THE IMPLICIT ASSOCIATION TEST FOR CONSCIENTIOUSNESS: AN INDIRECT METHOD OF MEASURING PERSONALITY

Ryan D. Anderson

A Thesis

Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

Master of Arts

August, 2009

Committee:

Michael J. Zickar, Advisor
Scott Highhouse
J. Devin McAuley
ABSTRACT

Michael J. Zickar, Advisor

The self-report method is commonly cited as a limitation in the personality literature. Therefore, this thesis was conducted to determine whether an Implicit Association Test (IAT) might serve as an alternative measure of conscientiousness; thus providing researchers with an additional measure to complement existing measures and to also provide a more intricate understanding of the personality construct. Similarly, this investigation was designed to extend the potential scope of the IAT, from predominantly attitude research, to the domain of personality; and to provide some clarification on conflicting studies that have incorporated the IAT as a measure of conscientiousness. An IAT was created to assess the relative strength of associations related to the self (vs. other) with conscientiousness adjectives (positive vs. negative). The aim of this thesis was to determine whether the IAT would relate to previously validated indicators of conscientiousness (i.e., self-report, academic motivation, a behavioral measure of attention, and several objective indicators of academic success). The results for the IAT were nonsignificant for all hypotheses. Further analysis, revealed several outliers and violations of normality, however, these problems did not contribute to the lack of findings. Therefore, the discussion section suggests several potential problems with the current investigation and offers several suggestions for future research.
I dedicate this thesis to my parents. In particular, I want to thank my mother for serving as my bedrock of love and security from which I have based my life. I want to thank my father for the countless hours he spent as my baseball coach instilling in me the self-confidence to persevere through adversity. Finally, I would be remiss if I did not thank my stepfather, Paul Wocken, whose continued confidence and support has provided me the opportunity to pursue my passions. I am truly grateful to all of you for your unconditional love and support.
ACKNOWLEDGMENTS

I would like to thank my advisor, Mike Zickar, for his insight, guidance, and the patience he provided while allowing me the opportunity to pursue my research interests throughout the thesis process. I would also like to thank Scott Highhouse and Devin McAuley for asking the tough questions and providing helpful feedback to enhance my research and ensure my continual development throughout the process.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER I. INTRODUCTION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limitations of Explicit Measures</td>
<td>1</td>
</tr>
<tr>
<td>The Implicit Association Test (IAT)</td>
<td>2</td>
</tr>
<tr>
<td>IAT Studies of Personality</td>
<td>4</td>
</tr>
<tr>
<td>Psychometric Properties of the IAT (and other alternative measures)</td>
<td>5</td>
</tr>
<tr>
<td>The Big-5 Traits</td>
<td>7</td>
</tr>
<tr>
<td>Conscientiousness and Academic Performance</td>
<td>9</td>
</tr>
<tr>
<td>The IAT-Conscientiousness</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER II. METHODS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>13</td>
</tr>
<tr>
<td>Measures</td>
<td>13</td>
</tr>
<tr>
<td>Explicit Measures</td>
<td>13</td>
</tr>
<tr>
<td>Implicit Measure</td>
<td>14</td>
</tr>
<tr>
<td>Criterion</td>
<td>14</td>
</tr>
<tr>
<td>Procedures</td>
<td>15</td>
</tr>
<tr>
<td>IAT Procedure</td>
<td>15</td>
</tr>
<tr>
<td>IAT Data Analysis</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER III. RESULTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive Statistics</td>
<td>17</td>
</tr>
<tr>
<td>Explicit Measures</td>
<td>17</td>
</tr>
<tr>
<td>Criterion from Self-Report Survey</td>
<td>18</td>
</tr>
</tbody>
</table>
Outlier Analysis ................................................................. 17

Sex .................................................................................. 18

Ethnicity ........................................................................... 18

Order Effect of IAT ............................................................ 19

Hypotheses ....................................................................... 19

CHAPTER IV. DISCUSSION ...................................................... 22

Similar Findings ................................................................. 23

Implicit Versus Explicit Measures ....................................... 24

Progression of Implicit Measures ....................................... 25

Implicit-Explicit Correspondence and Possible Moderators ...... 27

Translation ......................................................................... 28

Explicit Measurement .......................................................... 29

Representation Strength ..................................................... 29

Implicit Self-Observation ..................................................... 30

Dimensionality and Personality .......................................... 30

Manipulating the Situation .................................................. 31

Future Research ................................................................. 32

Concluding Remarks .......................................................... 33

REFERENCES ................................................................. 34
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Attribute and Concept Stimuli for IAT</td>
<td>44</td>
</tr>
<tr>
<td>2</td>
<td>Conceptual Example of IAT Task</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Detailed Example of IAT Task</td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td>Descriptive Statistics and Zero-Order Correlations</td>
<td>47</td>
</tr>
</tbody>
</table>
CHAPTER I: INTRODUCTION

Predicting performance using personality inventories remains an important topic in academic and applied research in Industrial and Organizational psychology. This focus has increased considerably since Barrick and Mount’s (1991) meta-analysis of the Big-5, which revealed that personality instruments can be a vital tool for identifying consistent behavioral differences between individuals. More precisely, the authors found that the conscientiousness dimension was a significant predictor on all job performance criteria (i.e., job proficiency, personnel data, and training proficiency) across a range of occupations. Despite these robust findings, some researchers have noted that there appears to be a ‘0.30 prediction barrier’ that remains, for the most part, unbroken (Pervin, 1994). The aim of this study is to investigate a relatively new method of personality assessment called the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998) in order to increase the predictive reliability and validity of personality measures that have already been empirically established.

The predominant method used in assessing personality is self-report surveys. Other researchers have taken notice of this fact and have argued that multiple methods should be used for a variety of reasons. For example, researchers have noted the limitations of studies in which there is a common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). This confound is particularly concerning when there is reason to believe that the measurement instrument (e.g., self-report surveys) can introduce a bias (e.g., impression management) simply as a consequence of the instruments inherent limitations. Similarly, researchers in social psychology have illustrated the inherent limitations with self-report measures, noting that self-report surveys require an honest and accurate appraisal of oneself (Nosek, Greenwald, & Banaji, 2006). For instance, these authors state that if an individual is interested in measuring math
ability, “…one approach…would be to ask ‘how good are you at math?’ whereas an alternative approach is to infer math ability via performance on a math skills test. The former requires introspection to assess the relevant construct; the latter does not. And yet, the latter is accepted as a measure of math ability, and is even preferred to one requiring self-assessment.” (p.2). In other words, unlike explicit measures, which are generally based on the assumption that honest and accurate responses will be provided upon completion of a comprehensive introspective process, implicit measures such as the IAT are designed to indirectly assess a construct by measuring differences in response latency, thus providing researchers with a measure of the differing strength levels of relatively automatic cognitive associations. Therefore, it is proposed that the personality literature would benefit from an indirect measure of conscientiousness that does not require introspection and is resistant to the common barriers that are associated with self-report measures. Thus, the goal of the current study is to develop an implicit measure of conscientiousness and collect initial validation data. However, before reviewing the relevant literature on the IAT it is important to review the limitations of explicit measures.

Limitations of Explicit Measures

Researchers have identified at least two barriers encountered when asking participants to provide information about themselves on self-report surveys that can potentially affect the accuracy and validity of the construct being measured (Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott, D. S., 2002; Haines & Sumner, 2006). The first barrier, response factors, is an acknowledgement that when participants are asked to self-report, there may be differing levels of motivation to purposefully distort aspects of oneself depending upon the topic and context. Importantly, researchers have recognized the potential influence of this barrier for over forty years (for a review of faking, see Zickar & Gibby, 2006). Specifically, researchers have argued
that self-reports are susceptible to demand characteristics (Orne, 1962), evaluation apprehension (Rosenberg, 1969), and impression management (Tedeschi, Schlenker, & Bonoma, 1971; Weber & Cook, 1972, Paulhus, 1984). Importantly, this potential confound becomes an increasing concern when participants are aware that their personal interests are at stake. This concern is clearly justified upon examination of a book over fifty years old, which contains an appendix titled, “How to cheat on personality tests” (Whyte, 1956). This book perhaps set into motion an unofficial struggle between individual employees and/or job applicants and the researchers striving to stay ahead of the public in order to ensure validity of their instruments. Furthermore, the persistence of this struggle in the digital era is evident in a recent article available to anyone with access to the Internet, which describes how answer keys to personality tests have been made available on the Internet (e.g., Facebook and Wikipedia) (O’Connell, 2009, January 7).

The second barrier, introspective limits, warrants calling into question an individual’s ability to accurately assess their attitudes. This concern, illustrates the inherent measurement limitations in self-report surveys (Greenwald & Banaji; 1995; Nisbett & Wilson, 1977; Asendorpf, Banse, & Mucke, 2002; Haines & Sumner, 2006). Nosek et al. (2006) illustrate this point by noting that explicit measures require an ability to introspect, and that the very process of introspection creates a sense of ‘knowing.’ This barrier may help explain the incongruence often observed between participants’ self-ratings on a construct of interest and their actual behavior, despite genuine attempts by participants to provide accurate self-reports. Indeed, research indicates that in social situations with highly sensitive subject matter, the accuracy of explicit measures may be suspect (Nosek et al., 2006; Poehlman, Uhlmann, Greenwald, & Banaji, 2005). This research highlights the fact that even though an individual may respond genuinely, they may be unable to accurately assess themselves.
In light of these potential measurement confounds, a growing body of researchers have advocated the use of alternative measures. In a recent article, Haines and Sumner (2006) argue that researchers should begin to consider incorporating implicit measures such as the IAT because the different approach toward psychological measurement may act as an enhancement (or correction) to traditional measurement by advancing our understanding of relevant theory, prediction, and application. Similarly, the authors of a recent meta-analysis on the IAT concluded that a promising area for future research will be incorporating IAT measures with self-reports to improve criterion prediction (e.g., judgments, choices, physiological responses, and behaviors) (Poehlman, et al., 2005).

The Implicit Association Test (IAT)

The Implicit Association Test is a categorization task that measures the relative difference in association strength between pairs of concepts (Greenwald et al., 1998). More specifically, the authors state that the IAT is designed to measure the underlying automatic evaluations of a given affect or attitude by using the speed of participants’ responses as an indicator of how strongly (or weakly) constructs are associated. The procedure for the IAT, in brief, consists of asking participants to sort stimuli from one of four categories on a computer monitor into their respective category using just two response options on the keyboard. This task is designed to capitalize on the biological and cognitive structure of a target construct, thus making it extremely resistant to the common barriers encountered with self-reports. In other words, the IAT measures how strongly or weakly a category (e.g., I versus other) is associated in the brain with a concept (i.e., conscientious adjectives) using response latencies as an indicator.

In the eleven years since the development of the IAT, researchers have demonstrated the utility of these measures across an array of disciplines and domains. Haines and Sumner (2006)
highlight just some of the research areas that have incorporated the IAT. For example, the IAT has been employed in research pertaining to marketing (Brunel, Tietje, & Greenwald, 2004; Maison, Greenwald, & Bruin, 2001), health and clinical psychology (Marsh, Johnson, & Scott-Sheldon, 2001; Teachman, Gapinski, Brownell, Rawlins, & Jeyaram, 2003; Teachman, Gregg, & Woody, 2003), neuroscience (Phelps et al., 2000), social-psychology (Greenwald & Farnham, 2000; Rudman, Greenwald, Mellott, & Schwartz, 1999) and personality (e.g., anxiety and shyness) (Asendorpf, et al., 2002, Egloff & Schmukle, 2002; Egloff, Wilhelm, Neubauer, Mauss, & Gross, 2002). In addition, a recent meta-analysis of 61 studies provided further empirical support for previous IAT studies (Poehlman, et al, 2005). Importantly, the aim of this study is to further explore the potential of using the IAT as a measure of personality. Therefore, a more extensive review of the studies that have used the IAT to measure personality is warranted.

**IAT Studies of Personality**

Although the IAT was originally developed as a unique measure of attitudes in social psychology (i.e., prejudice), evidence suggests that it can also be applied to personality measurement and researchers have often acknowledged that this is an area that requires more attention in the literature (e.g., Nosek, 2005). In the first study designed to measure individual differences in participants’ self-concepts of trait shyness, Asendorpf et al. (2002) employed psychometrically verified adjectives of shyness in the design of an IAT of shyness. This study was based upon previous research findings that indicate that the self is processed in two different modes, generally distinguished by experts as an *explicit mode* and an *implicit mode*.

Importantly, the explicit mode is characterized by conscious, controlled, and reflective information processing, whereas the implicit mode is characterized by unconscious, automatic, and intuitive processes. As predicted, this study revealed what the authors termed a *double*
dissociation between the measures employed and the behaviors they predict. In other words, implicit and explicit measures were better predictors of spontaneous (i.e., highly automatized behavior) and controlled shy behavior, respectively. The authors also conducted a second study that included a faking condition (i.e., a job application scenario). Interestingly, those participants instructed to fake being non-shy were able to adjust their controlled behavior, but they were unable to adjust their spontaneous shy behavior. In addition, these same participants’ scores for explicit shyness decreased by more than one standard deviation, but their IAT scores remained stable. In their final discussion, Asendorpf et al. (2002) state, “…we have shown that it is possible to apply procedures such as the IAT to the assessment of implicit traits. Shyness is just one example; in principle, any trait that can be described by adjectives can be studied by an IAT procedure” (p. 392).

Similar studies that have incorporated the IAT have bolstered Asendorpf et al.’s. (2002) claim that the IAT is a viable tool that researchers may use to assess personality. Specifically, researchers have found evidence that the IAT can be applied to other domains of personality, such as anxiety (Egloff & Schmukle, 2002; Schnabel, Banse, & Asendorpf, 2006). It is important to note that these studies also demonstrated that the IAT added incremental validity to the prediction of the target behavior that is not typically captured by standard self-report measures. In all of the aforementioned studies of personality that have incorporated the IAT, researchers have noted that their particular domain of interest for these studies were chosen because they contain a social desirability component and that the IAT particularly resistant to this influence.

Similar to the previously mentioned studies, researchers recently revealed that four of the Big-5 dimensions of personality (i.e., neuroticism, extraversion, agreeableness, and
conscientiousness) predicted spontaneous behavior better when employing separate IAT’s for each dimension than did the respective explicit measures (Steffens & König, 2006). Importantly, the IAT for conscientiousness revealed the highest internal consistency and the largest correlation with spontaneous behavior among all the big-5 dimensions. More precisely, the authors used the d2 test of attention (Brickenkamp, 1981) as an indicator of spontaneous conscientious behavior to determine whether individuals high on conscientiousness would complete this mundane attention task more slowly and with fewer mistakes. As expected, the authors found a significant relationship between implicit scores of conscientiousness and the number mistakes on the d2 test. However, these findings were not analyzed to determine whether the IAT tests added incremental validity above and beyond the explicit measures. In this study, I created an IAT measure of conscientiousness to compare the results to an explicit measure.

Hypothesis 1: The IAT for conscientiousness will significantly predict spontaneous behavior better than the established self-report measures.

Psychometric Properties of the IAT (and other alternative measures)

According to Haines and Sumner (2006), the indirect nature of the IAT must be distinguished from other alternative indirect measures that have been used, such as projective tests (e.g., Rorschach Inkblot Method, Thematic Apperception Test, Draw a Person Test). These authors contrast the assumptions between these measures by noting that, whereas projective personality tests reveal aspects of the self by requiring individuals to evaluate and place their own structure on ambiguous stimuli (Anastasi & Urbina, 1997), the IAT’s development can be traced to research on social cognition that uses response latencies during a categorization task of common stimuli as an indicator association strength (Greenwald et al., 2002).
Overall, the IAT has consistently demonstrated acceptable psychometric properties that far surpass previous implicit measurement attempts such as projective tests (Haines & Sumner, 2006). On the one hand, projective tests across numerous studies have demonstrated undesirable psychometric characteristics (e.g., low reliability and low validity), thus warranting skepticism when drawing conclusions from studies employing this method of research (e.g., Garb, Wood, Lilienfeld, & Nezworski, 2002; Grove, Barden, Garb, & Lilienfeld, 2002). On the other hand, IAT research indicates that the psychometric properties of this measure generally fall within acceptable professional standards (e.g., Banse, Seise, & Zerbes, 2001; Greenwald & Nosek, 2001; Kim & Greenwald, 1998; Nosek et al., 2002). Specifically, the IAT measures are internally consistent (ranging from .7 to .9) (Bosson, Swann, & Pennebaker, 2000; Dasgupta & Greenwald, 2001; Greenwald & Nosek, 2001; Greenwald & Farnham, 2000, Schmukle & Egloff, 2004), are not confounded by subject’s familiarity with IAT stimuli (Dasgupta, McGhee, Greenwald, & Banaji, 2000; Ottaway, Hayden, & Oakes, 2001; Rudman et al., 1999), and are relatively insensitive to methodological factors (e.g., the number of trials and type of target stimuli) (Nosek, Greenwald, & Banaji, 2005; Greenwald et al., 1998).

Recent evidence suggests that implicit and explicit measures will often significantly correlate (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005a). However, research using the IAT has produced some mixed results, leaving researchers to question what the IAT is measuring in comparison to self-report measures. For example, some researchers interested in attitudes across a variety of domains have found substantial correlations between the IAT and self-report measures (e.g., Banse, Seise, & Zerbes, 2001), and yet others have found no significant relationship (e.g., Karpinski & Hilton, 2001). Similarly, there have been mixed results regarding the correlations between implicit and explicit measures of conscientiousness.
This latter study, which was released since the initial proposal, found negligible results whereas the former found positive results. It is also important to note that stronger correlations have been found for criterion when the IAT and explicit measures are strongly correlated; however, both measures can reliably predict behavior regardless of their correspondence (Poehlman et al., 2005). A recent meta-analysis investigating Implicit and Explicit (I-E) correspondence revealed a small, but significant correlation ($r = 0.24$) between self-report and IAT measures (Hofmann, et al., 2005a). The authors conclude that these two methods are tapping the same construct, but that they each provide unique variance when entered into multiple regression.

Hypothesis 2: It is predicted that the IAT-Conscientiousness will correlate with the IPIP (explicit) measure of conscientiousness.

Hypothesis 3: In addition to these measures correlating, it is predicted that the IAT measure of conscientiousness will provide a significant amount of unique variance when entered into multiple regression with the explicit measure.

Clearly, applying the IAT to organizational studies is an area ripe for exploration, and the domain of task performance is a natural starting point. Therefore, a review of the relevant literature on the relationship between personality and performance is necessary.

*The Big-5 Traits*

Years of personality research investigating individual differences have revealed five dimensions of personality that are considered to be universal, stable, and extremely influential on behavior (e.g., Cattel, Eber, & Tatsuoka, 1970; Goldberg, 1981; Digman & Inouye, 1986; Digman, 1989; Costa & McCrae, 1992). Of these five dimensions (i.e., Neuroticism, Extraversion, Openness, Conscientiousness, and Agreeableness), conscientiousness has received
considerable attention in the work performance literature (Barrick & Mount, 1991).
Interestingly, recent evidence suggests that the same individuals who are desirable employees
(i.e., because of their high scores on conscientiousness) are also the same individuals who are most likely to distort their scores (McFarland & Ryan, 2000). Thus, developing a measure that is resistant to faking would be an important advancement for the personality literature.

Among the numerous instruments developed to measure conscientiousness, one approach in particular is well suited to be adapted to the IAT format. The instrument developed by Goldberg (1992), employs a unipolar approach to the Big-5 that asks participants to respond on a Likert scale to adjectives that have been psychometrically validated and also demonstrate a significant relationship with the construct of interest (i.e., conscientiousness). Similarly, Saucier (1994) developed an abbreviated version of Goldberg’s adjective checklist called the ‘Mini-Markers.’ The item format of these scales is ideal for incorporating into the IAT because each item consists of a single word. Therefore, this study will draw from these empirically validated measures by selecting adjectives as stimulus items for the IAT-conscientiousness.

Conscientiousness and Academic Performance.

Understanding the relationship between personality and academic performance is an important question for both academic and applied researchers. This is becoming increasingly important as organizations make decisions among a diverse array of training programs (Towler & Dipoye, 2003). For example, recent evidence investigating how personality traits and learning style preferences interact, found the strongest effects between conscientiousness and preference for a structured learning environments (i.e., conscientious individuals prefer orderly and planned learning). Research has also demonstrated that conscientiousness is related to both goal orientation and achievement motivation, and that these relationships are predictive of academic
performance. Zweig and Webster (2004), found that goal orientation and personality are related, but distinct concepts and that goal-orientation mediates the relationship between personality and performance intentions. Specifically, conscientiousness was related to adaptive goal orientations (i.e., positively related to mastery learning, positively related to performance approach, and negatively related to performance avoidance). Similarly, conscientiousness was found to be the strongest predictor of achievement motivation, accounting for over 29% of the variance (Komarraju & Karau, 2005). This research suggests that conscientious individuals are likely to have an approach toward learning that is conducive to their academic success (i.e., structured learning style, goal-orientation, and achievement motivation).

Hypothesis 4a: The IAT for conscientiousness will positively correlate with indicators of academic success (e.g., GPA, ACT, and SAT) and negatively correlate with average days they skip class per month.

In order to provide a stronger test of the validity of the IAT for conscientiousness, this study will also test whether an existing measure of goal-orientation (GOALS-S, Dowson & McInerney, 2004) is related to conscientiousness.

Hypothesis 4b: Scores on the IAT will be significantly related to all three orientations as suggested by the previously cited literature (i.e., + for mastery learning, + performance approach, - performance avoidance).

The IAT-Conscientiousness

Although the actual task for the Implicit Association Test is relatively simple, great care must go in to the design of the test (Greenwald, Nosek, & Banaji, 2003). More specifically, the IAT requires researchers to select distinct category labels and stimuli that reflect the construct of interest (Nosek et al., 2006). These categories are referred to as the attribute and concept
categories. The central idea behind the IAT task is that response latencies that occur during the stimulus categorization task are indicative of the relative association strength between these two categories.

The current study will use previously validated items (i.e., adjectives) to comprise the *attribute category* of conscientiousness (Goldberg, 1992; Goldberg, 1999; Saucier, 1994) (see Table 1). The second category will be based upon the *concept category* developed by Greenwald et al. (2002) (i.e., ‘ideographic self’ and ‘other’ items). Karpinsky (2004) criticized the use of this type of category for assessing implicit attitudes, claiming that the *other* category is inappropriate when contrasted with the *self*. However, Pinter and Greenwald (2005) argued that Karpinsky erroneously assumed that the two categories have the same association strengths. Their results demonstrated that the categories of *self* and *other* (represented in several variations and across cultures) are in fact distinct concepts. Furthermore, they demonstrated that the *other* category is significantly indistinguishable from *middle*. Therefore, these results justify the *other* category as a neutral valence and an acceptable category for comparison with the *self*. 
CHAPTER II: METHODS

Participants

Ninety undergraduates enrolled in a psychology course from a public university in Ohio were recruited to participate in the experiment as part of their course requirements. Participants ranged in age between 18 and 63 years, with an average age of 19.6 years ($SD = 5.04$). The sample was comprised of 63% females, 30% males, and 7% did not disclose. The sample was predominantly Caucasian (74%). The remaining participants’ ethnic background was as follows: African American (12%), Latino (3%), Middle Eastern (2%), Asian (1%), and no response (8%).

Measures

Explicit Measures. The twenty-item scale of conscientiousness from the IPIP (Goldberg, 1999) served as the explicit personality measure ($\alpha = .90$). Participants were asked to self-report their perceived levels of conscientiousness from 1 (very inaccurate) to 5 (very accurate) for items that were both positively and negatively scored (i.e., “Do things according to plan” and “Waste my time” respectively).

The GOALS-S measure (Dowson & McInerney, 2004) served to assess participants’ academic goal orientation on three dimensions (i.e., mastery “I want to do well at school to show that I can learn new things,” performance approach “I want to do well at school because being better than others is important to me,” and performance avoidance “I want to do well at school, but only if the work is easy”). Participants completed the 18 item goal orientation measure on a 5-point scale, ranging from 1 (strongly disagree), 3 (not sure), to 5 (strongly agree). All three motivational subscales demonstrated acceptable levels of internal consistency (i.e., Mastery motivation $\alpha = .85$, Performance motivation $\alpha = .89$, and Avoidance motivation $\alpha = .80$).
Implicit Measures. The IAT-Conscientiousness consisted of two general categories. The attribute category contained conscientiousness adjectives selected from Saucier’s (1994) Mini-marker adjective checklist. These items have valences that are both positive (i.e., efficient, organized, prompt, systematic, careful, practical, thorough, and steady) and negative (i.e., inefficient, disorganized, unsystematic, careless, impractical, undependable, sloppy, and inconsistent). The second category, called the concept category, was based upon previous studies using the IAT (Greenwald et al., 2002; Egloff & Schmukle, 2002). More precisely, the stimuli represented either the idiographic self (i.e., I, self, my, me, and own) and other items (i.e., they, them, your, you, others). These two distinct categories are combined to create ‘sub-tests,’ commonly referred to as Word Blocks (WB), to assess the relative strength of associations (see Table 2). The reliability of participant’s response latencies on the four WB tasks was assessed and found to be acceptable. Specifically, each participant’s average response latencies were correlated between the congruent (WB’s 3 & 4)($r = .72, p < .01$) and incongruent tasks (WB’s 6 & 7)($r = .74, p < .01$).

Criterion. Participants completed the timed d2 Test of Attention (Brickenkamp & Zillmer, 1998) in order to objectively assess spontaneous indicators of conscientious behavior. Participants were given 20 seconds to work on each of the 14 lines that contained 47 stimuli. This task requires individuals to discriminate between the letter “d” or “p” accompanied with up to four dashes above or below, thus creating 16 combinations of similar stimuli. More precisely, participants were instructed to cross out the letter “d” only when it was accompanied by two dashes (i.e., two above, two below, or one above and one below) while ignoring all other variations. Two aspects of task performance were of interest, concentration performance (CP) and error percentage (E%), because they are related to the accuracy, quality, and carefulness of
participants’ work and internally consistent ($\alpha = .96$ and $\alpha = .79$, respectively) (Brickenkamp & Zillmer, 1998).

Participants were also asked to self-report objective correlates of conscientiousness. Specifically, participants were asked several questions to assess their level of academic achievement, such as the last score on their psychology exam, their current GPA, and their score(s) on the ACT and/or the SAT. In addition, participants were also asked to report how often they skipped class per month, how often they watched television per week, and whether they were employed.

Procedures

Participants arrived at their scheduled times in a reserved computer lab, in groups ranging from 2 to 10, and were asked to complete all of the implicit and explicit measures as part of their course requirement. Specifically, the experimenter first instructed participants on how to complete the d2 Test of Attention and answered any questions before proceeding. Participants then began the timed test (20 seconds for each of the 14 lines) by placing a single line through the appropriate stimuli. Next, the experimenter instructed participants on how to complete the IAT and again answered any questions before proceeding. Finally, the experimenter instructed participants to begin the self-report survey once they had finished the IAT and thanked them for their participation.

IAT Procedure. Instructions for completing the IAT were included for each of the seven blocks and the experimenter strongly urged participants to read each set of instructions because their task on each block would be different. On each trial, a randomly chosen stimulus was presented on the center screen in black letters with a white background. Participants were instructed to determine which category the stimuli belonged to and respond accordingly by
pressing the appropriate key (see Table 3). When participants assigned the stimulus to the incorrect category, a red ‘X’ appeared and participants had to respond appropriately before proceeding. Two versions of the IAT were created to counterbalance the target stimuli to determine whether there were any order effects (i.e., half of the participants began the IAT on word block 5).

**IAT Data Analysis**

The analysis of the implicit measure of conscientiousness was calculated in accordance with the improved scoring algorithm for the IAT (Greenwald, et al., 2003) because this algorithm has been shown to be consistently superior to alternative algorithms. More specifically, trials with reaction times below 400 milliseconds (ms) and above 10,000 ms were eliminated. For each participant, a mean reaction time for each Word Block (WB) 3, 4, 6, 7 was calculated for only correct responses. Next, the pooled $SD$ for the correct responses in WB3 & WB6 and WB4 & WB7 was calculated for each participant. Incorrect responses were replaced with the mean reaction time for the associated block plus a 600 ms penalty. Then, two difference scores were calculated (i.e., WB 6 - WB 3 and WB 7 - WB 4), using the average of both the correct and (recalculated) incorrect responses for each participant. Finally, these difference scores were divided by their associated pooled $SD$ and the remaining quotients were averaged to obtain the IAT effect.
CHAPTER III: RESULTS

Descriptive Statistics

Explicit Measures. Two explicit measures were administered using a five point Likert scale to assess self-reported levels of conscientiousness and motivation. Analyses revealed a mean of 3.53 (SD = 0.59) for the 20-item IPIP scale of conscientiousness (see Table 4). The GOAL-S measure of academic motivation is comprised of 18 items that form 3 subscales of motivation (mastery [M = 4.02, SD = 0.68], performance [M = 2.99, SD = 0.97], and avoidance [M = 1.89, SD = 0.68]).

Criterion from Self-Report Survey. Several questions in the survey were included as indicators of conscientious behavior in an academic context. Specifically, participants reported their scores on the ACT (M = 23.50, SD = 3.63), SAT (M = 1354.82, SD = 318.90), GPA (M = 3.17, SD = 0.52), and the percent correct on their last psychology exam (M = .80, SD = 0.12). Only 11 participants reported their SAT score, therefore this criterion was dropped from the study. In addition, participants were also asked to report how often they skipped class per month (M = 1.98, SD = 2.62), how often they watched television per week (M = 8.47, SD = 7.45), and if they had a job (37%), how many hours per week they worked (M = 16.32, SD = 9.04).

Outlier Analysis

Exploratory analyses were conducted to detect violations of normality and these analyses revealed that multiple measures and criterion were not normally distributed (i.e., Mastery motivation, avoidance motivation, and error percentage rate on the d2 test). In each instance, these outliers were examined and removed to determine their effect on the central purpose of the current study (i.e., the predictive ability of the IAT). Specifically, several iterations of outlier removal were conducted based upon casewise diagnostics and several criteria guidelines for the
removal of outliers (i.e., Standardized residual > 3.3, Externally studentized residual > ± 2.0, Cook’s D > 1.0, and DFFITS > ±1.0) (Cohen, Cohen, West, & Aiken, 2003). In each instance, outlier removal did not significantly change the results; therefore the following results reported are based upon the original untreated data.

**Sex.** An exploratory analysis revealed several significant differences for sex. More specifically, females explicit responses for conscientiousness \( (M = 3.65, SD = 0.57) \) were significantly higher than males \( (M = 3.23, SD = 0.53) \), \( t(82) = -2.81, p < .01 \) (two-tailed). Females self-reported GPA \( (M = 3.25, SD = 0.48) \) was significantly higher than males \( (M = 2.98, SD = 0.53) \), \( t(72) = -2.19, p < .05 \), (two-tailed). Females also demonstrated significantly higher error percentage rates on the d2 test of attention \( (M = 4.08, SD = 3.49) \) than did males \( (M = 2.62, SD = 2.08) \), \( t(77.51) = -2.39, p < .05 \) (two-tailed); and a Levene’s test indicated an inequality in the variances between females and males \( (F = 8.46, p < .01) \). Similarly, a Levene’s test for the hours spent watching television revealed a significant difference in variance \( (F = 19.66, p < .01) \) between females \( (M = 6.92, SD = 5.21) \) and males \( (M = 11.06, SD = 10.42) \); however the t-test with equal variances not assumed was not statistically significant, \( t(31.17) = 1.92, p > .05 \). The means and variances for the remaining measures and criterion were not statistically significant for sex.

**Ethnicity.** Caucasians comprised nearly 74% of the participants \( (n = 67) \), thus minority participants were collapsed into a single group for analyses. An independent samples t-test revealed a significant difference between Caucasians and minorities for ACT and error percentage on the d2 test. Specifically, analysis of the self-reported ACT scores revealed that an inequality in the variance between Caucasians and minorities \( (F = 5.77, p < .05) \) and the t-test revealed that Caucasians ACT scores \( (M = 23.78, SD = 3.63) \) were significantly higher than
minorities ($M = 21.00, SD = 2.16$), $t (29.48) = 3.63, p < .01$ (two-tailed). An analysis of the error percentage on the d2 test revealed that Caucasians error percentage rates ($M = 3.24, SD = 3.10$) were significantly lower than minorities ($M = 5.34, SD = 3.47$), $t (82) = -2.43, p < .05$ (two-tailed). The means and variance for the remaining measures and criterion were not statistically significant for ethnicity.

**Order Effect of IAT.** As previously mentioned, two versions of the IAT for conscientiousness were created to test whether there was an order effect present for the IAT stimulus categories. Specifically, version 1 contained the Self/Positive combination first whereas version 2 contained the Self/Negative combination first. An independent samples t-test revealed that there was a significant difference in the IAT effect for the two versions, $t (88) = 2.30, p < .05$ (two-tailed) ($M = 0.55, SD = 0.31$ for version 1 and $M = 0.39, SD = 0.34$ for version 2). Therefore, the scores from the two versions were standardized so that analyses could be conducted for the following hypotheses.

**Hypotheses.** Significance tests for *a priori* predictions were performed with one-tailed tests.

Hypothesis 1 states that the IAT for conscientiousness will significantly predict spontaneous behavior (i.e., performance on the d2 test of attention) better than the self-report measure of conscientiousness. Neither the implicit nor explicit measures of conscientiousness were significantly related to performance on the d2 test of attention ($p > .05$). Specifically, for both of the d2 performance indicators, *concentration performance* and *error percentage*, the correlations with the standardized IAT scores were $r = 0.06$ and $r = -0.03$, respectively; the explicit test correlations were $r = 0.10$ and $r = -0.05$, respectively.
Hypothesis 2 predicted that the correlation between implicit and explicit measures of conscientiousness would be positive and significant. Although the correlation was positive, it failed to reach a level of statistical significance ($r = 0.10, p > .05$).

In light of the unexpected findings that the IAT measure did not significantly correlate with any dependent variable, it was inappropriate to use multiple regression to test hypothesis 3. This hypothesis stated that the IAT for conscientiousness would provide significant predictive variance above and beyond the explicit measure of conscientiousness for the dependent variables.

Hypothesis 4a was proposed to validate that the IAT for conscientiousness is tapping the same construct that the explicit measure taps (i.e., academic performance indicators, such as: GPA, last exam grade, and ACT score). The results did not support the hypothesis that the IAT for conscientiousness was significantly related to these indicators of academic performance. The following are the nonsignificant ($p > .05$) correlations for the IAT scores and these academic performance indicators: $r = 0.12$ for GPA, $r = -0.06$ for last exam, and $r = 0.14$ for ACT. However, the explicit measure of conscientiousness was significantly related to GPA ($r = 0.37, p < .01$) and last exam grade ($r = 0.21, p < .05$), but failed to reach significance for ACT score ($r = .13, p > .05$). Finally, participants were also asked to indicate on average how often they skipped class per month. The analysis revealed similar findings as above; the IAT was not significantly related to average days skipped per month ($r = -0.07, p > .05$), however the explicit measure significantly predicted how often participants reported that they skipped class per month ($r = -0.37, p < .01$).

Hypothesis 4b was meant as a stronger test of the validity of the IAT for conscientiousness. This hypothesis proposed that IAT scores would be related to an academic
motivational scale comprised of three dimensions related to academic performance (i.e., mastery learning, performance approach, and performance avoidance). Again, the results did not support the hypothesis that the implicit measure of conscientiousness was significantly correlated with these dimensions of academic performance (mastery learning: $r = -0.02, p > .05$; performance approach: $r = -0.002, p > .05$; and performance avoidance: $r = -0.01, p > .05$). However, the explicit measure of conscientiousness was significantly related to 2 out of 3 of these dimensions. Specifically, the correlation between explicitly measured conscientiousness was significant for both mastery motivation ($r = 0.42, p < .01$) and avoidance motivation ($r = -0.50, p < .01$), but the correlation for performance motivation was nonsignificant ($r = 0.09, p > .05$).
CHAPTER IV: DISCUSSION

The IAT developed for the current study clearly did not function as expected and a review of the likely causes of these results is warranted. Although there were some problems with outliers and assumptions of normality with the dependent variables, these problems did not contribute to the lack of findings. Therefore, the results from the current study are indicative of substantial problems that require further examination.

The purpose of hypothesis 1 was to demonstrate one important characteristic of the IAT, which is the ability to predict spontaneous behavior better than explicit measures (Asendorpf et al., 2002). Surprisingly, the current results conflict with the only known study that investigated the relationship between an IAT for conscientiousness and spontaneous indicators of conscientious behavior. Recall that Steffens and König (2006), found a significant correlation between the Implicit and Explicit (I-E) measures of conscientiousness and the IAT was significantly correlated with spontaneous behavioral indicators of conscientiousness (i.e., d2 test). However, a thorough review of this study revealed several important methodological discrepancies. Specifically, their study did not counterbalance the IAT word blocks, and order effect was an issue in the current study. In addition, they did not use an error penalty; and they did not randomize stimuli; and finally they calculated the IAT effect by using the mean reaction times between congruent and incongruent word blocks and divided by the overall standard deviation for each participant. These deviations are important for two main reasons. First, the currently recommended algorithm has been suggested after several refinements due to superior performance (Greenwald et al., 2003). Second, research has revealed that deviations from the recommended algorithm can introduce a cognitive skill confound (McFarland & Crouch, 2002). The authors did note that the “IAT effects were computed similar to” the currently recommended
algorithm; however no justification for these deviations were provided. (McFarland & Crouch, 2002, p. 16) Finally, the review also revealed that the authors deviated from the instructions for the d2 test of attention by briefly stating ‘there was no time constraint.’ This discrepancy might explain why the d2 test of attention was not a significant indicator of conscientiousness in the current study because the time constraint likely introduced a ceiling on task variability between participants. Unfortunately, the current author did not immediately detect these deviations until after searching for an explanation for the current results. In light of these discrepancies between the current study and Steffens and König (2006), comparisons with the current study and conclusions regarding the results of these studies should be interpreted with caution.

Hypotheses 2 and 3 were proposed under the assumption that the I-E measures would correlate and that the implicit measure would add unique predictive variance in multiple regression. However, the current results did not detect a significant correlation between these measures and the IAT did not predict any dependent variables. In light of these results, it is not entirely surprising that hypotheses 4a and 4b were also not supported because these hypotheses were proposed as stronger tests that the I-E measures would predict similar criteria (i.e., academic indicators of success and goal orientation). Therefore, a more extensive review of the literature on the IAT is necessary to understand why the current results failed to support these hypotheses and to offer suggestions for future research efforts.

Similar Findings

One unpublished study presented at SIOP (2008) revealed remarkably similar results. In an investigation of the big five personality dimensions, Siers and Christiansen (2008, April), investigated whether an IAT for conscientiousness could be a useful addition for personality measurement. Like the current study, these authors used the IPIP scale as an explicit measure of
conscientiousness, in addition to also collecting two peer ratings for each participant. Similar to the present results, the authors did not detect a significant I-E correlation, whether it was a self-report or peer ratings (i.e., $r = -.01$ and $r = .08$, respectively). Similar results were also found regarding the dependent variables in their study. More specifically, although the explicit measure was significantly related to GPA ($r = 0.38$), the implicit measure did not reach significance ($r = 0.08$). In addition to these similar results, the authors also collected supervisor ratings for each participant, but both the I-E measures of conscientiousness did not significantly predict supervisors’ ratings ($r = 0.12$ and $r = 0.07$, respectively).

*Implicit Versus Explicit Measures*

The IAT has received considerable methodological attention over the last decade; however many questions regarding its psychological nature remain unresolved. A review of implicit measures that focused on the IAT compare the current progress of the IAT to the debates regarding attitudes and behavioral in/consistency (Hofmann, Gschwendner, Nosek, & Schmitt, 2005b) compare the current progress of the IAT to the debates regarding attitudes and behavioral in/consistency. The first set of debates is similar to the accumulating evidence from various IAT studies that have focused on first demonstrating whether (and how much) a relationship exists with meaningful criterion. Since the inception of the IAT 11 years ago, there has been substantial empirical progress that suggests that the IAT can be successfully applied to a variety of domains; however, many of these experts acknowledge that more work needs to be conducted regarding the IAT’s utility for personality assessment. Hence, one of the primary goals of this study was to determine whether an IAT for conscientiousness would correlate with an explicit measure (hypothesis 2) and whether these two forms of measurement could be used to add unique predictive variance when entered into multiple regression for several criteria (hypothesis
3). Similarly, hypotheses 4a and 4b were included to test whether the IAT could predict distal criteria (e.g., GPA) and motivational approaches related to academic success that have been shown to correlate with explicit measures of conscientiousness. Unexpectedly these hypotheses were not supported, thus begging the question “Why did the implicit and explicit measures not correlate?” In order to attempt to answer this question we must return to the ‘first set of debates;’ however, it is also important to remain aware that very few IAT studies have involved personality assessment.

**Progression of Implicit Measures**

The fact that the implicit measure did not support the hypotheses forces one to consider, “What, if anything, did the IAT for conscientiousness assess?” This is not a new question in IAT research and a review of the IAT’s progression will help shed light on the potential utility of the IAT as a personality measure. The last decade of research on the IAT has revealed that there is substantial variability between I-E measures (Hofmann et al., 2005a). Early research placed the latent construct(s) at the center stage of this debate and two explanations have been offered for the large variability in I-E correspondence (Hofmann, et al., 2005b). The main question is whether implicit and explicit measures assess ‘independent aspects of the same construct’ or whether these measures assess ‘completely unrelated and independent constructs.’ On the one hand, early research led some experts to believe that implicit and explicit measures assess independent aspects of the same construct (Devine, 1989; Dovidio, Kawakami, Johnson, & Johnson, 1997; Greenwald & Banaji, 1995). However, it is important to recognize that these researchers did not rule out the possibility of I-E correspondence and that this position is theoretically in keeping with recent research regarding dual attitudes (Wilson, Lindsey, & Schooler, 2000). On the other hand, some researchers have suggested that I-E measures assess a
singular construct and that the variability in I-E correspondence is due to reflective processes and measurement error (Fazio & Olson, 2003). The latter position asserts that implicit measures are designed to measure representations beyond conscious control; explicit measures will deviate from this point to the extent that they are altered through controlled processes. Although there is no definitive answer on the nature of the latent construct, a better understanding of I-E correspondence can be gathered from empirical studies that have tested whether the latter account holds as an explanation for the observed I-E variability.

If I-E measures assess completely unrelated and independent constructs, it would still be possible to have significant, but spurious correlations, given the number of IAT studies that have been conducted. However, for this explanation to be true one would also expect the correlations across studies to randomly vary and be near zero. To investigate this claim, researchers have employed meta-analyses and MTMM to investigate the relation between I-E measures in the hopes of offering a better understanding that will help resolve the debates that have arisen from conflicting empirical studies of the IAT (Hofmann, et al., 2005a; Nosek, 2005; Nosek & Smyth, 2007; Poehlman, et al., 2005). These studies have revealed that I-E measures are indeed systematically related, with a mean corrected correlation of $r = .24$ across the studies included in the meta-analysis (Hofmann, et al., 2005a). In addition, Nosek and Smyth (2007) conducted an extensive MTMM investigation that found evidence of convergent and divergent validity and concluded that their findings support the assertion that these measures assess related but distinct aspects of the same construct that are distinguishable from systematic measurement error. In light of these findings, Nosek and Smyth (2007) argued that the extremes between these two positions are not useful (i.e., implicit and explicit representations are either completely
independent constructs or that they comprise one construct and the variability between these measures is due to reflective processes and measurement error).

These results have begun to shed light on the nature of the construct measured by I-E tests, but it is important to remember that this remains a work in progress. Although evidence suggests that implicit and explicit measures assess related but distinct constructs, Nosek and Smyth (2007) acknowledge single attitudes may exist for certain constructs (as suggested by Fazio & Olson, 2003), but they note that any domain that forms a single factor would be the exception to the variety of domains that do not. Regarding the current study, it is crucial to remember that the constructs of interest for many of the IAT studies (e.g., flowers-insects, Democrats-Republicans, straight-gay, etc) upon which these results are based are different from the current study. In fact, IAT researchers often acknowledge that other areas (e.g., personality) are underrepresented and require empirical testing (e.g., Nosek, 2005). To this author’s knowledge the only direct test of the relationship between I-E measures of conscientiousness is the study previously mentioned (Siers and Christiansen, 2008, April), in which the authors concluded that there is little overlap between the I-E measures of conscientiousness. Thus, despite strong evidence from meta-analyses and MTMM investigations across domains that suggest that I-E measures are generally related but distinct constructs, more research is required to understand how these findings relate to personality and in particular conscientiousness.

Implicit-Explicit Correspondence and Possible Moderators

Although meta-analysis is a relatively objective means that is useful for interpreting conflicting findings from the numerous IAT studies, the authors of a recent IAT meta-analysis note that it does not capture the complexity of the variability in I-E correspondence (Hofmann et al., 2005a). They suggest that a more promising approach that is emerging is to identify
moderator variables that influence when I-E correlations are high or low (Hofmann et al., 2005b; Nosek, 2005). Some researchers have compared this effort to the second set of debates regarding the attitude and behavioral in/consistency debates that shifted attention, from whether, to when a relationship exists (Hofmann et al., 2005b).

One factor affecting I-E correspondence that has been suggested is the principle of correspondence (Ajzen & Fishbein, 1977; Hofmann et al., 2005b). This principle states that correspondence will be higher when two measures are comparable in content and specificity and was originally proposed as an explanation of the attitude and behavior in/consistency phenomenon in a review of over 100 studies (Ajzen and Fishbein, 1977). They found that when attitudinal and behavioral measures were judged to be similar they often obtained correlations above .40; however, when these measures were judged to be low in similarity they found that the results often were nonsignificant. At this point, it is helpful to recall that both the current study and the study by Siers and Christensen (2008, April) used propositional statements as the explicit measure and neither study found a significant I-E correlation. Perhaps if some other type of measure more consistent with the IAT were used, results would have been more favorable.

*Translation.* Evidence suggests that associative (i.e., implicit) and propositional (i.e., explicit) representations are distinct, but related constructs (Hofmann, et al., 2005a; Nosek, 2005; Nosek & Smyth, 2007). Researchers have also suggested that this relationship is bi-directional (Hofmann, et al., 2005b). In other words, associations that are automatically activated can form the ‘building blocks’ of our propositional thinking and propositional thinking can influence the associative representations in long-term memory. This is what researchers have referred to as translation (Hofmann, et al. 2005b). There are several factors that can influence this potential moderator.
Explicit Measurement. One factor that can affect I-E correspondence concerns the effort an explicit item requires for a response (i.e., adjectives versus sentences). This is because the associative representations that the IAT is designed to measure are just one source of information that form the base of a propositional evaluation (Hofmann, et al., 2005b); and more effort requires engaging in reflective thinking; and a major characteristic of the reflective system is to generate and transform information (Bargh, 1994; Nosek, 2005). In other words, I-E correspondence is likely to decrease as individuals engage in deliberative thinking that intentionally incorporates other sources of information, such as, autobiographical memory (Koole, Dijksterhuis, & Knippenberg, 2001; Nosek, 2005; Fazio & Olson, 2003). Thus, one could argue that a semantic differential scale is preferable to assess I-E correspondence because, unlike items that contain propositional sentences (i.e., the IPIP scale used in this study); responses to semantic differential scales (i.e., adjectives) encourages more spontaneous judgments (Hoffman et al., 2005a). This assumption received considerable support in a meta-analysis that tested several types of explicit response formats (Hoffman et al, 2005a). Their results revealed that I-E correspondence was higher for explicit items that were judged to be more ‘gut-level’ reactions compared to those items that required more effortful and self-reflective thought processes.

Representation Strength. Strong representations show higher I-E consistency because they are more practiced, persistent, and less susceptible to change (Hofmann, et al., 2005b). In an investigation designed to test the importance of representation strength, Nosek (2005), found that a strength factor (comprised of attitude importance, thought frequency, and familiarity) moderated I-E correspondence across 57 attitude domains. However, the extent to which a strength factor mediates an IAT designed to assess personality remains an empirical question.
Implicit Self-Observation. Nosek (2005) note that a crucial determinant of I-E correspondence is likely to be the quality of self-observation, rather than introspective access to implicit representations, because by their very nature implicit representations are less accessible to introspection (Hofmann et al., 2005b; Greenwald and Banaji, 1995). This observation follows self-perception theory (Bem, 1972), which argues that the quality of self-observations will vary depending on whether they can be characterized as consistent, practiced, clear, and distinct or whether they are inconsistent, novel, ambiguous, and normative. This raises the possibility that implicit personality measurement could be more contextually contingent than other common constructs in the IAT literature (e.g., stereotypes) because it is less organized and stable. In other words, implicit personality measures in general, and conscientiousness in particular, may lack correspondence with explicit personality measures because individuals’ self-observations of their personality are inconsistent across situations.

Dimensionality and Personality. Dimensionality refers to whether a representation is conceived of as bipolar or unipolar (Hofmann, et al., 2005b). Research has shown that bipolar representations (e.g., being pro-life implies you are against pro-choice) are judged more quickly and better recalled (Bargh, 1994) suggesting that these representations simplify and organize information which leads to more efficient and consistent representations, whereas unipolar representations are more complex, less stable and therefore are less reliable and more likely to be influenced by situational factors. Again, although traditional conceptualizations of personality seem to reflect a unipolar conceptualization of traits (i.e., your are either efficient or inefficient), there is evidence that indicates that these traits are more contextually contingent than originally conceived and that this systematic variance is usually described away as error (e.g., Mishcel, 1968). One simple illustration of this point is the increase in predictive ability of measures that
add the words ‘…at work’ to the end of the explicit items (Schmit, Ryan, Stierwalt, & Powell, 1995); Robie, Schmit, Ryan, & Zickar, 2000)

Manipulating the Situation. Even though the IAT is designed to assess relatively stable associations, it is important to remember that, like explicit measures, evidence has shown that they can also be affected by situational influences that can make these associations more or less accessible (Blair, 2002). For example, researchers have shown that the IAT effect commonly found in investigations of stereotypes can be reduced when the exemplars (e.g., famous African Americans and infamous European Americans) are made salient before the IAT (Dasgupta & Greenwald, 2001). Importantly, this manipulation did not alter explicit responses. Similarly, Greenwald et al. (2002) proposed that representations of the self and various associations linked to the self comprise a Social Knowledge Structure (SKS); and that the IAT is designed to measure the strength of these associations. In a test of this assumption, researchers have found evidence across several methods and domains that the activation of individuals’ SKS prior to the IAT task can increase the predictive ability of relevant behaviors (Perugini, O’Gorman, & Prestwich, 2007). The authors suggested that this manipulation primarily influences the salience of associative evaluations. Again, it is worth noting that like the previously mentioned example (i.e., manipulating exemplars) this manipulation also did not affect the explicit representations as evidenced by self-reports.

There is also evidence that the provision of feedback for incorrect IAT responses alters the salience of relevant associations and therefore the response latencies that form the basis of the IAT measure (Olson & Fazio, 2004). This could be of particular importance for an IAT designed to assess conscientiousness because of the theoretical relationship between conscientiousness and the provision of negative feedback for incorrect responses. This
situational influence could have altered the salience of participants’ representations and may help explain the null results obtained in present study and nearly identical results of a similar study (Siers & Christiansen, 2008, April). Unfortunately, it is unclear whether the study that demonstrated a significant relationship between IAT scores and spontaneous behavior (Steffens and König, 2006) provided task feedback; however the authors did note that they did not use an error penalty in calculating the IAT effect (as previously mentioned). Nevertheless, the evidence suggests at least two ways that situational influences could have altered the salience of participants’ associations and thus decreased the predictive ability of the IAT for this study. Therefore, future studies should consider: 1) Employing similar self-activation manipulations before the IAT (e.g., activating an individuals SKS and/or activating relevant exemplars) and 2) purposefully deviating from the IAT recommendations by eliminating feedback for incorrect responses for theoretical reasons.

_Future Research._ Current understanding of the importance of representation strength as a moderator suggests several explanations that may have been responsible for the poor performance of the IAT. Recall that the stimulus items chosen for the IAT were selected because they are empirically validated items that reflected the explicit construct of conscientiousness. This investigation has highlighted research that suggests that I-E measures are distinct but related constructs and offered suggestions for future investigations to consider when choosing items for both measures (e.g., the extent to which they are ‘clear, consistent, practiced, and distinct’ and their ‘subjective importance, thought frequency, and familiarity’). In addition, evidence exists that certain personality traits may vary across time and context (Mischel, 1968) and future research may benefit from situational manipulations that can affect the predictive ability of the IAT for personality studies.
Concluding Remarks. A consensus on the precise nature of the underlying construct(s) of I-E personality (e.g., origin, manifestation and malleability) has remained elusive to researchers. Even the now well accepted explicit conceptualization of personality has had to overcome fundamental challenges to assumptions that led many experts to question its relevance; explicit personality research remains a construct that merits scientific attention because it is able to predict important psychological phenomenon. Similarly, early research investigating implicit measures (e.g., Rorschach tests, Thematic Apperception Tests, etc) were challenged to their core (and rightfully so because of their psychometric properties) and even the most novice of psychologists have probably shelved these tests for eternity. Despite sharing some superficial similarities with these measures, the IAT in general is accumulating evidence that given the right conditions this new measure can reliably predict meaningful criteria in novel ways. A brief examination of the literature clearly suggests that the IAT has generated a wave of exciting new possibilities in the scientific community after just over one decade of results. However, with excitement comes criticism, but interestingly these criticisms have often resulted in the refinement of our understanding of this measures nature, capabilities, and limitations. Although nonsignificant results are always disappointing, there remains the consolation prize of a fresh understanding for the probable causes and a new appreciation of how to proceed.
REFERENCES


Gottingen: Hogrefe.


Table 1

*Attribute and Concept Stimuli for IAT*

<table>
<thead>
<tr>
<th>Attribute and Concept Labels</th>
<th>Positive</th>
<th>Negative</th>
<th>Self</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient</td>
<td>Inefficient</td>
<td>I</td>
<td>They</td>
<td></td>
</tr>
<tr>
<td>Organized</td>
<td>Disorganized</td>
<td>Self</td>
<td>Them</td>
<td></td>
</tr>
<tr>
<td>Prompt</td>
<td>Inconsistent</td>
<td>My</td>
<td>Your</td>
<td></td>
</tr>
<tr>
<td>Systematic</td>
<td>Unsystematic</td>
<td>Me</td>
<td>You</td>
<td></td>
</tr>
<tr>
<td>Careful</td>
<td>Careless</td>
<td>Own</td>
<td>Others</td>
<td></td>
</tr>
<tr>
<td>Practical</td>
<td>Impractical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thorough</td>
<td>Undependable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steady</td>
<td>Sloppy</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Participants assigned the stimuli to the appropriate label by pressing either the ‘i’ or ‘e’ key. The attribute and concept labels remained visible for duration of each word block while individual stimuli appeared in the middle of the screen after each correct response.
Table 2

*Conceptual Example of IAT Task*

<table>
<thead>
<tr>
<th>Word block</th>
<th>Discrimination Task</th>
<th>Trials</th>
<th>‘i’ key</th>
<th>‘e’ key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Concept</td>
<td>20</td>
<td>Self</td>
<td>Other</td>
</tr>
<tr>
<td>2</td>
<td>Attribute</td>
<td>20</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>3</td>
<td>Combined</td>
<td>20</td>
<td>Self + Positive</td>
<td>Other + Negative</td>
</tr>
<tr>
<td>4</td>
<td>Combined</td>
<td>40</td>
<td>Self + Positive</td>
<td>Other + Negative</td>
</tr>
<tr>
<td>5</td>
<td>Concept (reversed)</td>
<td>40</td>
<td>Other</td>
<td>Self</td>
</tr>
<tr>
<td>6</td>
<td>Combined</td>
<td>20</td>
<td>Other + Positive</td>
<td>Self + Negative</td>
</tr>
<tr>
<td>7</td>
<td>Combined</td>
<td>40</td>
<td>Other + Positive</td>
<td>Self + Negative</td>
</tr>
</tbody>
</table>

*Note.* The order of the word blocks for version 2 was as follows: 5, 2, 6, 7, 1, 3, and 4.
Table 3

**Detailed Example of IAT Task**

<table>
<thead>
<tr>
<th>Word block</th>
<th>Discrimination Task</th>
<th>Left Label (‘i’ key)</th>
<th>Right Label (‘e’ key)</th>
<th>‘i’ Sample Stimuli</th>
<th>‘e’ Sample Stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Concept</td>
<td>Self</td>
<td>Other</td>
<td></td>
<td>They</td>
</tr>
<tr>
<td>2</td>
<td>Attribute</td>
<td>Positive</td>
<td>Negative</td>
<td></td>
<td>Prompt Sloppy</td>
</tr>
<tr>
<td>3</td>
<td>Combined</td>
<td>Self or</td>
<td>Other or</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Positive or</td>
<td>Negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Combined</td>
<td>Self or</td>
<td>Other or</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Positive or</td>
<td>Negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Concept</td>
<td>Other</td>
<td>Self</td>
<td></td>
<td>Them Me</td>
</tr>
<tr>
<td>6</td>
<td>Combined</td>
<td>Other or</td>
<td>Self or</td>
<td></td>
<td>Your Organized Me Sloppy</td>
</tr>
<tr>
<td>7</td>
<td>Combined</td>
<td>Other or</td>
<td>Self or</td>
<td></td>
<td>Others Systematic Self Impractical</td>
</tr>
</tbody>
</table>

*Note.* The order of the word blocks for version 2 was as follows: 5, 2, 6, 7, 1, 3, and 4.
Table 4

*Descriptive Statistics and Zero-Order Correlations*

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAT(z)</td>
<td>0.0</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPIP</td>
<td>3.53</td>
<td>0.59</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mastery</td>
<td>4.02</td>
<td>0.68</td>
<td>-0.02</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>2.99</td>
<td>0.97</td>
<td>-0.002</td>
<td>0.09</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td>1.89</td>
<td>0.68</td>
<td>-0.01</td>
<td>-0.50</td>
<td>-0.59</td>
<td>-0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>191</td>
<td>31.9</td>
<td>0.06</td>
<td>0.10</td>
<td>-0.06</td>
<td>-0.01</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E%</td>
<td>3.69</td>
<td>3.32</td>
<td>-0.03</td>
<td>-0.05</td>
<td>0.11</td>
<td>0.11</td>
<td>-0.07</td>
<td>-0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>3.17</td>
<td>0.52</td>
<td>0.12</td>
<td>0.37</td>
<td>0.20</td>
<td>-0.29</td>
<td>0.10</td>
<td>-0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>23.5</td>
<td>3.63</td>
<td>0.14</td>
<td>0.13</td>
<td>0.20</td>
<td>0.26</td>
<td>-0.21</td>
<td>0.21</td>
<td>-0.29</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Last Exam</td>
<td>0.80</td>
<td>0.12</td>
<td>-0.06</td>
<td>0.21</td>
<td>0.09</td>
<td>0.06</td>
<td>-0.18</td>
<td>0.20</td>
<td>-0.09</td>
<td>0.52</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skip</td>
<td>1.98</td>
<td>2.62</td>
<td>-0.07</td>
<td>-0.37</td>
<td>-0.07</td>
<td>0.08</td>
<td>0.20</td>
<td>0.08</td>
<td>0.10</td>
<td>-0.48</td>
<td>-0.02</td>
<td>-0.02</td>
<td></td>
</tr>
</tbody>
</table>

**Bold** = \( p \leq .05 \) level, one-tailed. **Bold Italicized** = \( p \leq .01 \) level, one-tailed