25**	-	*.22**	*.03*	*.04*
14. GPA	155	3.14	.52	.26**	.21**	.11	.22**	.00	.30**	.30**	.17*	.27**	.17*	.23**	.14	.22*	-	*-.19**	*-.22***
15. REMIND	174	2.37	2.07	-.15*	-.03	-.10	-.07	-.18*	-.11	.01	-.13	-.07	-.02	-.05	-.19**	.03	-.19*	-	*.95***
16. DAYS	175	1.4	1.70	-.17*	-.01	-.12	-.09	-.20**	-.11	.02	-.15*	-.06	-.01	-.02	-.20**	.04	-.22**	.95**	-

A-TOTAL: APPM Composite; **A-ACH:** APPM Achievement; **A-CAUT:** APPM Cautiousness; **A-DUTI:** APPM Dutifulness; **A-ORD:** APPM Orderliness; **N-TOTAL:** NEO Composite; **N-ACH:** NEO Achievement; **N-CAUT:** NEO Cautiousness; **N-DISCI:** NEO Discipline; **N-DUTI:** NEO Dutifulness; **N-EFFI:** NEO Efficacy; **N-ORD:** NEO Orderliness; **SAT:** SAT Scores; **GPA:** Cumulative GPA; **REMIND:** Number of reminders to complete the survey; **DAYS:** Number of days from study completion to study start.

Italicized numbers are the predictor faking variables.

Note: The same size for the faking condition is N = 177.

* p < .05; ** p < .01

Appendix E. Analysis Replication Using NEO-IPIP Composite That Matches the APPM Facets (Excludes Discipline and Dutifulness Facets)

(Similar to Table 5.) Intercorrelations of the APPM, NEO-IPIP (Composite 1), and the NEO-IPIP scale that match APPM constructs (Composite 2)

	APPM Composite	APPM Achieve	APPM Orderliness	APPM Cautious	APPM Dutifulness
NEO - Composite 1	.37***	.27***	.25***	.27***	.15**
NEO – Composite 2	.41***	.28***	.28***	.27***	.19**
NEO - Achieve	.29***	.29***	.12*	.16**	.14*
NEO - Cautious	.33***	.18**	.22**	.24***	.17*
NEO - Orderliness	.40***	.13*	.46***	.25***	.22**
NEO- Discipline	.26***	.13*	.19***	.18**	.18**
NEO- Dutifulness	.14*	.22***	.00	.21**	-.11
NEO-Efficacy	.12	.25***	-.05	.11	-.02

*** Significant $p < .001$; ** Significant $p < .05$; *Significant p .10

(Similar to Table 7.) Correlations of the APPM, NEO-IPIP (Composite 1), and the NEO-IPIP scale that match APPM constructs (Composite 2) with Explicit and Implicit Criteria

	GPA	DVTime	DVRemind
NEO – Composite 1	.30**	-.12	-.11
NEO – Composite 2	**.29****	**-.13**	**-.13**
APPM – Composite	.26**	-.15*	-.17*

Note: **DVTime:** Implicit Criteria 1 (Difference in time); **DVRemind:** Implicit Criteria 2 (# of reminders)
** Significant $p < .01$; *Significant $p < .05$
N = 155-175

(Similar to Table 8.) Summary of Hierarchical Regression Analysis for NEO-IPIP with the deliberate and implicit criteria

	Step 1		Step 2
Variable	SAT	APPM	NEO - Composite 2
DV: GPA			
β	**.24*** **	**.30*** **	**.16***
ΔF	**9.37*** **		**3.10***
ΔR^2	.14		**.02***
R^2	.14		.16
DV: Time Difference			
β	.01	**-.16***	.07
ΔF	1.46		.45
ΔR^2	.03		.00
R^2	.03		.00
DV: Reminder Count			
β	.02	**-.17***	.08
ΔF	1.86		.71
ΔR^2	.03		.00
R^2	.03		.01

Note: **SAT:** Scholastic Aptitude Test; **APPM:** Anchored Preferential Personality Measure; **NEO:** Neuroticism-Extraversion-Openness Inventory (Includes Conscientiousness & Agreeableness; **GPA:** Cumulative Grade Point Average; **Time Difference:** Time difference from starting first survey to the completion of the second survey; **Reminder Count:** Number of total reminders sent before second survey was complete.

* $p < .10$. ** $p < .05$. *** $p < .01$

(Similar to Table 9). Summary of Hierarchical Regression Analysis for APPM with the deliberate and implicit criteria

	Step 1		Step 2
Variable	SAT	NEO – Composite 2	APPM
DV: GPA			
β	**.23*** **	**.25*** **	**.24** **
ΔF	**7.49*** **		**6.50** **
ΔR^2	.11		**.05** **
R^2	.11		.14
DV: Time Difference			
β	.03	.01	-.18
ΔF	.06		3.23
ΔR^2	.00		.03
R^2	.00		.03
DV: Reminder Count			
β	.04	.06	**-.20** **
ΔF	.09		**4.25** **
ΔR^2	.00		**.04** **
R^2	.00		.05

Note: **SAT:** Scholastic Aptitude Test; **APPM:** Anchored Preferential Personality Measure; **NEO:** Neuroticism-Extraversion-Openness Inventory (Includes Conscientiousness & Agreeableness; **GPA:** Cumulative Grade Point Average; **Time Difference:** Time difference from starting first survey to the completion of the second survey; **Reminder Count:** Number of total reminders sent before second survey was complete.
*$p < .05$ **$p < .01$

Table 10. Paired Sample T-Tests between Faking and Honest Conditions for the NEO-IPIP and the APPM.

	Fake		Honest		95% CI					
Variable	Mean	SD	Mean	SD	Lower	Upper	t	df	Sig. (2-tailed)	Cohen's *d*
NEO Faking – Honest (Composite 1)	3.87	.58	3.63	.47	0.16	0.32	6.06	176	.00	.45
NEO Faking –Honest (Composite 2)	3.86	.56	3.65	.48	.13	.29	5.39	176	.00	.60
APPM Faking – Honest	3.62	.32	3.59	.38	-0.02	0.08	1.18	176	.24	.09

Appendix F. Analysis Replication Using study order as a Covariate

(Similar to Table 8). Summary of Hierarchical Regression Analysis for NEO-IPIP with the deliberate and implicit criteria with Study Order as a Covariate

	Step 1	Step 2		Step 3
Variable	Order	SAT	APPM	NEO
DV: GPA				
β	.11	**.29*** **	**.26*** **	**.21****
ΔF	1.31	**9.67*** **		**5.17****
ΔR^2	.01	.14		.04
R^2	.01	.15		.19
DV: Time Difference				
β	-.02	.02	-.17	.13
ΔF	.03	1.83		1.67
ΔR^2	.00	.03		0.01
R^2	.00	.03		0.05
DV: Reminder Count				
β	-.04	0.01	-.18	0.12
ΔF	.16	1.41		.90
ΔR^2	.00	0.02		0.01
R^2	.00	0.03		0.03

Note: **SAT:** Scholastic Aptitude Test; **APPM:** Anchored Preferential Personality Measure; **NEO:** Neuroticism-Extraversion-Openness Inventory (Includes Conscientiousness & Agreeableness; **GPA:** Cumulative Grade Point Average; **Time Difference:** Time difference from starting first survey to the completion of the second survey; **Reminder Count:** Number of total reminders sent before second survey was complete.
*$p < .05$ **$p < .01$

(Similar to Table 9). Summary of Hierarchical Regression Analysis for APPM with the deliberate and implicit criteria

	Step 1	Step 2		Step 3
Variable	Order	SAT	NEO	APPM
DV: GPA				
β	.11	**.28*** **	**.27*** **	**.18** **
ΔF	1.31	**10.21*** **		**4.21** **
ΔR^2	.01	**.15*** **		**.03** **
R^2	.01	.19		.16
DV: Time Difference				
β	-.04	.03	**-.16***	-.18
ΔF	.16	.12		3.47
ΔR^2	.00	.00		.
R^2	.00	.00		.03
DV: Reminder Count				
β		.02	**-.17***	.12
ΔF		1.86		1.63
ΔR^2		.03		.01
R^2		.03		.05

Note: **SAT:** Scholastic Aptitude Test; **APPM:** Anchored Preferential Personality Measure; **NEO:** Neuroticism-Extraversion-Openness Inventory (Includes Conscientiousness & Agreeableness; **GPA:** Cumulative Grade Point Average; **Time Difference:** Time difference from starting first survey to the completion of the second survey; **Reminder Count:** Number of total reminders sent before second survey was complete.

** $p < .05$ *** $p < .01$