THE REFINEMENT OF A MULTIDIMENSIONAL COMPUTER BASED IMPLICIT ASSOCIATION TEST AS A MEASUREMENT OF ATTITUDES TOWARD PERSONS WITH DISABILITIES

Except where reference is made to the works of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

_____

E. Daly Vaughn

Certificate of Approval:

_____	_____
Daniel J. Svyantek Professor Psychology	Adrian Thomas, Chair Associate Professor Psychology
_____	_____
Bryan Edwards Assistant Professor Psychology	George T. Flowers Dean Graduate School

THE REFINEMENT OF A MULTIDIMENSIONAL COMPUTER BASED IMPLICIT ASSOCIATION TEST AS A MEASUREMENT OF ATTITUDES TOWARD PERSONS WITH DISABILITIES

E. Daly Vaughn

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Master of Science

Auburn, Alabama
August 10, 2009

THE REFINEMENT OF A MULTIDIMENSIONAL COMPUTER BASED IMPLICIT ASSOCIATION TEST AS A MEASUREMENT OF ATTITUDES TOWARD PERSONS WITH DISABILITIES

E. Daly Vaughn

_____
Signature of Author

_____
Date of Graduation

THESIS ABSTRACT

THE REFINEMENT OF A MULTIDIMENSIONAL COMPUTER BASED IMPLICIT ASSOCIATION TEST AS A MEASUREMENT OF ATTITUDES TOWARD PERSONS WITH DISABILITIES

E. Daly Vaughn

Master of Science, August 10, 2009
(B.A., Baylor University, 2005)

152 Typed Pages

Directed by Adrian Thomas

Currently, persons with disabilities (PWDs) account for approximately one-fifth of the adult US population (Steinmetz, 2006). Legislation has been passed to help protect PWDs (e.g., the Americans with Disabilities Act). However, disparity still exists within the workforce between persons with and without disabilities in percentage of persons employed (Taylor, 2000) and pay once employed (Steinmetz, 2006). Attitude research offers one approach to better understood and counteract barriers in the workforce for PWDs. The current study extended previous research investigating the utility of the Multiple Disability Implicit Association Test (MDIAT; Thomas, Doyle, & Vaughn, 2007). The MDIAT uses separate IAT administrations for four disability groups (i.e., persons with and without: cancer, paraplegia, mental illness, and alcoholism).

Two-hundred forty nine undergraduate participants completed the study. This multidimensional IAT was modified and the psychometric properties were investigated. The measure was scored using an updated scoring algorithm (Greenwald, Nosek, & Banaji, 2003) believed to maximize validity of the IAT measurement and minimize measurement error.

Results of the study indicated implicit bias was found for people with a disability versus people without a disability for all four of the disabilities. However, the MDIAT was susceptible to a practice or experience effect whereas individual IAT administration scores tended to decrease across successive IAT administrations. No relationship existed between the MDIAT scores and a commonly used and validated self-report measure, the Interaction with Disabled Persons (IDP) scale. In addition, the MDIAT was not found to be related to general cognitive ability, socially desirable responding, or the interaction between the two. The IDP was susceptible to socially desirable responding but not to cognitive ability.

The study provides mixed evidence concerning the potential utility for the MDIAT in research and practice. Although the MDIAT was not susceptible to socially desirable responding whereas the self report measure was, the existence of an experience or practice effect provides a threat to the usefulness or interpretability of the scores. Caution should be taken when using the IAT technique in training or work settings as previous experience or exposure to the procedure may falsely indicate less implicit bias. Future research should investigate the predictive validity of the MDIAT.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the members of my thesis committee. First, I would like to thank Dr. Adrian Thomas for assisting with design and analysis issues as well as assisting with conceptual considerations regarding aspects of the IAT procedure. I also want to thank Dr. Dan Svyantek and Dr. Bryan Edwards for their willingness to review the project and provide critical feedback.

Thanks are also due to the research assistants which have greatly assisted in this endeavor through helping collect and enter data for the current investigation. Lastly, I want to express my extreme gratitude to my family and close friends for their endless support and encouragement.

Style manual or journal used:

Publication Manual of the American Psychological Association, Fifth Edition

Computer software used:

Microsoft Excel 2007
Microsoft Word 2007
Inquisit 3.0.1.1
Statistical Package for the Social Sciences 16.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Traditionally, the majority of workplace discrimination research has focused on the areas of race, ethnicity, or gender. These forms of discrimination have received a great deal of attention from lawmakers, researchers, human resource managers, civil rights activists, and the public. This continued attention is justifiable, as researchers have supported that these forms of discrimination are still evident in the workplace today (see, for example, Coach & Daly, 2002; Wittenbrink, Judd, & Park, 1997; Budig, 2002; Schneider, Swan, & Fitzgerald, 1997). In addition to these commonly studied target groups, researchers are beginning to explore other groups experiencing potential intolerance. Evidence also supports the notion that discriminatory practices exist in the present day based on sexual orientation (e.g. Clain & Leppel, 2001; Hebl, Foster, Mannix, & Dovidio, 2002), age (e.g. Finkelstein & Burke, 1998; Perry & Finkelstein, 1999), level of physical attractiveness (and/or unattractiveness) (e.g. Hosoda, Stone-Romero, & Coats, 2003; Stone, Stone, & Dipboye, 1992), and disability status (e.g. Ravaud, Madiot, & Ville, 1992; Stone et al., 1992).

Researchers are making progress in each of these areas resulting in positive changes in the workplace. However, it is important to bring further understanding of the attitudes contributing to these forms of discrimination to the public, particularly to those that affect the composition of organizations. Furthermore, discrimination researchers need practical, psychometrically sound measures to assess and provide insight into the attitudes toward the lesser studied or only more recently studied groups (Antonak &

Livneh, 2000). Studying attitudes toward persons with a disability is important as persons with a disability make up an increasing proportion of the population in the United States. In fact, the United States Census Bureau recently reported that 18.1 percent of persons in the United States report having some form of disability and 11.5 percent report having a severe disability (Steinmetz, 2006).

The present study explored the validity and practicality of refining a measure of attitudes toward persons with disabilities (PWDs). More specifically, the measure investigated assesses *implicit* attitudes toward PWDs. Before introducing the methodology used to refine and investigate the measure under examination, a broad review of relevant background information is provided. Current employment legislation in place to protect PWDs from discrimination in the workplace is reviewed. Then, the role of attitude measurement and a review of the various measures and techniques employed for assessing attitudes toward PWDs is provided. Specific aspects and technical issues related to the measurement technique employed in the current study are then presented. Then the design used in the current study is provided, followed by a section detailing the results of the current study. Lastly, conclusions are drawn and the implications and limitations of the current study are discussed.

Legislation in Place to Protect PWDs

Currently, the most broadly applied protective policies and procedures for employees with disabilities in the United States includes the Rehabilitation Act of 1973 and the Americans with Disabilities Act (ADA; 1990). The Rehabilitation Act of 1973 requires federal contractors and subcontractors to seek out qualified PWDs (Cascio & Aguinis, 2005). The Rehabilitation Act of 1973 is limited by the fact that it only applies

to organizations that are considered federal contractors and subcontractors. Thus, many private businesses are not under these regulations. The Office of Federal Contract Compliance Programs (OFCCP) regulates the Rehabilitation Act of 1973. In many ways, the Rehabilitation Act of 1973 merely provided the framework for the ADA (1990).

The Americans with Disabilities Act (ADA; 1990) began to be enforced in organizations in 1992. The first purpose of the ADA was "to provide a clear and comprehensive mandate for the elimination of discrimination against individuals with disabilities." Unfortunately, as Papinchock (2005) evidenced in her review of Supreme Court case litigation related to Title I of the ADA, this initial purpose has yet to be realized. The ADA defined that a qualified person with a disability is one whom can perform the essential functions of the job with or without "reasonable accommodation." The ADA explained that reasonable accommodations are required up to the point of "undue hardship." Action that results in significant difficulty or expense to the company or entity falls under the definition of undue hardship. Other factors considered when determining if accommodations result in undue hardship include an employer's financial resources, size, and the nature and structure of its operations (EEOC ADA handbook, 2005).

One of the criticisms of the ADA is the difficulty in defining what constitutes a person as having a "disability" (Stone & Colella, 1996; Colella & Stone, 2005; Wells, 2001; Papinchock, 2005). Originally, employers were worried about what would be considered a disability and the broad ramifications that the ADA would have toward spending to accomplish reasonable accommodations for employees. To the contrary, Supreme Court litigation previously greatly narrowed and in some respects limited what

qualifies for a disability and what is required by the company for accommodation (Papinchock, 2005; see Sutton v. United Airlines, 1999; Murphy v. UPS, 1999; Albertsons Inc. v. Kirkingburg, 1999). In fact, researchers have found that determination of a lack of a disability is the number one reason that plaintiffs lose their cases (Colella & Stone, 2005; Lee, 1998). Colella and Stone (2005) pointed out that this problem is unique from other types of civil rights legislation. It is much easier for an individual to classify him or herself into a specific gender, race, and so on than in terms of a person with a disability or a person without a disability.

Further complicating the current act, congress recently passed a new amendment to the ADA which increases the coverage for individuals with less severe impairments (i.e., the ADA Amendments Act of 2008 [ADAAA]). The interpretation by the courts will ultimately determine the extent to which the ADAAA of 2008 improves the coverage for persons with disabilities; however, the amendment certainly has the potential to improve working conditions and accommodations for PWDs. The ADAAA began being enforced on January 1, 2009. The ADAAA redefined definitions of PWDs and overturned many of the previous rulings which limited the scope of the ADA's coverage, especially in regard to the use of mitigating measures (e.g., Sutton v. United Airlines, 1999; Murphy v. UPS, 1999).

Current Employment Discrimination toward PWDs

Researchers have found that although legislation has been put in place to protect applicants and employees with disabilities, PWDs are still not receiving equal employment opportunities or equal levels of pay once employed as people without disabilities (Colella & Stone, 2005; Thomas, 2001; Thomas et al., 2007; Wells, 2001).

According to a recent survey conducted by the Harris Poll organization (Taylor, 2000), only 31% of PWDs were employed part or full time (cited in Colella & Stone, 2005). In another Harris Poll survey (Taylor, 1998), 72% of PWDs that were unemployed responded that they wanted to work (cited in Colella & Stone, 2005). Additionally, the U.S. Census Bureau (Steinmetz, 2006) reported that in 2002 persons with severe and non-severe disabilities had lower annual incomes than people with no disability. Persons with a severe disability reported median earnings of $12,800 annually while persons with a non-severe disability earned $22,000 and persons with no disability earned $25,000. In addition, Steinmetz (2006) reported that in 2002 the poverty rate for people 25 to 64 with no disability was 7.7 percent as opposed to 11.2 percent for people with a non-severe disability and 25.9 percent for people with a severe disability.

Other sources have reported similar statistics. For example, the Rehabilitation Research and Training Center on Disability Demographics and Statistics (2005) reported that in 2004, the employment rate for people with a disability was 37.5 percent as opposed to 77.8 percent for people without a disability. In addition to surveys and census data, researchers have found empirical evidence to support that once PWDs secure employment, discriminatory treatment at the workplace potentially leads to self-selecting out, self-consciousness, isolation, anxiety, and depression for PWDs (e.g., Stone et al., 1992; Stone & Colella, 1996). Taken together, the evidence appears to support that although many PWDs want to work, many of these individuals still face difficult barriers to fulfilling and satisfactory employment.

Call for Change

The ADA (1990) has been limited in scope by recent interpretation of the act by the judicial system (Papinchock, 2005). A combination of methods must be utilized to help learn more about the barriers to equal treatment and how to improve the current status of PWDs relative to persons without a disability. One such method is through attitude research. An attitude can be defined as "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly & Chaiken, 1993, p. 1). It is well established that an individual's attitude toward a target object or concept is a possible determinant of behavior toward that object or concept (Brehm, Kassin, & Fein, 2005; Kraus, 1995). Antonak and Livneh (1988) noted that under certain conditions, attitudes might strongly influence behavior. Individual's attitudes and the composite organizational and societal climate that these attitudes produce toward PWDs are one such barrier linked to discriminatory behavior (Antonak & Livneh, 1988; Colella & Stone, 2005).

In order to improve the conditions for PWDs and accomplish a better understanding of the relationship that attitudes may play, Antonak and Livneh have called for "innovative experimental methods and psychometrically sound instruments that are reliable, valid, and multidiminsional (p. 211, 2000)." In an effort to respond to the call of Antonak et al. (2000), this study intends to substantially modify and further investigate the psychometric properties of a previously developed implicit attitude measurement tool, the Multiple Disability Implicit Association Test (MDIAT; Thomas et al., 2007).

Potential Benefits of Establishing a Sound Measure

Establishing a psychometrically sound measure of attitudes toward PWDs, particularly if it could be linked to employee discriminatory behavior, could be beneficial to practitioners and human resource managers for a number of reasons. A valid and reliable measure could be used to evaluate outcomes for a program that emphasizes attitude change or to assist in the needs assessment phase of training by serving as a baseline assessment of current attitudes held by employees. Consultants and practitioners might also use such a measure across a group of employees to assist in understanding the underlying assumptions of an organizational or departmental culture towards PWDs. In the 2004 fiscal year, the EEOC resolved 16,949 cases of disability discrimination and recovered $47.7 million for those making charges and other aggrieved parties not including monetary benefits resulting from litigation (EEOC ADA handbook, 2005). If used effectively and innovatively, such a measure could translate into potential savings by organizations.

Disability researchers would also benefit from the availability of a potentially improved measure of implicit attitudes. Researchers have recognized the need to use measures of attitudes toward PWDs and are using the available measures to investigate important research questions (Antonak et al., 2000). Studies have investigated the underlying constructs that compose these attitudes and behaviors (e.g., Hirschberger, Florian, & Mikulincer, 2005; Thomas, 2001), as well as possible interventions that would diminish or reverse the negative aspects of these attitudes and behaviors (e.g., Fichten, Schipper, & Cutler, 2005). For example, Hirschberger et al. (2005) conducted a study which supported that the human need to manage the terror of death contributed to the

emotional reactions toward PWDs. In addition, Fichten et al. (2005) found that working with children with disabilities diminished social distance from PWDs and improved certain attitudes, thoughts, and feelings. These studies and others (e.g., Loo, 2004) have shown that measures can be utilized to assist in empirically investigating important research questions.

Current Measurement of Attitudes

Although attitude research has a long-standing history, recently attitude measurement techniques have become increasingly complex theoretically (Antonak et al., 2000). The current measures for attitudes toward PWDs are often classified into two broad categories: direct and indirect measures. Before presenting the proposed implicit measure, a brief review will be given of the measures currently available. The review begins with an overview of each technique and is followed by the limitations associated with the technique. To assist in structuring the overview, the grouping of direct measures will be presented and discussed first, followed by the available indirect measures.

Direct Measures

When employing direct measures, the respondent is aware that he or she is participating in an attitude measurement. The direct measurement approaches are still the most widely known and commonly employed technique for attitude measurement toward PWDs (Antonak et al., 2000). Antonak et al. recently provided a summary of the ten most commonly used direct measurement techniques available. The first direct measurement technique reviewed by these authors was opinion surveys (e.g., Caruso & Hodapp, 1988). In an opinion survey, respondents are asked to provide a structured (forced choice) or unstructured response expressing their attitudes toward the referent. A second direct

measurement technique available is interviews (e.g., Salend & Giek, 1988). The researcher or practitioner can opt to utilize an interview with a fixed set of questions and sequence for all respondents (structured) or choose to have the freedom to vary and ask additional questions from interviewee to interviewee (unstructured). A third available direct measure technique is the ranking method (e.g., Abroms & Kodera, 1979). In this technique, the respondent arranges a small set of items into an ordered sequence based on a specified criterion (Antonak et al., 2000). This technique has been used, for example, to rank order various types and levels of a disability (Richardson, Goodman, Hastorf, & Dornbusch, 1961).

A fourth direct measure technique is the Q methodology (Stephenson, 1953; Shaver & Scheibe, 1967). The Q methodology requires the participant to sort a set of statements or phrases into various stacks based on some specified criterion (e.g., favorability, intensity of agreement, descriptiveness). The researcher then analyzes the resulting clusters of statements or phrases and compares them with that obtained by other respondents (e.g., different groups, pre- and post-test design, previous research) (Antonak et al., 2000). A fifth option available using a direct measure method is sociometrics (e.g., MacMillan & Morrison, 1984). Sociometrics are designed to reveal how a participant behaves or would intend to behave in a group setting toward a person with a disability given a choice of behaviors (Antonak et al., 2000). For example, the participant might be given a picture roster of hypothetical coworkers and be asked to whom they would most like to sit next to at work. They might then be asked whom they would like to sit next to second most and so on until they select whom they would least like to sit next to at work. A sixth direct measurement technique used is the Adjective CheckList (ACL; Gough,

1960). Gough originally developed the ACL to assess personality and self-concept, but it has been applied to measure attitudes toward PWDs (e.g., Gottlieb, Corman, & Curci, 1984).

A seventh direct measure technique presented by Antonak et al. (2000) was the method of paired comparisons. Following this technique, the participant would rate all possible pairs of referents presented by the researcher based on some predetermined criterion. The data would be analyzed to produce an ordering based on a participant's or a group of participants' responses. Janicki (1970) used this method to study the attitudes of hospital employees toward 12 different groups of PWDs. The eighth technique reviewed was the semantic differential method (Osgood, Suci, & Tannenbaum, 1957). Using the semantic differential technique, the participant is presented with a single concept followed by 7 to 20 scales anchored at each end by bipolar adjectives and connected by a line that is marked at intervals. The participant is then asked to mark the point representing her or his rating of the concept on the line. A ninth direct measure technique used in measuring attitudes toward PWDs is based on consensual location scaling (Thurstone, 1928). In consensual location scaling an index value is assigned to each item in a group of items characterizing an attitude referent. The researcher then designs a scale using the indexed items believed to create a "continuum of agreement" (p. 215, Antonak et al., 2000). Tringo (1970) used this technique to produce the Disability Social Distance Scale. Based on this technique, Tringo developed the hierarchy of preference for disability groups.

Lastly, the tenth and most commonly used direct measure technique used in attitude research (Antonak et al., 2000, 1988) is rating scales (e.g., Yuker, Block, &

Younng, 1966; Gething, 1994). Often these scales follow a Likert (1932) type rating format. Given that rating scales are such commonly used tools by attitude researchers, it is necessary to provide a more in depth review and discussion of three of the most well known self-report scales used in disabilities research.

The Attitude Toward Disabled Persons Scale. The Attitude Toward Disabled Persons Scale (ATDP) (Yuker, Block, & Campbell, 1960; Yuker, Block, & Younng, 1966) is the most widely used measure in disability research (Antonak et al., 1988, 2000; Pruett & Chan, 2006; White, Gordon, & Jackson, 2006; Thomas et al., 2007). Antonak et al. (1988) reported that three variations of the ATDP were created by Yuker and colleagues: the original 20 item scale, Form O (Yuker et al., 1960), in addition to two other equivalent 30 item variations of the ATDP, Forms A and B (Yuker et al., 1966). The three forms have shown to have fairly stable test-retest and internal reliability (Antonak et al., 1988). The ATDP measures attitudes in terms of perceived differences between people with and without a disability at the group or societal level (Gething, 1994). It asks the respondent to report their level of agreement regarding a series of statements related to the appropriate treatment and characteristics of PWDs in society.

Although the ATDP is still being utilized in the literature (e.g., Alghazo, Dodeen, & Algaryouti, 2003; Feldman, Gordon, White, & Weber, 2002), researchers have criticized the measure for a number of reasons related to the validity of the measure. Several researchers have suggested a need for a more contemporary measure that would be easier to use and also more psychometrically sound (Antonak et al., 1988; Gething, 1994). One specific criticism of the ATDP is its susceptibility to faking (e.g., Cannon &

Szuhay, 1986). All three forms of the ATDP have been found susceptible to socially desirable responding (see Antonak et al., 1988).

Another critique of the ATDP is related to the proposed unidimensional measure of affect (Thomas et al., 2007; Gething, 1994). The problems regarding dimensionality in using the ATDP are two-fold. First, attitude researchers (e.g., Antonak et al., 1988, 2000; Gething, 1994) have requested measurement scales intended to be multidimensional in nature as opposed to the more traditional unidimensional structure described by Yuker and colleagues (Yuker et al., 1960; Yuker et al., 1966). In addition, Antonak et al. (1988) discussed that although the developers of the ATDP have reported that the measure is based on a unidimensional structure (i.e., favorable-unfavorable attitudinal dimension), factor analytic investigation has evidenced two or more dimensions on at least two (Form O and A) of the three forms of the ATDP. Taken together, the criticisms presented warrant consideration of other direct measurement scales as the ATDP has several undesirable characteristics.

The Scale of Attitudes Toward Disabled Persons. In response to several of the previously mentioned criticisms of the ATDP, Antonak (1981, 1982) developed an alternative multidimensional rating scale entitled the Scale of Attitudes Toward Disabled Persons (SADP). The SADP is a 24-item summated rating scale composed of three subscales (Antonak et al., 1988). These dimensions were labeled Optimism-Human Rights (11 items), Behavioral Misconceptions (seven items), and Pessimism-Hopelessness (six items) (Antonak et al., 1988). Although less commonly utilized than the ATDP, the SADP continues to be used as a measurement tool in the literature (e.g., Tervo, Palmer, & Redinius, 2004; Barkheit & Shanmugalingam, 1997).

Antonak and Livneh (1988) reported the scale having Spearman-Brown corrected reliability coefficients from .81 to .85. In support of the validity of the ATDP, Antonak et al. (1988) reported finding moderate relationships with other attitude scales for PWDs (e.g., Form-O of the ATDP, $r = .54$). The SADP, like the ATDP, emphasizes the perceived differences between people with and without a disability at the group or societal level (Gething, 1994).

The Interaction with Disabled Persons Scale. Gething (1994) developed the third rating scale reviewed herein. She sought to develop a measure that could offer a different perspective and provide additional insight into attitude measurement from that given by the two previously discussed scales (the ATDP and SADP). She created a 20-item multidimensional paper-and-pencil measure called the Interaction with Disabled Persons (IDP) scale. Gething (1994) proposed that by asking respondents to reflect on their specific personal experiences with PWDs rather than asking about blanket attitudes toward PWDs at a macro level, the scale would reduce socially desirable responding. This focus is the primary difference in the IDP and the two previously discussed self-report measures (i.e., the SADP and the ATDP). Gething (1994) claimed and found empirical evidence to support that the measure generalizes across several countries (including the United States and Canada). It is believed to be a multidimensional measure which is not disability specific, but instead gets at the underlying constructs behind the general attitudes such as experiences of discomfort and uncertainty.

Based on the original Australian standardization project, the measure has been found to have adequate test-retest reliability with coefficients ranging from .82 given a two week period to .51 over a one-year period (Gething, 1994). In addition, Gething

supported that the measure was internally consistent after assessing the measure on 15 different occasions (Cronbach's alpha coefficient between .74 and .86). The measure was related to the ATDP (-.22, $p < .001$); however, it was not significantly related to the SADP (-.08, *ns*; Gething, 1994)[1]. The measure was investigated in eight additional countries and the validity and reliability information obtained was relatively consistent with the Australian norming population. In the United States sample ($n = 159$), the mean reported score was 62.52 ($SD = 11.75$) with an alpha coefficient of .77. This result is comparable to the Australian sample ($n = 4180$) which found a mean score of 64.14 ($SD = 12.21$) and an alpha coefficient of .79 (Gething, 1994).

Thomas, Palmer, Coker-Juneau, and Williams (2003) supported that the IDP is multidimensional. Thomas et al. conducted an exploratory factor analysis (EFA) and found three underlying dimensions to the IDP: social discomfort, empathy, and fear of having the disability. These researchers then confirmed the three latent constructs through a confirmatory factor analysis (CFA). Although some researchers will question the validity of any type of direct paper-and-pencil measure, researchers using this measure thus far support its reliability and validity as well as its resistance to self presentation artifacts such as socially desirable responding (e.g., Gething, 1994; Loo, 2004; Thomas et al., 2007). The reason this measure is believed to be more resistant is its unique personal level focus versus the general, societal level focus of the previously reviewed summated rating scale measures.

[1] Note: relationships between scales are negative based on the direction of the scoring method used in the different measures.

Summary and Limitations of Direct Measurement Techniques

As reviewed, a number of interesting methods have been established to directly measure attitudes toward PWDs. The most commonly used technique involves summated rating scales (i.e., self-report, paper-and-pencil measures). These measures are preferred because in general, summated rating scales have higher levels of reliability than the other direct methods developed. There are several threats to the validity of the responses from direct method approaches (Antonak et al., 2000, 1988). Particular social processes might occur which provide potential invalidity in the data. For example, the participant might wish to give a good impression, also known as "faking good" or "the evaluation apprehension effect" (p. 215, Antonak et al., 2000). Another disrupting social process sometimes found in direct measurement research is attempting to sabotage the study by providing inaccurate attitudes (i.e., "faking bad") (p. 215, Antonak et al., 2000). In addition, researchers have evidenced that sometimes participants will endorse ideas or statements that one believes is representative of the socially appropriate response (i.e., the social desirability bias) (Strahan & Gerbasi, 1972; Antonak et al., 2000). Another potential problem is that participants might refuse to provide responses due to the fear of revealing controversial or nontypical views (Antonak et al., 2000).

Direct measures are especially vulnerable in a workplace setting where it is likely that employees will be concerned of the job related ramifications of expressing a negative attitude toward PWDs (Haines & Sumner, 2006). These limitations do not suggest that a direct measure cannot be a useful tool under certain conditions or in conjunction with other methods. However, researchers and practitioners should be cautioned of the potential vulnerabilities of direct attitude measurement techniques. Indirect measures

provide a potential alternative or supplementary attitude measurement technique to disability researchers and organizational practitioners to the traditional direct measures.

Indirect Measures

Indirect measures are those in which the respondent is unaware of the measurement, the respondent believes the researcher will know if the respondent is not being honest, or in which the attitude measurement is beyond the respondent's control (Antonak et al., 1988, 2000). These alternative methods to direct measure approaches have increased due to several purported advantages that the indirect approaches offer. When using more direct, obtrusive, and overt measures for attitudes toward PWDs, the nature of the targeted attitude referent is socially sensitive and the individual's measured attitude is potentially altered due to conscious or unconscious mechanisms (Antonak et al., 2000). Indirect measures are believed to avoid or attenuate these mechanisms which could interfere with true attitude measurement. In previous research, indirect attitude measures have been classified into four categories:

> ...those in which the respondents: (1) are unaware that they are being observed or measured (nonobtrusive behavioral observations); (2) are aware that they are being observed or measured, but are unaware of or unclear about the purpose of the measurement situation (projective techniques); (3) are purposefully deceived as to the true purpose of the measurement situation (disguised techniques); and (4) are aware of being measured but are inactive participants in the measurement process (physiological methods). (p. 216, Antonak et al., 2000)

The four categories of indirect measures used in attitude research will be presented followed by a summary and review of associated limitations.

Nonobtrusive Behavioral Observation Techniques. The first category, nonobtrusive behavioral observations, is one way in which attitudes have been assessed indirectly, particularly toward persons with physical disabilities. Using this technique, the researcher intends to observe attitudes of the participants in natural settings. For example, Kleck, Ono, and Hastorf (1966) led participants to an interview room where they were met by either a confederate with a physical disability (using a wheelchair) or a person without a disability (Antonak et al., 2000). Researchers then analyzed participants' responses to this situation as an attitudinal measure that focuses on the behavioral component of an attitude.

Projective Techniques. A second indirect attitude measurement category is the projective technique method. Attitude measurements using projective techniques involve presenting an ambiguous stimulus or a task in which only vague clues have been provided. A trained clinician or psychometrician must then interpret the participant's responses to the technique. It is believed that under this condition, the likelihood of conscious response distortion is lessened and the participant is projecting her or his true attitudes onto the task (e.g., Noonan, Barry, & Davis, 1970).

Disguised Techniques. Antonak et al. (2000) also presented an indirect measurement category classified as the disguised procedures. Using disguised procedures, the researcher attempts to direct the participant's attention away from the attitude for which a measurement is being sought. Under this category, there is a specific structure to the task that is to be performed and the task performance is not based on the assumption of psychodynamic processes driving the attitudes (unlike the projective technique method) (Antonak et al., 2000). The disguised technique may be conducted in

one of three ways: the participant is either (1) unclear to the true purpose of the investigation, (2) under the impression that no control can be exerted over responses, or (3) led to believe that the purpose of the study is different than it really is (Antonak et al., 2000). An example of a disguised technique procedure is the bogus pipeline procedure.

Physiological Techniques. The last category of indirect measures presented in Antonak et al.'s (2000) recent review were the physiological methods. This category of measurement techniques uses instruments to measure reactions believed to be outside of the conscious or voluntary control of the participant. The magnitude of the physiological reaction is believed to be directly related to the intensity of the underlying attitude (see Cacioppo & Tassinary, 1990). The most common physiological method used assesses electrical conductiveness of the skin in response to a stimulus (e.g., interaction with a confederate using a wheelchair versus a confederate not using a wheelchair). In addition, pupil dilation, heart rate, vocal modulations, and blood pressure have also been applied as a physiological measure of attitudes.

Summary and Limitations of Indirect Measurement Techniques

Although indirect measures present a unique and innovative approach to attitude measurement, they also suffer from several limitations (Antonak et al., 2000). First, they tend to be more costly and require more resources than direct measurement techniques. In addition to cost, other questions of practicality also arise. For example, the measures often require trained confederates, administrators, and/or assessors. The indirect methods are also often time consuming to administer and analyze. In addition, the measures have failed to show high levels of reliability or validity (Antonak et al., 1988, 2000) when

compared with the traditional direct measures which brings into question their utility beyond experimental techniques at the current time.

Fortunately, more measures are beginning to be developed which are attempting to obtain the advantageous properties of indirect measures (i.e., create less systematic bias in the measures caused by socially desirable responding and other mechanisms) while at the same time attempting to mirror some of the advantageous properties of direct measures (i.e., higher levels of objectivity, reliability, and psychometric validity). Implicit attitude measurement techniques are one such measurement method that has recently become popular amongst attitude researchers (Antonak et al., 2000; Greenwald et al., 1998; Greenwald et al., 2003; Nosek & Banaji, 2001).

Implicit Attitude Measurement

Over the past few decades, attitude researchers have shifted their attention toward the study of implicit attitude measurement. Although these measurement techniques use terminology taken from the cognitive memory research domain (i.e., implicit versus explicit), measures of implicit attitudes can be considered forms of indirect measures (Fazio & Olsen, 2003; Haines & Sumner, 2006; Pruett & Chan, 2006). In their review of implicit measures, Fazio and Olsen (2003) suggest that implicit measures should be considered as "projective" or "unobtrusive" techniques because they "provide estimates of individuals' attitudes without our having to directly ask them for such information (p. 303)."

Greenwald and his colleagues (e.g., Greenwald & Banaji, 1995; Greenwald, McGhee, & Schwartz, 1998; Greenwald, Nosek, & Banaji, 2003) provided much of the framework for the implicit attitude measurement movement. Greenwald and Banaji's

(1995) keystone paper on implicit social cognition began an expansion of many areas of attitude research to a renewed exploration of implicit attitudes. These researchers described that implicit attitude measures are concerned with measuring the more automatic, and sometimes unconscious, attitudes that individuals hold. Greenwald and Banaji (1995) provided that traces of past experience can affect some performance. Implicit attitudes are those "activated outside of conscious attention (p. 5)" in which the activation happens more quickly than can be consciously mediated. An implicit attitude is thus believed to be an existing attitude projected upon a novel object automatically.

One implicit measuring device developed using this framework that may prove beneficial for research on discrimination against PWDs is the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998). The expansion of implicit attitude research has been facilitated to a large degree by the development of the IAT, which is the most commonly used implicit attitude measurement technique (Fazio & Olsen, 2003).

The Implicit Association Test

Greenwald, McGhee, and Schwartz (1998) developed and investigated the IAT to examine if it would be useful for measuring evaluative associations underlying implicit attitudes. The IAT operates on the premise that it will be cognitively less difficult for an individual to match up two congruent ideas compared to two incongruent ideas. Introspective access to the associations is believed to be prevented by having participants respond to the task in as timely a fashion as possible while still eliciting the correct responses. Thus, the techniques "automatic evaluation" component is derived from the prevention of introspective access to the association strength being measured. Through this procedure, the measurement remains at what the disability researchers have

traditionally referred to as an indirect level of measurement. For this purpose, and to avoid confusion, the IAT can be thought of as an indirect attitude measurement technique while as previously discussed, explicit measures such as paper-and-pencil self-report scales can be considered direct attitude measurement techniques (Fazio & Olsen, 2003; Pruett & Chan, 2006).

Since the measures introduction (Greenwald et al., 1998), the IAT has been used to classify various forms of implicit cognitions beyond implicit attitudes such as self esteem, stereotypes, political views, and consumer preferences (Greenwald, Poehlman, Uhlmann, & Banaji, in press; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). Within the implicit attitude measurement research, it has been adopted for a variety of constructs providing a flexible technique. Greenwald (2007, May) recently reported that an online database search resulted in 350 studies published that reference the IAT, the majority of which were empirical. In addition, between 1998 and early 2007, more than 5 million IATs had been completed online using the Project Implicit website (Greenwald, May 2007).

Introduction to the IAT Process and the "Conventional" Scoring Algorithm

The IAT is a latency based measure in which a participant responds to a series of items (i.e., "exemplars") that are classified into one of four categories using only two responses. In a typical administration, participants must classify a concept discrimination (e.g., instruments and weapons) and an attribute discrimination (e.g., pleasant and unpleasant valence). The IAT measure is derived from the latencies of responses to these two categorization tasks. The resulting differences in response times are interpreted in terms of association strengths. The technique follows the assumption that participants

respond more quickly to concepts and attributes that are assigned to the same response (i.e., right-hand key press or left-hand key press) if the concepts and attributes are strongly associated (e.g., instruments and pleasant) as opposed to when they are weakly associated (e.g., weapons and pleasant) (Greenwald et al., 1998).

An Example Following the Conventional Algorithm.

The original scoring process and administration procedure for the IAT will first be introduced using one of the most highly referenced experiments as an example (found in Greenwald et al., 1998). This process and procedure is now referred to by Greenwald and colleagues (e.g., Greenwald et al., 2003) as the "conventional" IAT scoring algorithm. At the beginning of the procedure, participants are placed in front of a computer and given instructions. Participants are asked to classify stimulus words as quickly as possible while still classifying them correctly.

Trial Blocks. The IAT follows a seven step process (Greenwald et al., 2003). Greenwald and colleagues refer to these steps as blocks (e.g., Lane, Banaji, Nosek, & Greenwald, 2007). In the first block, participants are given instructions to categorize stimulus words (i.e., exemplars) by the target concept that they represent. The exemplars are presented in the middle of a computer screen and the target concept categories (e.g., "Black" versus "White") are displayed on the top left- and right-hand side of the screen. The participants classify the stimulus items using the corresponding key response. A few examples of stimulus words used to represent the categories are "Tyrone" and "Darnell"

or African American names and "Ian" and "Hank" for European American names (Greenwald et al., 1998, p. 1479)[2].

In the second block, the participants are given a new discrimination task. In this block of trials, participants classify exemplars representing the attribute discrimination of pleasant versus unpleasant valence. Examples of stimulus words used during this block are "pleasure," "love," and "peace" for pleasant and "murder," "hatred," and "stink" for unpleasant. Once again, one category response is assigned to the left-hand (e.g., pleasant) and one category response is assigned to the right-hand (e.g., unpleasant).

In the third and fourth blocks, the tasks from the first two blocks are combined (e.g., black and pleasant for left-hand key response versus white and unpleasant for right-hand key response) and participants are presented with stimulus words for the target concept and attribute interspersed randomly in sequential trials. As will be shown in the analysis section, the fourth block is one of the essential, data-collecting blocks as the average response time latency for this block helps derive the participants' scores of associative strength (referenced to as the IAT effect; Greenwald et al., 1998).

In the fifth block, the participant learns a reversal of hand assignment for the target discrimination. In this block, the target discrimination task (e.g., Black versus White) is reversed from the hand response assigned in blocks one, three, and four. The fifth block serves as practice trials for learning the new hand assignment for the target concept and no attribute discrimination task is included in these trials.

[2] In the original study, two IATs were administered, one using only black male and white male names as exemplars and another using only black female and white female names. For the purpose of the current example, only the race IAT using male names will be discussed.

In the sixth and seventh block, participants once again complete the combined task; however, during these trials, the participant is presented with the reversed combined task (e.g., white and pleasant with left-hand key response versus black and unpleasant with right-hand key response). The implicit attitudinal difference in the target categories is believed to be assessed by measuring an effect value derived from the differences in mean latencies for a participant from the fourth block and the seventh block. All tables utilized in the present study are located in Appendix A. See Table 1 for an overview of the blocks following the conventional scoring algorithm.

Technical and Timing Aspects in Conventional Scoring Algorithm. Following the conventional scoring algorithm, several technical aspects were set as guidelines. For example, practice blocks (i.e., blocks one, two, three, five, six) consist of 20 trials. The two test blocks (i.e., blocks four and seven) consist of 40 trials. In the race IAT, Greenwald et al. (1998) set a 250 millisecond interval between response to one stimulus and presentation of the next. In addition, data collection is only recorded for analysis purposes during blocks four and seven. Blocks one, two, three, five, and six serve as practice trials to prepare the participant for the critical combined categorization tasks (i.e., blocks four and seven). Stimulus words during trials are chosen randomly, without replacement so that all items can be used once before any of the stimulus items can be reused.[3] The first two trials in blocks four and seven are dropped from analysis as they typically have longer trial latencies than the remaining trials. In addition, trials with

[3] Note: In Greenwald et al. (1998), 25 stimulus words for each of the four categories were used (i.e., Black names, White names, Pleasant, Unpleasant). Subsequent studies employing the conventional scoring algorithm have varied the number of representative stimulus words.

latencies less than 300 milliseconds or greater than 3,000 milliseconds are set to these lower and upper limits for analysis purposes.

When a participant makes an error in classification under the conventional method, the word "error" appears in the middle of the screen and remains for 300 milliseconds, which expands the time of that intertrial interval by 300 milliseconds. The researcher still uses the latencies when errors are made on trials; however, the researcher sets a total error rate that if surpassed, would eliminate the participant from analysis. For example, in Greenwald et al.'s (1998) race IAT, participants that made errors on over 25 percent of trials in blocks four and seven were eliminated from analysis. Participants are also dropped from analysis if they consistently respond unusually slow to trials (Greenwald et al., 2003).

Analysis Using the Conventional Scoring Algorithm. After administration of the IAT, the participants' latencies from trials four and seven are analyzed in the following manner using the conventional scoring algorithm. First, all latencies undergo a natural logarithm transformation to help normalize the distribution of latencies. Next, the trial averages of the log-transformed data are computed for blocks four and seven. The IAT effect is then calculated by subtracting the hypothesized congruent pairings' (e.g., white names + pleasant, black names + unpleasant) log-transformed average from the hypothesized incongruent pairings' (e.g., black names + pleasant, white names + unpleasant) log-transformed average. The larger this difference value, the stronger the preference for the hypothesized congruent condition target concepts and attributes (e.g., white + pleasant; black + unpleasant) relative to the hypothesized incongruent condition

target concepts and attributes (e.g., black + pleasant; white + unpleasant) (Greenwald et al., 1998).

In the last few years, implicit measurement has established a strong position within psychological research, with the IAT being the most popular measure (Perugini, Gorman, & Prestwich, 2007). Due in part to the wide spread adaptation and use of the IAT, researchers have brought forth many important challenges to the procedure's utility.

Challenges to the Validity of the IAT

Researchers have found that certain extraneous influences can affect the results of participants' IAT score (Nosek, Greenwald, & Banaji, 2007). These challenges will now be introduced. Later, following the introduction of the new scoring algorithm, methods to reduce these methodological flaws will be discussed.

Order of Combined Trial Blocks. One noted flaw in the IAT administration is the order in which the combined tasks are presented to the participant has been shown to affect the magnitude of the IAT effect (Greenwald et al., 1998). An order effect does appear to exist as it appears that the previous combined task (block four) affects participants' performance on the later reversed combined task (block seven).

Cognitive Fluency. Another known extraneous influence is cognitive fluency (Nosek et al., 2007). Cognitive fluency refers to the individual differences observed in respondents overall ability to respond (average response latency) (Nosek et al., 2007). Researchers have found that participants with slower overall response time tended to also have larger IAT effects than those that responded more quickly overall (McFarland & Crouch, 2002; Greenwald et al., 2003).

Age. The age of the participant has also been found to be related to IAT effects, with older participants showing larger IAT effects than younger participants (Nosek et al., 2007; Greenwald & Nosek, 2001; Hummert, Garstka, O'Brien, Greenwald, & Mellott, 2002). Nosek et al. (2007) suggested that this finding might be related to the cognitive fluency influence described in the previous paragraph. That is, that older respondents might have less cognitive fluency.

IAT Experience. Researchers (e.g., Greenwald & Nosek, 2001; Thomas et al., 2007) have also found that increased experience with the IAT affects scores. Particularly, effect magnitudes tend to decline with repeated administrations (Greenwald et al., 2001; Greenwald et al., 2003; Nosek et al., 2007; Thomas et al., 2007). The concern here is for both multiple IATs in a single session (as in the current design) and longitudinal designs with multiple IAT administrations.

Order of IAT and Explicit Measures. Nosek et al. (2007) noted that the order that IAT and self report measures are administered can affect IAT performance as well as self report scores. The order of measure influence is not completely understood at the current time and in most cases it is minimal or non-existent (e.g., Hofmann, Gawronski, Gschwendner, Le, & Schmidt, 2005).

Introduction to the New Scoring Algorithm

Given some of the substantial flaws found when using the conventional scoring algorithm, Greenwald et al. (2003) were driven to develop a new scoring method which would eliminate or reduce some of the challenges posited toward the technique. The improved algorithm has been found by these researchers to outperform the original scoring method on five criteria: "(a) magnitude of implicit-explicit correlation, (b)

resistance to contamination by response speed differences, (c) resistance to the IAT-score-reducing effect of prior experience with the IAT, (d) sensitivity to known effects on IAT measures, and (e) latent implicit-explicit path in CFAs [confirmatory factor analysis]" (Greenwald et al., 2003, p. 212).

Since the introduction of the new algorithm which employs an analysis using what Greenwald et al. (2003) denoted as a *D* measure of effect, researchers have reanalyzed data or replicated studies found to be problematic under the "conventional" scoring algorithm. This method parallels the traditional IAT scoring procedure and will be detailed in the analysis section. Using the *D* measure of effect, researchers have found that many of the original problems were in fact reduced or eliminated (to non-significance) (e.g., Back, Schmukle, & Egloff, 2005; Cai, Sriram, Greenwald, & McFarland, 2004). The methodological and technical aspects of the new scoring algorithm will be introduced in more detail in the next section.

Description of the New IAT Process.

The trial blocks in the improved algorithm follow a very similar pattern to the conventional IAT block administration. As in the conventional IAT, the participant responds to a series of items (i.e., "exemplars") that are classified into one of four categories using only two responses. The seven block administration is almost identical to the blocks presented in the original process. There are only a few differences in the actual administration of the seven blocks. One of the notable suggested changes to the administration of the IAT is to include more practice trials during block five. It is suggested that to reduce the extraneous influence of order of combined blocks, more practice trials can be included in block five (Greenwald et al., 2003). Thus, to prevent

negative transfer, the number of trials during this block can be increased from 20 to 40. All other blocks continue to have the same number of trials as in the conventional method.

Overview of Critical Test Blocks. Within the improved scoring algorithm, data is recorded for all blocks in which one target concept and one attribute is paired on one side and the remaining target block and attribute is paired on the other side (i.e., blocks three, four, six, seven). The critical test blocks have changed from the conventional scoring algorithm. The *D* measure of effect uses the latencies from these four blocks to calculate the IAT effect. Under the conventional scoring procedure, only latencies from blocks four and seven were used in analysis, while blocks three and six served as practice trials of the combined tasks.

In the first set of primary test blocks (blocks three and four), participants are asked to respond quickly with a right-hand key press to items representing one concept and one attribute (e.g., instruments + pleasant), and with a left-hand key press to items representing the remaining concept and attribute (e.g., weapons + unpleasant). In a second grouping of critical blocks (blocks six and seven), the trials are given again; however, as in the conventional scoring algorithm, the key assignment for the target concept has been reversed (e.g., weapons + pleasant for right-hand key press and instruments + unpleasant for left-hand key press) (example from Greenwald et al., 1998; Greenwald et al., 2003). See Table 2 for the present example which overviews the seven block sequence using the improved scoring procedure.

Technical Aspects of the New Scoring Algorithm.

As will be evidenced in the next section, the improved algorithm (Greenwald et al., 2003) has many advantages to the conventional scoring algorithm (Greenwald et al., 1998) in overcoming or reducing some of the methodological flaws found in the IAT. Greenwald et al. (2003) chose the current algorithm from several alternatives for its ability to maximize preferred psychometric properties (e.g., correlations to self-report measures, internal consistency, lessening of the decline in IAT effect sizes across subsequent IAT administrations, etc.) using a very large web-based sample.

Critical Blocks and Screening Details. In the new algorithm, latencies from trials in blocks 3, 4, 6, and 7 are all used in the scoring procedure. Trials with latencies greater than 10,000 milliseconds are eliminated. Furthermore, data from participants which have latencies of less than 300 milliseconds in more than 10% of trials are eliminated. Beyond these two screening procedures, no extreme value treatment is used (Greenwald et al., 2003). Unlike the conventional scoring algorithm, in the improved scoring algorithm no logarithmic transformations are performed on the raw millisecond latency data.

Treatment of Error Latencies. The new algorithm utilizes a penalizing feature for trials in which the respondent incorrectly classifies an exemplar during any of the critical blocks (i.e., blocks 3, 4, 6, and 7). In the original algorithm, the latencies from incorrect responses, without penalty, were still used in calculation of the mean response times for each of the critical blocks. Under the conventional procedure, a participant's data was completely eliminated if more than a specific percent (e.g., 25 percent) of incorrect responses were made. In the new scoring algorithm, the latencies for incorrect responses should be included in the analysis with an error penalty component. Greenwald et al.

(2003) reported that including error penalties did result in a better measure than when error latencies were left unchanged or eliminated from analysis. Nosek et al. (2007) recommended two potential options for introducing an error penalty.

The two recommended procedures are as follows. If a built-in error penalty exists in the software being used (i.e., the respondent is forced to press the correct response before latency timing for that trial ends), this built-in procedure is sufficient (Nosek et al., 2007; Greenwald et al., 2003). The following display will occur for participants using software that has a built-in correction procedure. A red 'X' will appear on the screen below the stimulus item following an incorrect response and the respondent must hit the correct key to move to the next trial. Not all IAT software provides a built-in correction procedure. If the software being used does not provide a built-in error penalty, trials containing errors should be replaced with the mean latency from correct responses in that block, plus a 600 millisecond error penalty.

Greenwald et al. (2003) found that these two methods resulted in relatively similar measurement outcomes with the built-in error penalty being slightly superior to the mean plus 600 millisecond correction procedure. Both of the penalizing procedures were superior to traditional guidelines for handling incorrect trial responses. The current study utilized Inquisit 3.0 (2008) developed by Millisecond Software. Inquisit 3.0 included the built in error latency component so this procedure was chosen. At this point, the researcher can calculate what Greenwald and colleagues refer to as the *D* measure.

The IAT D Measure. The main difference between the conventional and the improved algorithm is that the new algorithm uses what Greenwald and colleagues refer to as the *D* measure as the main outcome measure produced from an IAT administration.

The *D* measure can be used in place of the conventional method which calculates the differences in log-transformed mean latencies found in test blocks 4 and 7 (Greenwald et al., 2003; Lane, Banaji, Nosek, & Greenwald, 2007). Greenwald et al. (2003) chose the italicized capital *D* because it represents that the IAT measure of effect is related to Cohen's *d* (see Cohen, 1992) although there are substantial differences. The main difference between the *D* measure and Cohen's *d* is that the IAT *D* measure replaces the traditional denominator (i.e., within-treatment standard deviation) found in Cohen's *d* with a pooled standard deviation.

To compute the *D* measure, the researcher first computes a single, pooled standard deviation for all trials in blocks 3 and 6 and another pooled standard deviation for all trials in blocks 4 and 7. Means are then computed for each of the four blocks being analyzed (i.e., blocks 3, 4, 6, and 7). Two difference scores are then computed (between blocks 3 and 6 as well as between blocks 4 and 7). In this step, the hypothesized congruent pairings are subtracted from the hypothesized incongruent pairings (Nosek et al., 2007). Both difference scores calculated within subject are then divided by its associated pooled standard deviation. Lastly, an unweighted average of the two resultant quotients is calculated. The resulting average is what Greenwald et al. (2003) describe as the *D* measure.

The formula used for calculating the two *D* scores to be averaged can be expressed as:

$$D = \frac{M_{HypInc} - M_{HypCon}}{\text{"inclusive" } SD}, \quad (1)$$

where M_{HypInc} is the mean latency from trials in a hypothesized incongruent sorting block, M_{HypCon} is the mean latency from trials in the parallel hypothesized congruent sorting

block, and the *"inclusive" SD* is the pooled standard deviation between all trials in blocks 3 and 6 or 4 and 7. To calculate the "inclusive" SD, the following formula was used:

$$"inclusive"\ SD = \sqrt{\frac{\left[(n_1-1)SD_1^2+(n_2-1)SD_2^2+(n_1+n_2)*\frac{(\bar{X}_2-\bar{X}_1)^2}{4}\right]}{(n_1+n_2-1)}} \quad (2)$$

where n_1 is the number of trial latencies retained for inclusion in the congruent block, n_2 is the number of trial latencies retained for inclusion in the incongruent block, $\bar{X}_1$ is the mean trial latency from trials included in the congruent block, and $\bar{X}_2$ is the mean trial latency from trials included in the incongruent block.

Large positive *D* scores suggest higher implicit bias in the direction hypothesized by the researcher. Large absolute value negative *D* scores indicate implicit preference in the reverse direction than that hypothesized. *D* measures at or near zero indicate little to no implicit evaluative preference for either end of the bipolar target concept under investigation.

Additional Considerations Following the New Scoring Algorithm. In addition, the intertrial interval (i.e., the time from the end of one stimulus presentation to the beginning of another) using the improved scoring algorithm ranges from 150 to 750 milliseconds. Researchers are encouraged toward a 250 milliseconds intertrial interval as this is the most common time between stimulus presentations and allows the measure to be completed rather rapidly (Nosek et al., 2007).

In addition, since the attribute and concepts are presented in the same modality (i.e., words), task confusion will be reduced by presenting the attribute stimulus words in one color (e.g., white) and the concept dimensions in another (e.g., green) (Nosek et al., 2007). The above section has outlined the major changes in administration and analysis

for the improved scoring algorithm. It is now time to present evidence for the psychometric benefits of using the new procedure, especially as related to previously discussed flaws in IAT research and methodology.

Benefits of the New Scoring Algorithm

Order of Combined Tasks. One of the most prevalent problems found in the IAT methodology has been that the order in which the combined tasks are completed biases an individual's total IAT effect score. Greenwald and colleagues (e.g., Greenwald et al., 2003, Nosek et al., 2005) have suggested and found evidence to support that adding 20 additional practice trials to block five greatly reduces or eliminates the order effect found under the conventional scoring method. Because this effect is not always completely eliminated, these researchers also suggest continuing to counterbalance order of combined task blocks between subjects (to allow for statistical removal of the biasing effects if one exists).

Cognitive Fluency. The new scoring algorithm has also been shown to reduce or eliminate the cognitive confound that has been found to influence participants' performance on the IAT (Cai, Sriram, Greenwald, & McFarland, 2004). By using the *D* measure which uses a participant's pooled standard deviation, slow overall responding by a participant has been shown to have less influence on IAT effect scores (Greenwald et al., 2003).

Age. The new scoring algorithm (Greenwald et al., 2003) also reduces the relationship between age and IAT score (Nosek et al., 2007). Age has been hypothesized to be related to cognitive fluency. Furthermore, researchers have consistently found a positive relationship between age and reaction times (Der & Deary, 2006). Therefore, the

same corrective procedure of using pooled standard deviation to decrease the influence of cognitive fluency is believed to attenuate the effects of age (Nosek et al., 2007).

Experience with IAT. In addition, the new scoring algorithm (Greenwald et al., 2003) reduces the influence of past experience with IAT administrations. However, at the present time no clear methodological solution exists to deter the influence of a participant's experience with the IAT completely. Greenwald et al. (2003) recommend that this limitation cannot be ignored and must be considered when interpreting results if using multiple IATs, either longitudinally or during multiple administrations at one setting.

Order of Explicit and Implicit Measures. Researchers have found that in general, the order of explicit versus implicit measure administration is minimal or non-existent. In fact, Hofmann et al. (2005) found no effect of order in a meta-analysis of IATs. Additionally, Nosek et al. (2005) found little to no effect based on measurement order using a large, web-based sample. However, Nosek et al. (2007) suggested counterbalancing order of IATs and self report measures if feasible as the order of measure influence is not completely understood at the current time.

Continued Limitation under the New Scoring Algorithm

Although the new scoring algorithm has improved many of the initial problems reported with the IAT methodology, one particularly relevant problem to the current study is the concern with faking the IAT. For example, Steffens (2004) questioned and found support that the IAT is susceptible to faking. If this finding proves to be robust through replications using other constructs, a major problem exists in the framework

behind using the IAT, a supposedly indirect measure which is assumed to be resistant to self presentation artifacts.

Greenwald (2007, May) responded to the faking threat to the validity of the IAT in a presentation given at the 19th annual meeting for the Association for Psychological Science. Greenwald noted that faking was indeed possible for the measure but only if the respondent slows the response time for the congruent pairings. The respondent can then achieve more favorable results by having latencies in the congruent pairings more similar to those in the incongruent pairings. The respondent is not capable of speeding up the incongruent pairings in an effort to fake a more favorable implicit attitude without resulting in an increase in number of errors. Greenwald also noted that although faking the IAT was possible, it cannot occur unless the respondent discovers the methodology of the IAT. Without prior experience to the method or instructions on how the method is scored, this task would be difficult to accomplish.

Accumulating Evidence of the IATs Validity and Usefulness

Despite the continued concerns presented by some researchers, evidence is accumulating in support of the validity and usefulness of the IAT (Greenwald, Nosek, Banaji, & Klauer, 2005). Hofmann et al. (2005) noted in a meta-analysis based on a sample of 126 studies a mean correlation between various IATs and corresponding explicit measures that was significant and increased as a function of moderator variables such as method related aspects of the IAT and conceptual correspondence between implicit and explicit measures. In addition, Nosek and Smyth (2007) used Campbell and Fiske's (1959) multitrait-multimethod design to test the construct validity of implicit attitudes as measured by the IAT. These researchers evidenced that implicit (particularly,

the IAT) and explicit measures do measure related but distinct constructs and the IAT is very adaptive across many different social categories.

In addition, Greenwald, et al. (in press) found in a meta-analysis of 61 studies that the IAT has predictive validity of behavioral criterion measures. In fact, in the area of stereotyping and prejudice, the IAT outperformed explicit measures in predictive ability for various behaviors (average r = .25 and .13, respectively). Although the debate over the usefulness of the IAT is rather young and appears to be far from an ending point, there is still a large quantity of research being done using the method (Greenwald, May 2007). It remains a viable and popular option in attitude and stereotype research despite the many criticisms that it has faced thus far. One area where the IAT is just beginning to be utilized is in the domain of research on attitudes toward PWDs. Currently only three published studies utilizing the IAT to measure attitudes toward PWDs exist. These three studies will now be reviewed.

The IAT Applied to Attitudes toward PWDs

Pruett and Chan's (2006) DA-IAT. Although many different constructs and attitudes toward specific groups have already been measured using the IAT method, the test's application as a means of measuring attitudes toward PWDs in relation to attitudes toward persons without a disability has only recently begun (see Pruett & Chan, 2006; White, Gordon, & Jackson, 2006; Thomas et al., 2007). Pruett and Chan developed a paper-and-pencil version of the IAT to measure attitudes toward PWDs called the Disability Attitude Implicit Association Test (DA-IAT). The DA-IAT measures the number of correctly categorized items in 20-second blocks in a paper-and-pencil format. The DA-IAT utilizes pictures taken from the Project Implicit web site to represent the

disability versus nondisability constructs and uses positive and negative words to represent the evaluative construct. The pictures were obtained from a demonstration version of the IAT applied to PWDs. The disability version used in Pruett and Chan consisted of pictures representing persons with physical disabilities (e.g., a stick figure using a cane) versus pictures representing able bodied persons (e.g., a stick figure in the running position).

Pruett et al. (2006) did find that, as would be expected, congruent associations (nondisabled + positive and disabled + negative) were correctly categorized more frequently than incongruent associations (nondisabled + negative and disabled + positive). In addition, these researchers found that demographic variables (i.e., race, gender, and age) did not predict outcomes on the DA-IAT. Neither the explicit measure used in this study (i.e., the ATDP) nor the DA-IAT was found to be susceptible to socially desirable responding. Lastly, Pruett and Chan did not find a significant relationship between the DA-IAT and the most popular explicit (direct) measure, the ATDP.

White, Gordon, and Jackson (2006). White, Gordon, and Jackson (2006) recently developed a computer-based version of the IAT to measure implicit attitudes toward athletes with disabilities relative to athletes without disabilities. These researchers used actual photographs of "able bodied" athletes and athletes with disabilities drawn from university Web sites (any identifying information was removed) as the target stimuli. Normatively identified pleasant and unpleasant words were used for the attribute dimensions.

White et al. did find participants to have less implicit preference for athletes with disabilities than for athletes without a disability. That is, it took participants substantially longer, on average, to perform the hypothesized incongruent combined tasks (i.e., athletes with disabilities + pleasant and abled athletes + unpleasant) than the congruent pairings. Additionally, unlike the Pruett and Chan (2006) study, White et al. found one version of the IAT to be slightly related to one of two explicit measures used in this study, the ATDP. No relationship was found between the IAT and a different explicit measure, the SADP (Antonak, 1982). Social desirability was not found to be related to the IAT for athletes with disabilities, nor to the two explicit measures, the ATDP or the SADP.

Thomas, Doyle, and Vaughn (2007). Thomas et al. (2007) have also developed a computer-based version of the IAT for PWDs. These researchers used multiple administrations of the IAT using four distinct disability and nondisability categories. There were significant preferences for persons without disabilities for all four disability groups investigated. These four disability categories were: (a) Alcoholic, (b) Cancer, (c) Mental Illness, and (d) Paraplegic [sic]. The researchers conducted a pilot study to obtain the four decided upon target groups which originally resulted in the following five disability target groups: (1) HIV positive, (2) Paraplegic, (3) Alcoholic, (4) Cancer, and (5) Mental Illness. [Please note that Thomas et al. (2007) felt the use of less handicapping language such as person with alcoholism, person with paraplegia, etc. would be problematic given the IAT involves reaction times to stimuli.] HIV positive was subsequently thrown out due to the confusion some participants had in discriminating that HIV positive indicated having the disability and the increased cognitive processing time that this disability group required for classification.

Thomas et al. (2007) did find convergent validity with the use of the IAT that they developed as it was related to the self-report measure included in the study, the IDP (Gething, 1994). Surprisingly, the researchers found that the IAT was also significantly correlated with the Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960). This finding supports Steffens (2004) proposal that the IAT might be vulnerable to faking. The direct measure, the IDP did not significantly correlate to socially desirable responding.

Limitations of Previous IAT and Disability Research

Pruett and Chan's (2006) DA-IAT. The DA-IAT suffered from several limitations. First, as the researchers noted, computer based IATs have produced larger and more robust findings than paper-and-pencil IATs (Greenwald, 2004; cited in Pruett & Chan, 2006). Pruett and Chan, in fact, found the DA-IAT to only marginally relate to a battery of psychosocial variables and called for the development of a computer based version of the IAT. In addition, the test was not found to correlate to the explicit measure, the ATDP. The fact that no relationship was found between the two measures does not support the convergent validity argument that implicit and explicit measures of parallel constructs should be related, although distinct (Nosek & Smyth, 2007). Researchers suggest that the relationship between implicit and explicit measures should be consistently positive, although they can be somewhat variable (Greenwald & Nosek, 2001). The fact that no relationship was found in this case might fall back to the limitation of the less sensitive nature of the paper-and-pencil IAT verses the computer-based measures. Additionally, it might be a result of the previously mentioned limitations of the ATDP as a self-report measure or a combination of both.

Another limitation of the DA-IAT is that due to the modality of the concept stimulus used in this IAT (i.e., using picture stimuli to represent disability versus no disability), the article focused only on displays of physical disabilities. A multidimensional/multiple disability type approach is needed in assessing attitudes toward PWDs as researchers have found that attitudes vary across different disabilities with certain less directly observable disabilities such as mental illness being highly stigmatized (Thomas, 2001; Tringo, 1970). Thomas (2001) directly addressed this issue and supported that an individual's perceptions toward a person with a disability are affected by the overtness of the disability, the degree of risk associated with it, and the PWD's response toward their environment.

White, Gordon, and Jackson (2006). White et al. (2006) suffered similar limitations as observed in Pruett and Chan (2006). First, although White et al. used a computer-based measure, the use of photographs of athletes with disabilities once again limited the study to physical disabilities as opposed to examining multiple diverse types of disabilities. Furthermore, the use of only athletes limited the generalizability further to that of only attitudes toward *athletes* with disabilities in relation to attitudes toward *athletes* without disabilities. Secondly, the lack of a robust relationship brings in to question the construct validity of the measures utilized. As the researchers noted, studying attitudes toward athletes with disabilities (as targeted in the IAT) with the indirect measure versus PWDs in general (as targeted in the two explicit measures) with direct measures, the researchers might have inadvertently been examining two different constructs producing the lack of consistent relationships. Third, White et al. followed the conventional scoring algorithm when administering and analyzing their data which has

already been evidenced in general to be inferior to the new scoring algorithm (e.g., Nosek et al., 2005; Greenwald et al., 2003).

Thomas, Doyle, and Vaughn (2007). Thomas et al. (2007) reported results indicating serious threats to the validity of the IAT. These researchers found a modest relationship between the four IAT effects and socially desirable responding. Although other studies have reported the IAT to be vulnerable to a "cognitive skill confound" (e.g., McFarland & Crouch, 2002) and faking (Steffens, 2004), this study actually showed a relation among IAT effects and respondents' ratings of social desirability. This finding was aberrant from previous IAT research and represents a serious potential threat as one of the proposed advantages of the IAT is its resistance to self-presentational bias in responding (Greenwald et al., 1998).

Thomas et al. (2007) proposed several possible explanations for the resulting relationship between the IAT and the MCSDS. One possible explanation was that there is a "practice" or "learning effect" across administrations of multiple IATs. This explanation was supported by the finding that the results according to the IAT effects did not follow the "hierarchy of preference" as established in previous research (Thomas, 2000; Jones & Stone, 1995; Lyons & Hayes, 1993; Schmelkin, 1984; Tringo, 1970). Instead, in general, disability groups that were administered earlier in the study had larger effect sizes than groups that occurred in later administrations. As noted previously throughout the manuscript, the phenomenon of reduced effects across repeated administrations is widely observed and acknowledged (see Nosek et al., 2007); however, the moderating variables for this extraneous influence is not completely understood as of yet.

Thomas et al. (2007) also proposed that general intelligence, when combined with the predisposition to respond in a socially desirable manner, might be responsible for the improvements that were observed over the multiple IAT administrations conducted in the study. In other words, participants higher in "*g*" (general intelligence) might be able to discover the purpose or nature of the response times between congruent and incongruent trials over successive IAT administrations and actually answer more appropriately from one administration to the next. Furthermore, those higher in socially desirable responding might be more likely to then adjust their responding accordingly. This hypothesis has yet to be tested formally although other researchers have previously noted the phenomenon of the score reducing effect of prior exposure to other IATs (Greenwald et al., 2003; Nosek et al., 2007) as well as the existence of a cognitive confound that affects participants effect sizes (McFarland & Crouch, 2002).

Further limitations found in Thomas et al. included that all participants received the four IAT administrations in the same order. Uniform order of IAT administration did not allow the researchers to formally test if an order effect was really acting on the observed phenomenon of a general decrease in IAT scores across administrations. In addition, the method of analysis for this study relied on the Greenwald et al. (1998) conventional scoring algorithm; however, as previously discussed Greenwald, Nosek, and Banaji (2003) have provided strong support for the benefits of using their improved scoring algorithm. Implementing these changes might improve and offer support for the potential usefulness of this measure as well as provide an opportunity to empirically investigate some of the suggested explanations for the findings from the original study conducted by Thomas et al. (2007). In fact, it is likely that many of the previous problems

found in Thomas et al.'s MDIAT would have been reduced or eliminated using the new scoring algorithm as has been supported by other studies using the new algorithm (Cai, Sriram, Greenwald, & McFarland, 2004; Mierke & Klauer, 2003). The stage for the current study has now been set.

Current Study

Although the IDP has been found to be a reliable self-report measure thus far, a more effective implicit measure would be beneficial to the field of disabilities research. Greenwald et al. (in press) demonstrated that implicit and explicit measures were often differentially related to various behaviors. Furthermore, these researchers found that when applied to socially sensitive target objects, implicit measures were better predictors of behavior than explicit measures. The previously developed disability measures utilizing the IAT technique (i.e., Thomas et al., 2007; Pruett et al., 2006; White et al., 2006) are in need of further modification and refinement incorporating current IAT methodology and considerations. Test refinement is the primary objective of the current study, while the investigation of potential moderator variables for the experience effect found in Thomas et al. (2007) serves as the secondary objective.

The four IATs administered in the current study, collectively referred to as the Multiple Disability IAT (MDIAT), underwent several key changes from the original measure developed by Thomas et al. (2007). These changes included modification of category labels, modification of stimulus words, and other adjustments following the previously discussed recommendations from Nosek et al.'s (2003) new scoring algorithm.

Modification of Category Labels.

The first alteration in the current administration of the IAT from Thomas et al.'s (2007) IAT was the presentation of category labels. Category labels refer to the labels displayed on the top left and top right portion of the monitor that participants use to categorize exemplars. In the present study, the attention shifted more directly and definitively to measurement of attitudes toward *persons* with disabilities rather than attitudes toward disability groups as seen in Thomas et al. Although attitudes toward PWDs was the focus in Thomas et al. (2007), the choice of disability categories (e.g., Cancer vs. Cancer Free) left some room for interpretation by the participant. That is, it was not completely clear whether the participant displayed his or her implicit attitude toward a person with the disability or attitude toward the disability itself. The current study clarified the focus by changing the categories that were presented on the top left and right of the screen to specifically reflect if a stimulus item should be categorized under a person with or a person without the particular disability being studied (e.g., person *with* cancer versus person *without* cancer).

In addition to clarifying the target objects under investigation, the category changes also eliminated the problem in Thomas et al. (2007) of using category words that presented handicapping language. The removal of handicapping language within the measure was the most salient issue in the decision to use the less stigmatizing labels. Lane et al. (2007) reported that "it is the construal of the category that determines how it is evaluated (p. 85)." All judgments are made within a context and it is of utmost importance that the category labels are chosen in a way that does not facilitate a more biasing implicit attitude due to handicapping language. Implicit cognition researchers

have provided evidence that the category labels are more important in determining response latencies than the valence of the target concepts' exemplars (e.g., De Houwer, 2001; Mitchell, Nosek, & Banaji, 2003). By changing the presentation for the eight target concepts used in Thomas et al. (2007), the current study strives to more accurately define the target objects while also serving to decrease the bias that might be evoked by the category label itself.

Modification of Stimulus Words.

In addition to changing the category presentation for the four disability and four non-disability groups, the stimulus words (i.e., exemplars) for each of the groups underwent reliability investigation and modification in a pilot study. Lane et al. (2007) indicated that the stimulus items must be representative of the construct under investigation. Furthermore, they suggest that pilot testing should be used to ensure participants can easily classify the exemplars into the correct categories. If stimulus items do not alter the way the groups are construed, exemplars will have little to no effect on implicit attitudes (Lane et al., 2007). The stimulus items obtained in Thomas et al. (2007) were determined using a different sample taken from a different university. Replicating the pilot study in addition to subjecting the originally used stimulus words to ratings of construct relatedness supported the justification and modification of the chosen stimulus words used in the current study. These exemplars were chosen based on two considerations. First, Nosek et al. (2005) established that there was no loss in reliability with as few as four stimulus words and only slight loss with as few as two stimulus words. Second, the use of fewer words has been recommended when doing so can best capture the concept being investigated (Nosek et al., 2007; Lane et al., 2007). Therefore,

four, rather than six stimulus words per target concept were used in the present study. Details of the method used to obtain the stimulus words are provided in the methods section under "Pilot Study."

Use of the New Scoring Algorithm

The scoring procedure for the current study will follow the previously detailed new scoring algorithm (Greenwald et al., 2003). The revised algorithm has been chosen for its shown increased psychometric performance across several categories in comparison with the conventional scoring algorithm. Instead of reporting the IAT effect in terms of logarithmically transformed milliseconds, the new *D* measure will be reported for each of the individual IATs in the MDIAT as well as a composite *D* score computed by taking the mean of the four individual IAT administrations' *D* values.

Hypotheses for the Current Study.

After implementing the previously covered changes to the MDIAT, several hypotheses were tested. First, as seen in Thomas et al. (2007), participants were expected to respond more slowly on trials in the hypothesized incongruent combined tasks (i.e., sorting words under person with a disability or pleasant on one side and person without a disability or unpleasant on the other side) than on congruent combined sorting tasks (i.e., sorting words under person with a disability or unpleasant on one side and persons without a disability or pleasant on the other side).

> *Hypothesis 1a*: The *D* measure of effect will be significantly greater than zero (i.e., the score obtained if no preference exists) for the person with alcoholism versus person without alcoholism IAT (i.e., the alcoholism IAT).[4]

[4] A positive *D* score is only found in the IAT when quicker average response times occurred in the hypothesized congruent pairing condition than average response times in the hypothesized incongruent pairing conditions.

Hypothesis 1b: The *D* measure of effect will be significantly greater than zero for the person with cancer versus person without cancer IAT (i.e., the cancer IAT).

Hypothesis 1c: The *D* measure of effect will be significantly greater than zero for the person with mental illness versus person without mental illness IAT (i.e., the mental illness IAT).

Hypothesis 1d: The *D* measure of effect will be significantly greater than zero for the person with paraplegia versus person without paraplegia IAT (i.e., the paraplegia IAT).

Hypothesis 1e: The composite MDIAT *D* score (found by averaging the four *D* measures for each participant) will be significantly greater than zero.

In addition, the current study sought to formally test the existence of an order effect on IAT scores in the MDIAT. As noted earlier, researchers have acknowledged the existence of a general decrease in IAT *D* scores across successive IAT administrations (Greenwald et al., 2003; Nosek et al., 2007; Thomas et al., 2007). The new scoring algorithm is believed to reduce this known effect although it is not believed to completely eliminate the experience effect (Greenwald et al., 2003). The current study randomly selected four disability condition orders to test if the order participants completed the four individual IATs affected participants *D* scores for the various disability conditions.

Hypothesis 2: The order of administration of the four IATs composing the MDIAT will affect the magnitude of the *D* scores obtained for each administration time. Specifically, there will be a general decline in magnitude of IAT *D* scores across successive completions of IATs by participants.

In the present study, the relationship between the MDIAT and an explicit measure was also examined. As done in Thomas et al. (2007), the IDP (Gething, 1994) was used to see if the significant relationship found in the previous study could be replicated. The relationship between the MDIAT and the IDP total scores as well as the three underlying factors measured by the scale, social discomfort, empathy, and fear of having the

disability (see Thomas, Palmer, Coker-Juneau, & Williams, 2003) were assessed. The new scoring algorithm is believed to maximize the correlation between implicit and explicit measures through the loss of method/error variance (Greenwald et al., 2003). For this reason, it was hypothesized that the current study would replicate the findings of Thomas et al. (2007) by indicating a significant correlation between the MDIAT *D* scores and the IDP.

> *Hypothesis 3a*: A significant relationship will be found with the IDP total score as well as the three factors identified by Thomas et al. (2003) and the alcoholism IAT.
>
> *Hypothesis 3b*: A significant relationship will be found with the IDP total score as well as the three factors identified by Thomas et al. (2003) and the cancer IAT.
>
> *Hypothesis 3c*: A significant relationship will be found with the IDP total score as well as the three factors identified by Thomas et al. (2003) and the mental illness IAT.
>
> *Hypothesis 3d*: A significant relationship will be found with the IDP total score as well as the three factors identified by Thomas et al. (2003) and the paraplegia IAT.
>
> *Hypothesis 3e*: A significant relationship will be found with the IDP total score as well as the three factors identified by Thomas et al. (2003) and the composite MDIAT scores.

In addition to the previous three research hypotheses, the current study intended to further investigate the potential effects of two variables on the implicit and explicit measures. Under the conventional scoring algorithm, Thomas et al. (2007) found the MDIAT to be related to socially desirable responding. This relationship was once again examined utilizing the new scoring algorithm. Greenwald (May, 2007) recently addressed the concern for faking on the IAT under the new scoring algorithm and stated that the measure can only be faked if participants can figure out the methodology used in scoring

the IAT. If the participant discovered that the incongruent pairings were being compared with the congruent pairing latencies, the participant might then slow the responses for the congruent order conditions. Thomas et al. (2007) did not find the IDP to be susceptible to socially desirable responding. Other researchers have reached similar conclusions using the IDP (Gething, 1994; Loo, 2004). Based on these findings, the following hypotheses regarding socially desirable responding were proposed.

Hypothesis 4a: Socially desirable responding will not be significantly related with the alcoholism IAT.

Hypothesis 4b: Socially desirable responding will not be significantly related with the cancer IAT.

Hypothesis 4c: Socially desirable responding will not be significantly related with the mental illness IAT.

Hypothesis 4d: Socially desirable responding will not be significantly related with the paraplegia IAT.

Hypothesis 4e: Socially desirable responding will not be significantly related with the composite MDIAT scores.

Hypothesis 4f: Socially desirable responding will not be significantly related to the total IDP score, or any of the three factors of the IDP found by Thomas et al. (2003).

Thomas et al. (2007) and McFarland and Crouch (2002) both suggested the existence of a cognitive confound on the IAT. However, both of these researchers found this confound with the traditional scoring algorithm and subsequent research using the new algorithm has reduced or eliminated this effect (Lane et al., 2007). Despite the suggested existence of a cognitive confound, to date, no one has included a measure of general cognitive ability to see if it can account for "cognitive fluency." Building on the possible explanation discussed by Thomas et al. (2007), the fifth hypothesis intended to

investigate if cognitive ability facilitates a practice effect across successive administrations of the IAT where participants higher in cognitive ability could "figure out" the appropriate responses and become better at producing the correct or appropriate response across administrations. In addition, the relationship between cognitive ability and scores on the explicit measure were also investigated.

> *Hypothesis 5a*: General cognitive ability will be negatively correlated with performance scores on the alcoholism IAT.
>
> *Hypothesis 5b*: General cognitive ability will be negatively correlated with performance scores on the cancer IAT.
>
> *Hypothesis 5c*: General cognitive ability will be negatively correlated with performance scores on the mental illness IAT.
>
> *Hypothesis 5d*: General cognitive ability will be negatively correlated with performance scores on the paraplegia IAT.
>
> *Hypothesis 5e*: General cognitive ability will be negatively correlated with performance scores on the composite MDIAT.

The sixth and final hypothesis examined if an interaction existed between the two variables examined in hypotheses four and five. The sixth hypothesis examined if general cognitive ability and tendency toward socially desirable responding had an interactive effect on participants' scores on the MDIAT *D* composite score. This hypothesis built on Thomas et al.'s (2007) suggestion that persons high in both general cognitive ability and tendency toward socially desirable responding would be more likely than other individuals to have lower scores on the MDIAT. In other words, these individuals would be more likely to have both the ability to figure out the methodology of the IAT and the tendency to attempt to fake their responses on the MDIAT tasks to give a good impression.

Hypothesis 6a: There will be a significant interaction between general cognitive ability and socially desirable responding on participants' performance scores on the alcoholism IAT.

Hypothesis 6b: There will be a significant interaction between general cognitive ability and socially desirable responding on participants' performance scores on the cancer IAT.

Hypothesis 6c: There will be a significant interaction between general cognitive ability and socially desirable responding on participants' performance scores on the mental illness IAT.

Hypothesis 6d: There will be a significant interaction between general cognitive ability and socially desirable responding on participants' performance scores on the paraplegia IAT.

Hypothesis 6e: There will be a significant interaction between general cognitive ability and socially desirable responding on participants' performance scores on the composite MDIAT.

Demographics variables (gender, race, age) were also investigated for effects on MDIAT and IDP scores. However, the current study intended to look at these variables in an exploratory fashion. In addition, following the recommendations of Nosek et al. (2007), the current study examined some process and administration variables related to the IAT. Specifically, this study tested if the order of the attitude measure administration (i.e., implicit first or explicit first) affected MDIAT or IDP scores. Additionally, within the MDIAT, the current study tested if the order of the combined categorization task (i.e., congruent first or incongruent first) affected scores on the MDIAT or the IDP.

METHODS

Pilot Study

Participants. A pilot study was conducted to select the stimulus words to be presented during administrations of the MDIAT. Seventy-two students enrolled in psychology courses at a large southeastern university participated in the pilot study for the current stimulus word investigation and modification in exchange for extra credit. The sample consisted of 56 females (77.8%) and 16 males (22.2%). The mean age of participants was 19.9 years (*SD* = 2.7). Fifty-three (73.6%) of the participants were Caucasian, 14 (19.4%) were African American, and three (4.2%) were Asian American.

Procedure. Before beginning the pilot study, participants read and signed an informed consent form explaining that their participation was completely voluntary and that they could withdraw without penalty from the pilot study at any time. A copy of the informed consent used in the pilot study may be viewed in Appendix B. Some of the participants requested clarification for one or more of the disability conditions (most commonly, "person with paraplegia"). For this reason, following completion of the informed consent form, participants were read brief definitions of each of the disability conditions (see Appendix C).

Three of the definitions (i.e., person with alcoholism, person with cancer, person with paraplegia) were drawn from definitions of the disability found in *Merriam-Webster's Pocket Dictionary* (2006). Mental illness was not listed in *Merriam-Webster's*

Pocket Dictionary (2006), so a shortened definition drawn from *The American Heritage Dictionary* (2004) was used to describe this disability condition. To make the definitions applicable to PWDs rather than the disability itself, "Person with" or some similar modification was added to the beginning of each of the disability definitions. Lastly, a fifth definition for "person without the listed disability" was given for the purposes of the present study to describe persons without one of the four disabilities examined in the pilot study.

Participants were next requested to report up to eight words per disability and non-disability condition that they believed were descriptors or adjectives of persons with that condition. The eight conditions were persons with and persons without: (a) alcoholism, (b) cancer, (c) mental illness, and (d) paraplegia. At the end of this paper-and-pencil packet, participants were also asked to write down descriptors of the attributes pleasant and unpleasant. Participants were then instructed to turn in the first packet.

Following creation of descriptive words, participants rated pre-existing stimulus words from those used in the first administration of the MDIAT (from Thomas et al., 2007). Participants rated the degree of representativeness of six descriptors for each of the disability and non-disability conditions. Ratings followed a seven point Likert-type scale (i.e., 1 = Very Unrepresentative; 4 = Neutral; 7 = Very Representative). Lastly, participants rated the 12 attribute descriptors (i.e., six pleasant valence and six unpleasant valence) used in Thomas et al. following the same Likert-type scale. To view the materials completed by participants during the pilot study see Appendix D.

Results of the Pilot Study. After completing the pilot study, the researchers selected stimulus items based on the guidelines of Nosek et al. (2007) and Lane et al.

(2007). The resulting stimulus words may be viewed in Tables 3 and 4. For each category, four rather than six stimulus words were chosen. The selected words were chosen based on their ability to provide for easy identification with the category, strong representativeness of the concept of interest, agreement with previous disability prototype research (e.g., McCaughey & Strohmer, 2005), and (to the extent possible) freedom from confounded meaning with the attribute dimension (i.e., the pleasant and unpleasant categories). The four stimulus words (i.e., exemplars) for the various disability and non-disability conditions were chosen by subject matter experts (SMEs) (one professor and two graduate students) based on several criteria. First, SMEs reviewed the pilot study results from the information provided during the first packet of the pilot study. This portion involved the generation of stimulus words by disability or non-disability condition. Words or phrases containing the same or similar meaning were collapsed into simple one word descriptors (e.g. chemo and chemotherapy under the person with cancer condition). Those words that were most commonly found under each of the concept dimensions were given extended consideration through the next phase of selection.

During the next phase, SMEs reviewed the ratings of the previously existing stimulus words used in Thomas et al. (2007). Stimulus words were ranked highest to lowest by mean representativeness ratings within each of the disability and non-disability conditions. The best potential exemplars for each category were compared both with the most commonly occurring words from the first phase as well as conformation with previous disability literature to further support conceptual and theoretical fit. Next, the words were screened to prevent words causing ambiguity, confounding characteristics with the attribute dimension, or characteristics that would slow down the processing time

(Lane et al., 2007). The resultant stimulus words were obtained by SMEs guided by the results of the pilot study and the recommendations of researchers (Lane et al., 2007; Nosek et al., 2007) to be most associated with a person with or a person without the given disability (see Table 3).

For the attribute dimension (pleasant vs. unpleasant valence), the words created by participants in the first phase were found to have too much overlap with disability related words. For this reason, the top rated stimulus words from those used in Thomas et al. (2007) were chosen as these words were unambiguous and appropriate for the current investigation. In keeping with using four stimulus words for each category condition, only the top four rated words of the six used in Thomas et al. (2007) for both pleasant and unpleasant were retained in the current study (see Table 4). The pilot study allowed for the stimulus words to be updated and modified from those used in Thomas et al. (2007). With the stimulus words now determined, the primary study could be administered.

Participants

Two hundred forty-nine undergraduate students enrolled in psychology courses at a large southeastern university were recruited to participate in the study in exchange for extra credit. Of the 249 participants, three did not meet the set criterion for inclusion on at least one of the four IAT administrations in the MDIAT. In addition, one participant did not complete one of the four IAT administrations. These four participants' data (1.6%) were removed from all subsequent analyses. The resulting sample included 245 usable participants' data. For the self-report questionnaires, mean imputation was used to replace missing items. Five participants did not rate one or more items on the BIDR. Five participants did not rate one or more items on the IDP. The sample consisted of 181

(73.9%) females and 64 (26.1%) males. The mean age of participants was 20.2 years (*SD* = 3.7). Two hundred and one (82.0%) of the participants were Caucasian, 26 (10.6%) were African American, two (0.8%) were Native American, five (2.0%) were Hispanic or Latino, and five (2.0%) were Asian American.

Materials

The materials given to participants included an informed consent form, a demographics questionnaire, the four separate computer administered IATs (collectively, the MDIAT), a self-report attitude measure (IDP), a measure of general intelligence (WPT-R), and a measure of socially desirable responding (BIDR).

Informed Consent Form. The informed consent form explained important information to the participant including key issues involved in participating in the study as well as contact information of the researchers and the institutional review board. Participants were informed that the study involved disability research. In addition, participants were informed that their responses were completely confidential and that they were free to withdraw from the study at any time without penalty. The informed consent form is located in Appendix E.

Demographic Questionnaire. The demographic questionnaire consisted of three items that assessed participants' gender, race, and age. The demographic questionnaire is located in Appendix F.

Interaction with Disabled Persons Scale. The IDP scale consists of 20 items that asked participants to describe how they generally felt during an interaction with a person with a disability. Responses range from 1, "I disagree very much" to 6, "I agree very much." Three items are reverse coded (i.e., items 10, 14, and 15). Scores are summated to

produce a total score. Higher scores on the IDP indicate participants felt greater discomfort. The United States sample found by Gething (1994) appears to be similar to the Australian sample on which it was normed. In the United States sample (n = 159), the mean reported score was 62.52 (SD = 11.75) with an alpha coefficient of .77. In the present study, the mean total score was 72.74 (SD = 11.11). Internal consistency in the present sample resulted in an alpha coefficient of .79. The scale has been found by several researchers to be multidimensional (Thomas, Palmer, Coker-Juneau, & Williams, 2003; Wallymahmed, McKay-Moffat, & Cunningham, 2007). The present study computed IDP total scores and Thomas et al.'s (2003) three factor summated scores (i.e., social discomfort, empathy, and fear of having the disability). The IDP is located in Appendix G.

The Multiple Disability Implicit Association Test (MDIAT). The MDIAT is composed of four separate, disability specific IATs. The measure assesses relative preference between persons with and persons without four specific disability conditions. The four disability conditions (listed in alphabetical order) are: (a) alcoholism, (b) cancer, (c) mental illness, and (d) paraplegia. For the attribute dimension, the MDIAT used pleasant and unpleasant valence. Stimulus words for the concepts and attributes were drawn from the pilot study. The words for the attribute pleasant and unpleasant were the four most highly rated words of six representing each category taken from Thomas et al. (2007). Thomas et al. (2007) originally chose these words from previous IAT studies using this category (Cunningham, Preacher, & Mahzarin, 2001; Greenwald et al., 1998; Rudman, Greenwald, Mellott, & Schwartz, 1999).

The software used to administer the MDIAT was Inquisit 3.0 produced by *Millisecond Software* (2008). The outcome scores for participants' MDIAT administration included the previously detailed *D* scores for each of the four disability-specific IATs as well as a composite *D* score calculated by summing the four *D* scores and taking the arithmetic mean. The development, administration, and analysis of the MDIAT closely followed the updated suggestions provided by leading IAT researchers (e.g., Greenwald et al., 2003; Lane et al., 2007; Nosek et al., 2007). For an example of a typical administration using person with cancer and person without cancer as the target concepts refer to Table 2. The stimulus words used in categorization tasks involving persons with and persons without the various disabilities are located in Table 3. The stimulus words used in categorization tasks involving the pleasant and unpleasant attribute dimension are located in Table 4.

Wonderlic Personnel Test. The Wonderlic Personnel Test-Revised (WPT-R; E. F. Wonderlic Associates, Inc., 1983) will be used to assess cognitive ability. The measure is primarily intended for use in industry and education, but is commonly used in research due to its favorable qualities such as short administration time, abundance of reliability and validity data, and ease of assessment. The WPT-R was administered in the current study using the computer administration option. The measure is 50 items, multiple-choice format and the participant has 12 minutes to respond to as many of the test questions as possible. The measure is scored from 0-50 with one point for every correct response. The construct validity of the measure is supported with generally high correlations with other measures of general mental ability such as the WAIS Full Scale IQ and others (correlation coefficients range from .70-.92; Geisinger, 2001). The scale also has well

documented test-retest reliability coefficients of .82-.94 and alternate form reliability coefficients in the range of .73-.95. (E. F. Wonderlic Associates, Inc., 1983).

Balanced Inventory of Desirable Responding. Paulhus' (1984) balanced inventory of desirable responding, version 6 (BIDR) was used to measure socially desirable responding. The BIDR Version 6 (Paulhus, 1994) consists of 40 items which are ranked on a seven point scale with one denoting "not true" and seven denoting "very true." The BIDR consists of two subscales, impression management (IM) and self-deceptive enhancement (SDE; also referred to as self-deceptive positivity, Paulhus, 1991). According to Paulhus (1984), impression management denotes intentionally distorting self-descriptions to be viewed more favorably by others while self-deceptive enhancement refers to an unconscious propensity to view oneself in a favorable light. Half of the items in both subscale are reverse coded.

Scoring the BIDR can follow one of two methods, continuous or dichotomous scoring, as outlined by Paulhus (1994). In continuous scoring, all answers on the response scale are counted. In the dichotomous scoring scheme, only extreme scores are counted. The present study selected continuous scoring as this method has been found superior to the dichotomous scoring scheme in several dimensions (Stöber, Dette, & Musch, 2002). For example, the continuous scoring scheme has resulted in higher Cronbach's alphas, higher correlations with other measures of socially desirable responding, and higher correlations with personality traits believed to be associated with tendency to respond in a socially desirable manner than the dichotomous scoring scheme (Stöber et al., 2002).

Previous research has demonstrated the following mean internal consistency estimates: .68 for the SDE subscale, .74 for the IM subscale, and alpha of .80 for the

summed total measure (Li & Bagger, 2007). The BIDR has been shown to have good convergent validity as it correlates highly with the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960) and other similar measures (Stöber et al., 2002). In the current study, the SDE subscale had an alpha coefficient of .60, the IM subscale had an alpha coefficient of .79, and the total scale had an alpha coefficient of .79. The BIDR is located in Appendix H.

Procedure

Upon arrival to the laboratory, participants entered a quiet room containing a computer and a desk with space for writing. Data collection occurred in a single session and lasted approximately 60 minutes. Participants in the current study were asked if they were 19 or older. For participants under the age of 19, a parental assent form was collected. Participants age 19 or older received an informed consent form which they were instructed to review and sign. After collecting the appropriate forms, the participants were read the same list of disability condition definitions as were used in the pilot study (see Appendix C). Participants then completed the demographic questionnaire. Next, participants either completed the explicit, paper-and-pencil attitude measure (i.e., IDP) or the computer administrated implicit attitude measure (i.e., MDIAT).

The MDIAT administration procedure warrants further discussion. As previously mentioned, the MDIAT used Inquisit 3.0 software (2008) to run the four IATs. The computer programming code used in the current study was modified from sample code provided by the software developers. The code implemented the process and analysis recommendations detailed earlier from Greenwald et al. (2003) and others (Lane et al., 2007; Nosek et al., 2007). Participants were instructed to press the “e” or “i” key to

categorize exemplars appearing in the middle of the screen to the left or right respectively. During an IAT administration, the computer screen used a black backdrop and concept labels and exemplars appeared in white font while attribute labels and exemplars appeared in green font. Category labels were posted on the top left and top right of the screen. Intertrial intervals were set at 250 milliseconds. Additionally, the software recorded trial latency from the onset of a trial until the correct response was given. This timing feature allowed for the use of the new scoring algorithm to compute the *D* scores which takes advantage of a built in error penalty for incorrect responses.

At the beginning of the MDIAT procedure, the administrator of the study read the instructions located on the computer screen for the first portion of the procedure. Following reading this section, the administrator asked if the participant had any questions about the instructions. The administrator then asked the participant to complete the IAT administration and to open the laboratory door when the participant arrived on the screen stating that this portion of the study was completed. The administrator then left the room and closed the door. Participants then completed the seven blocks associated with completing one IAT administration. After completing each block, a screen presented instructions for the next categorization task and requested the participant to press the space bar to begin the next block. After the participant opened the laboratory door the administrator then set up the next IAT administration.

The order in which participants completed the four IAT administrations composing the MDIAT was determined based on a randomly selected administration order. Four of 24 possible IAT administration orders were randomly selected and included in the current study. The randomly selected orders are located in Table 5.

Following completion of the MDIAT and the IDP, participants completed the computer administered cognitive ability test (WPT-R). The participants were provided with two blank sheets of paper and access to pens and pencils to make notes. The administrator reviewed the instructions for the WPT-R with the participant and then asked them if they had any questions. Participants first read through instructions for the test and were given a few example questions. Participants completed the WPT-R using the mouse to navigate through an online computer interface managed by Wonderlic, Inc. The online administration controlled the time of the administration (i.e., 12 minutes) and following completion, Wonderlic, Inc. reported scores of participant performance back to the primary researcher.

Following completion of the cognitive ability measure, participants completed the paper-and-pencil measure of socially desirable responding (i.e., BIDR). After participants completed this packet, they were informed that this concluded participation in the current study and participants were debriefed and asked if they had any questions.

The current investigation examined two process effects, the current study counterbalanced the order of implicit (i.e., MDIAT) and explicit (i.e., IDP) attitude measurement completion between-subjects. The second order variable investigated applies specifically to the MDIAT. This factor counterbalanced the order of presentation of the combined sorting tasks. Approximately half of the participants were presented with the hypothesized congruent combined categorization tasks first for all four IAT administrations and the other half were presented with the hypothesized incongruent combined categorization tasks first for all four IAT administrations. These process factors were examined in an exploratory fashion as it is currently not clear when and how these

variables affect outcome scores (see Lane et al., 2007). Restricted randomization was used to acquire balanced cell sizes for the first 240 participants. The remaining nine participants were assigned to order conditions completely at random. A random number generator was used to create an order assignment list prior to any participant's arrival to the laboratory.

Materials were stored in the following fashion. The informed consent form was stored separately from the remaining paper-and-pencil and computer administered material. Computer administered materials were stored in a designated electronic folder on the desktop of the laboratory computers and backed up on the jump drive of the primary investigator. Paper-and-pencil materials were stored in a designated filing cabinet in the laboratory space. All study material was assigned a participant number to link participants' material together. No identifying information besides participants' identification numbers were listed on any of the study materials.

RESULTS

Descriptive Statistics and Intercorrelations of Study Variables

In preparing the data to conduct the appropriate analyses to test the proposed hypotheses, descriptive statistics and intercorrelations of the study variables were calculated. These data are presented in Table 6. *D* scores for each of the IAT administrations were computed as well as the composite *D* score. In addition, an IDP total score was calculated as well as the three subscale scores from Thomas et al. (2003; i.e., social discomfort, empathy, and fear of having the disability). From the BIDR, total socially desirable responding scores were calculated in addition to the two subscale scores (i.e., SDE and IM). Lastly, computations from the general cognitive ability measure, the WPT-R, were reported.

Analyses of Implicit Bias (Hypothesis 1)

Hypothesis 1 posited that significantly more implicit bias will be shown toward persons with each of the disability conditions compared to persons without the disability conditions. Additionally, implicit bias will be shown overall across all IAT administrations toward persons with the various disabilities compared to persons without the four disabilities. Hypothesis 1 was supported. Implicit preference was assessed by examining outcome scores from the MDIAT. As a reminder, these *D* scores are calculated by subtracting the mean response latency in the hypothesized congruent critical blocks from the mean response time in the incongruent critical blocks and

dividing by the pooled standard deviation. To test this hypothesis, one-tailed simple *t*-tests (H_a: $\mu > 0$) were run for each of the four disability specific IATs as well as the composite MDIAT scores setting the hypothesized population mean to zero (i.e., no preference exists).

Results from the analyses indicated that for the alcoholism IAT, implicit bias was displayed by participants, $t(244) = 27.22$, $p < .001$. The mean effect score indicated a moderate to strong preference for persons without alcoholism in relation to persons with alcoholism ($D = .60$, $SD = .35$). In addition, analysis of the cancer IAT displayed implicit preference for persons without cancer in relation to persons with cancer, $t(244) = 27.31$, $p < .001$. As with the alcoholism IAT, mean effect scores indicated a moderate to strong preference ($D = .68$, $SD = .39$). For the mental illness IAT, effect scores were once again significantly larger than zero, $t(244) = 26.09$, $p < .001$. The mean effect scores were indicative of a moderate to strong preference for persons without mental illness versus persons with mental illness ($D = .65$, $SD = .39$). The paraplegia IAT also resulted in effect scores significantly larger than zero, $t(244) = 30.06$, $p < .001$. As with the previous three disability condition IATs, moderate to strong bias toward persons with paraplegia versus persons without paraplegia was displayed ($D = .64$, $SD = .35$). Lastly, the composite MDIAT *D* score (found by averaging the four *D* measures for each participant) was found to be indicative of implicit bias, $t(244) = 41.85$, $p < .001$. The composite *D* score indicated a moderate to strong preference for persons without a disability relative to persons with a disability ($D = .65$, $SD = .24$). Table 7 displays the means and standard deviations for the four disability-specific IATs as well as the composite IAT.

Analysis of MDIAT Disability Order Effect (Hypothesis 2)

Hypothesis 2 predicted that order of IAT administration would affect *D* outcome scores. More specifically, it was suggested that a general trend toward a decline in disability-specific *D* scores would occur following successive IAT administrations. The current study used a random effects variable to assign participants to one of four randomly selected order conditions. The individual IAT *D* scores were collected for times one, two, three, and four (regardless of disability specific IAT) for each participant. To test if order of disability administration had any effect on IAT outcome scores, a repeated measures ANOVA was conducted with individual IAT outcome scores serving as the dependent variable and administration time one through four serving as the within-subjects factor. The results of the repeated measures ANOVA indicated that order of disability administration had an effect on outcome scores, $F(3, 242) = 51.54$, $p < .001$, partial $\eta^2 = .39$.

To further investigate if the expected downward trend in *D* scores existed in the data, post hoc tests using the Scheffé post hoc criterion for significance were conducted. Results of the post hoc tests indicated that administration time one (mean $D = .86$, $SD = .38$) resulted in a significantly larger mean *D* score than found in times two, three, and four. Administration time two (mean $D = .66$, $SD = .34$) resulted in significantly higher *D* scores than times three and four. Administration time three (mean $D = .55$, $SD = .33$) and four (mean $D = .52$, $SD = .32$) did not significantly differ from one another. A plot was constructed of the mean *D* scores for each IAT administration order further illustrating the existence of an order effect. Along the *x*-axis of the plot, the order in which the disability IAT was given is labeled. Along the *y*-axis, mean *D* scores were plotted. All

figures used in the current study are located in Appendix I. The plot demonstrating the experience effect is shown in Figure 1.

The order of administration did appear to produce the expected trend in the data. All disability-specific IAT scores were found to significantly decline following each successive IAT administration except for the fourth IAT administration. Because IAT administration order within the MDIAT affected disability-specific outcome scores, subsequent analyses were conducted using only the composite MDIAT *D* scores (calculated by taking the arithmetic average of the four disability-specific *D* scores).

Analysis of Implicit and Explicit Measure Relationship (Hypothesis 3)

Hypothesis 3 predicted that the correlation between the MDIAT and the IDP would be significant. This hypothesis was not supported, $r(243) = .09$, *ns*. In addition, none of the three subscales within the IDP identified by Thomas et al. (2003) were significantly related to the MDIAT composite score.

Analysis of Relationships with Socially Desirable Responding (Hypothesis 4)

Hypothesis 4 suggested that the MDIAT will not be related to socially desirable responding and that the IDP will not be related to socially desirable responding. The MDIAT was not significantly related to the BIDR total scale score ($r(243) = .00$, *ns*) or the SDE ($r(243) = .02$, *ns*) or IM ($r(243) = -.01$, *ns*) subscales. The IDP total score was significantly related to the BIDR total scale score ($r(243) = -.18$, $p < .01$) as well as the SDE subscale, $r(243) = -.25$, $p < .01$. The IDP was not related to the IM subscale, $r(243) = -.08$, *ns*. The current investigation further assessed the three factors composing the IDP for relationship with the BIDR total score and two subscales. Factor 1, Social Discomfort, was significantly related to BIDR total scores ($r(243) = -.19$, $p < .01$) and the SDE

subscale ($r(243) = -.27$, $p < .01$) but not IM ($r(243) = -.08$, *ns*). Factor 2, Empathy, was not significantly related to the BIDR total score or either of the two subscales. Factor 3, Fear of Having the Disability, was related to the SDE subscale ($r(243) = -.15$, $p < .05$) but not the BIDR total scores or the IM subscale. To view the correlations presented herein, please refer to Table 6.

Analysis of Relationships with General Cognitive Ability (Hypothesis 5)

Hypothesis 5 suggested that a significant relationship would exist between general cognitive ability and MDIAT scores. In addition, the relationship between general cognitive ability and IDP scores was investigated in an exploratory fashion. The specific hypothesis was that there would be a negative relationship between cognitive ability and MDIAT scores; thus participants higher in cognitive ability would tend to express less biased implicit attitudes toward PWDs than participants lower in cognitive ability. This hypothesis was not supported, $r\ (243) = .12$, *ns*.

The relationship between cognitive ability and explicit attitudes toward PWDs was also examined. Results of this analysis indicated that no significant relationship existed between the IDP total score and the WPT-R, $r\ (243) = .02$, *ns*. Furthermore, cognitive ability was not related to any of the three subscales composing the IDP.

Analysis of Interaction on MDIAT Scores (Hypothesis 6)

The final prediction, hypothesis 6, stated that participants' cognitive ability and tendency to respond in a socially desirable manner would create a significant interaction effect on MDIAT outcome scores. Multiple regression analysis was used to test this hypothesis. Guidelines provided by Hair, Black, Babin, Anderson, and Tatham (2006) were used to test for a significant interaction effect. Hierarchical regression was utilized.

In the first step, cognitive ability total scores and socially desirable responding total scores were entered in to the model as independent variables with MDIAT composite scores entered as the dependent variable. In the next step, the interaction variable (i.e., the product of the first two independent variables) was entered in to the model. The result indicated no significant difference in R^2 change, $F^*(1, 241) = .007$, *ns*. The interaction effect on MDIAT outcome scores was not supported. Table 8 summarizes the results of the hierarchical regression analysis.

Analyses for Gender and Race Effects

The effects due to gender and race have been mixed in previous research. For this reason, these demographic variables were investigated in an exploratory fashion. Due to the sample, age was not examined as the restriction of range in age limited any meaningful analysis on this variable. The demographics analyses will begin with examining effects on the MDIAT.

MDIAT Demographics Analysis. In examining the effects of the demographic variables on the MDIAT, the gender and race variables were analyzed using a 2 x 2 ANOVA. The dependent variable was the composite MDIAT score. Gender (male vs. female) and race (Caucasian vs. African American) were the between subjects factors. In the current sample, as with the sample in Thomas et al., (2007) there were a limited number of participants in several of the race minority groups. Therefore, the current analysis examined only Caucasian ($n = 201$) and African American ($n = 26$) participants' data. The means for the composite MDIAT score by gender and race are presented in Table 9.

The two-way ANOVA indicated that there was no effect due to gender, $F(1,223) = 2.72$, *ns*. Race also did not produce an effect on MDIAT scores, $F(1,223) = 3.85$, *ns.* Also, the interaction between race and gender was not significant, $F(1,223) = 2.31$, *ns*. These results suggested that gender, race, and the gender x race interaction did not produce significant effects on the composite MDIAT scores.

IDP Demographics Analysis. The IDP was also examined for race and gender effects. A two-way ANOVA using a 2 (Caucasian vs. African American) x 2 (male vs. female) fixed factor design was used to assess if these variables effected IDP total scores. Table 10 displays the descriptive statistics for the IDP by race and gender. The two-way ANOVA indicated no significant effect due to race, $F(1,223) = 1.83$, *ns*, gender, $F(1,223) = 0.22$, *ns*, or the race by gender interaction, $F(1, 223) = 1.77$, *ns*.

Analyses of Procedural Effects on MDIAT and IDP

Hypothesis 2 was previously confirmed indicating that the order of IAT administrations within the MDIAT affected individual IAT outcome scores. Two other potential order effect variables were introduced in the current study to examine their impact on MDIAT and IDP outcome scores. These variables were manipulated between subjects and they include order of attitude measurement administration (i.e., implicit vs. explicit first) and order of combined sorting task administration (i.e., congruent vs. incongruent first). The next section will examine the impact of these manipulations on MDIAT scores.

Procedural Order Effects on the MDIAT. A 2 x 2 ANOVA was used to investigate the order effects on composite MDIAT scores. The attitude measure administration order (implicit or explicit completed first) and the combined sorting task

order (incongruent or congruent combined tasks presented first) were entered as between subjects factors. The means and standard deviations by the two procedural order variables are presented in Table 11.

The main effect of implicit versus explicit measure completion order did not affect MDIAT composite scores, $F(1,241) = 2.14$, *ns*. The main effect of combined sorting task within the MDIAT (hypothesized congruent pairings versus hypothesized incongruent pairings first) did affect MDIAT scores, $F(1,241) = 15.86$, $p < .001$. Participants that completed the hypothesized congruent sorting task first expressed more implicit bias than participants receiving the hypothesized incongruent sorting task first. Lastly, a significant interaction existed between attitude administration order and order of combined sorting task on composite MDIAT scores, $F(1,241) = 4.85$, $p < .05$. It appears that participants that completed the IDP first and also received the incongruent sorting tasks first within the MDIAT administration had markedly lower MDIAT composite scores than participants receiving different administration orders. The interaction plot can be viewed in Figure 2.

Procedural Order Effects on the IDP. As done in the MDIAT analysis, a 2 x 2 ANOVA was used to investigate the order effects on total IDP scores as well as the three factors composing the IDP. The attitude measure administration order (implicit or explicit completed first) and the combined sorting task order (incongruent or congruent combined tasks presented first) were entered as between subjects factors. The means and standard deviations by the two procedural order variables for the IDP total scores and the three factors identified by Thomas et al. (2003) are presented in Table 12.

The main effect of implicit versus explicit measure completion order did not affect IDP total scores, $F(1,241) = 2.38$, *ns*. Additionally, as might be expected the main effect of combined sorting task within the MDIAT (hypothesized congruent pairings versus hypothesized incongruent pairings first) did not affect IDP total scores, $F(1,241) = .017$, *ns*. Lastly, there was no significant interaction between attitude administration order and order of combined sorting task on IDP total scores, $F(1,241) = 2.77$, *ns*. It appears that participants' IDP total scores were not affected by implicit-explicit administration order, combined task administration order, or the interaction between the two procedural variables.

Next, the three subscales composing the IDP were investigated by the two between subjects procedural variables. For the first factor, Social Discomfort, implicit-explicit order had a significant effect on scores, $F(1,241) = 4.10$, $p < .05$. Participants that completed the implicit measure first scored significantly lower on the social discomfort factor. Participants did not differ on level of factor 1 by combined task order or by the interaction between the two procedural variables. The second factor, Empathy, did not show any significant differences in level by the two procedural variables or the interaction between the two. Additionally, the procedural variables had no effect on the third variable, Fear of Having the Disability.

DISCUSSION

The IAT is believed to assess implicit associative strength through the examination of response times to combined sorting tasks. Proponents of the IAT technique have expressed that the technique allows for a flexible method that can assess various forms of implicit cognition including implicit attitudes and stereotypes (Greenwald et al., in press). Furthermore, the implicit nature of the measurement is believed to be less susceptible to confounds that might affect explicit measurement such as socially desirable responding. The current study yielded mixed results regarding the validity of the MDIAT as a measure of implicit attitudes toward persons with disabilities. The study also presented new information regarding a common explicit attitude measure, the IDP. The results found in the current study will now be discussed, followed by implications, limitations, and suggestions for future research.

Analysis of Demographic Variables

The results of the implicit and explicit measurement administration indicated no significant differences in attitudes by gender, race, or gender x race interaction. The analyses for race were constricted to examining Caucasian versus African American participants. For the implicit measure, the MDIAT, the lack of significant differences by demographic variables for MDIAT outcome scores is congruent with previous IAT research examining attitudes toward persons with disabilities (Pruett et al., 2006; Thomas et al., 2007; White et al., 2006).

In addition, the explicit measure, the IDP, did not find differences by gender, race, or the gender x race interaction. Traditionally, explicit measures have indicated that women hold less biased attitudes toward persons with disabilities than do men (Yuker & Block, 1986). However, recent empirical evidence investigating the effects of gender on self-report attitude measures in the disability literature has been mixed. For example, some studies have indicated that females have displayed more favorable attitudes on self report measures than do men (Chism & Satcher, 1997; Loo, 2004; MacLean & Gannon, 1995) while others have found no significant difference in attitudes by gender (e.g., Loo, 2001; Thomas et al., 2007). Some researchers have suggested that this finding may be due to a real shift in societal views toward persons with disabilities (Loo, 2001; Thomas et al., 2007; Yuker et al., 1986).

Persistence of Implicit Bias

Hypothesis 1 was supported. As found in previous empirical investigations utilizing the IAT to measure attitudes toward persons with disabilities (e.g., Pruett et al., 2006; Thomas et al., 2007; White et al., 2006), overall, participants' scores were indicative of implicit preference for persons without disabilities in comparison to persons with disabilities. At the most basic level, this result suggests that participants have a stronger association with persons without the four disabilities examined and pleasant and persons with the four disabilities and unpleasant than when the attribute discrimination task is reversed.

Observation of Practice or Learning Effect

In addition to refining and modifying the measure used in Thomas et al. (2007), the present study allowed to formally test the existence of the practice or learning effect

suggested to be found in the first administration of the MDIAT. Hypothesis 2 stated that the practice effect would persist under the new scoring algorithm. In general, disability-specific IATs would have larger effect sizes when administered earlier in the MDIAT process rather than later. This hypothesis was supported. It does appear that participants were "getting better" (i.e., displaying less bias) at successive IAT administrations across subsequent administrations. The continued existence of this practice effect presents a real challenge to the validity and interpretability of individual, disability-specific IAT outcome scores calculated within the MDIAT. Presently, researchers are aware that experience with the IAT does reduce IAT outcome scores but have no explanation to account for the reduction in scores (Nosek et al., 2007). Additionally, the new scoring algorithm was expected to reduce the strength of the experience effect (Greenwald et al., 2003). Of course, an alternative explanation for the observed practice effect could be that participants' attitudes are becoming less biased through the process of taking the MDIAT; however, the plausibility of this explanation is suspect.

Relationship between Implicit and Explicit Attitude Measures

Hypothesis 3 suggested that after implementing the new scoring algorithm, the MDIAT would be significantly related to the IDP as found in Thomas et al. (2007). The lack of a significant relationship between the MDIAT and the IDP in the present study was unexpected as the new scoring algorithm was predicted to increase implicit-explicit measurement correlations (Greenwald et al., 2003). This finding brings another significant challenge to the validity of the MDIAT as implicit and explicit measures are believed to assess related but distinct constructs (Nosek & Smyth, 2007).

Two previous meta-analyses have examined the strength of implicit-explicit correlations in IAT studies (Hofmann, Gawronski, Gschwendner, & Schmitt 2005; Nosek, 2005). Hofmann et al. (2005) reported average implicit-explicit correlations of .24 whereas Nosek (2005) reported average correlations of .37. However, the lack of a significant relationship between implicit and explicit measures of attitudes toward persons with disabilities is not completely inconsistent with previous IAT disability research. That is, Thomas et al. (2007) found a significant albeit modest relationship between the original MDIAT and the IDP. Furthermore, Pruett et al. (2006) found no significant relationship between the IAT they developed and the most commonly used explicit measure, the ATDP. In addition, White et al. (2006) only found one version of their IAT measure to have a small relationship with one of the two explicit measures included in their study, the ATDP.

One explanation for the weak to non-existent relationship found in the disability IAT studies to date is that the self-report measures used in the disability literature are absolute measures directed specifically toward persons with disabilities. For example, the IDP assesses interpersonal discomfort experienced during contact with persons with disabilities (Gething, 1994). The IAT on the other hand provides a relative measurement of association strength (Nosek, Greenwald, & Banaji, 2007). That is, the outcome score not only provides information about attitudes toward persons with the disability in question but also the associative strength relative to persons without the disability. The most common self-report attitude measures generally do not ask about attitudes toward persons without disabilities.

In fact, the noticeably higher correlation reported from the meta-analysis conducted by Nosek (2005) is due in part to the explicit measures used in the studies that Nosek (2005) examined. Nosek (2005) examined IAT measures that used relative feeling thermometers as the explicit, self-report attitude measurement which better corresponds to the IAT measurement. In addition, Nosek (2005) posited several moderating variables that affect the degree of relationship between automatic (implicit) evaluations and controlled (explicit) evaluations. One such moderating variable pertinent to measures of attitudes toward persons with disabilities is the self-presentational nature of the target construct. That is, the degree to which participants might be likely to modify responses on a self report measure for personal or social purposes. The case could certainly be made for altering responses on a controlled evaluation measure for a socially sensitive topic such as attitudes toward persons with disabilities. In fact, as will be discussed, the present study found the IDP to be related to socially desirable responding.

Examination of Potential Moderating Variables affecting MDIAT Outcome Scores

As noted previously, the phenomenon of reduced effects across repeated administrations of the IAT has been observed by other researchers (e.g., Greenwald & Nosek, 2001; Greenwald et al., 2003; Nosek et al., 2007); however, the moderating variables for this extraneous influence is not completely understood as of yet. Two potential individual difference variables, socially desirable responding and cognitive ability, were examined in the current study following the suggestions provided by Thomas et al. (2007). These variables were examined to investigate if they were related to overall outcome scores. In addition, the interaction between cognitive ability and

socially desirable responding was examined to see if these two variables in combination accounted for incremental variance in MDIAT composite scores.

Socially Desirable Responding as Extraneous Influence on MDIAT. The current study reexamined the anomalous finding of Thomas et al. (2007) which reported that the MDIAT was related to socially desirable responding. This previous finding posed a serious threat to the IAT methodology as one of the proposed benefits of the method is its "automatic" evaluative nature and its resistance to self presentation confounds such as socially desirable responding. The current study did not replicate this finding. Neither the socially desirable responding total score, nor the impression management or self deception enhancement subscales were significantly related to MDIAT composite scores. This finding addresses a major limitation of the original MDIAT measure. This result is also congruent with other disability related IAT research expressing no significant relationship between the disability IATs developed and socially desirable responding (e.g., Pruett et al., 2006; White et al., 2006). It is noteworthy that the current study assessed socially desirable responding using the BIDR (Paulhus, 1994) where as Thomas et al., (2007) used the Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960).

Cognitive Ability as Extraneous Influence on MDIAT. Hypothesis 5 suggested that cognitive ability would be related to MDIAT outcome scores. More specifically, hypothesis 4 stated that those higher in cognitive ability would exhibit lower MDIAT composite scores (thus producing a negative relationship between composite MDIAT scores and cognitive ability). This hypothesis was not supported. No significant relationship was found between these two variables. Cognitive ability was expected to

effect MDIAT scores because more intelligent individuals might be able to learn the purpose of the IAT procedure across administrations and then have the cognitive resources to tailor their responses to exhibit less bias whereas less intelligent individuals were predicted to be less likely to identify the purpose of the IAT and then to have the ability to tailor responses accordingly (Thomas et al., 2007).

Other researchers have postulated similar potential moderating variables. For example, McFarland & Crouch (2002) identified what they termed as a cognitive skill confound where participants that responded slower overall would produce larger outcome scores indicating more biased evaluation. In addition, Mierke and Klauer (2003) have extensively studied the effects of task switching ability on IAT outcome scores. Nosek, Greenwald, and Banaji (2007) refer to this grouping of extraneous influences as "cognitive fluency" and purport that the new scoring algorithm has reduced these effects. The lack of a significant effect is positive in that cognitive ability was not found to be a confounding variable on MDIAT scores. However, this result also was unable to account for the practice or learning effect identified in the current study.

Interaction of Socially Desirable Responding and Cognitive Ability. Hypothesis 6 tested whether there was a significant interaction between socially desirable responding and cognitive ability which affected MDIAT outcome scores. Thomas et al. (2007) suggested that participants with the disposition to respond in a socially desirable manner as well as with the capability to understand the methodology from previous administrations would alter (consciously or unconsciously) their reaction times in subsequent IAT administrations to portray less bias. That is, those participants high in both cognitive ability and tendency to respond in a socially desirable manner were

expected to demonstrate the least amount of implicit bias as assessed by the MDIAT. This hypothesis was not supported. It appears that these potential extraneous variables did not explain differences in MDIAT scores and did not account for the persistent practice or learning effect found when using multiple administrations of the IAT. Although the processes underlying the experience effect were not identified in the present study, it is encouraging that it appears that following the new scoring algorithm, neither cognitive ability nor socially desirable responding significantly affected MDIAT outcome scores.

The IDP

In addition to investigating the psychometric properties of the MDIAT, the current study had a secondary goal of further investigating properties of the IDP (Gething, 1994). Previous research has demonstrated that the IDP is not susceptible to socially desirable responding unlike other explicit attitude scales such as the most widely used self-report measure, the ATDP (e.g., Gething, 1994; Thomas et al., 2007; see Loo, 2001 for an exception). The current study did find that the IDP was related to socially desirable responding total scores and the self deception enhancement subscale but not the impression management subscale. Although IDP evaluations are not directly affected by the desire to portray a better image of self to others (impression management), it does appear that the scale is fallible in that it is affected by the propensity to view oneself in a more favorable light (self deception enhancement). It is worth noting that the previous investigations of susceptibility to socially desirable responding using the IDP have measured the construct using the MCSDS (Crowne & Marlowe, 1960) rather than the BIDR (Paulhus, 1994).

Implications

Within industrial-organizational (I-O) psychology, the IAT as a measurement technique has received a substantial amount of consideration and attention. As an example, the first volume of the Society for Industrial and Organizational Psychology's (SIOP's) new journal, *Industrial and Organizational Psychology: Perspectives on Science and Practice*, Frank Landy (2008) addressed stereotype research in general, and the IAT in particular, through a focal article. The article was followed by 13 commentaries. Although Landy (2008) criticized I-O psychology research using the IAT in laboratory settings, many of the commentaries strongly defended the usefulness of laboratory research utilizing the IAT (e.g., Greenwald, 2008; Hanges & Ziegert, 2008; Rudman, 2008). The IAT's ingression into disability research is more recent, and as is often the case with nascent research streams, many questions remain to be answered regarding the potential usefulness of this technique in assessing attitudes toward persons with disabilities.

Although the modifications and new scoring algorithm appeared to attenuate some problems (e.g., confounds due to socially desirable responding), the practice effect found in Thomas et al. (2007) persisted. In general, disability-specific IATs had larger effect sizes when administered earlier in the MDIAT process rather than later. However, by presenting multiple disability-specific IAT administrations, a broader perspective on disability attitude was captured rather than limiting to one IAT administration using only picture stimuli of persons with a physical disability as conducted in previous IAT disability research (e.g., Pruett et al., 2006; White et al., 2006). That is, the present study included a physical disability, a communicable disease, a mental disability, and a

disability involving substance abuse. Although thus far the IAT appears to be effected by previous experience, the dimensionality included in the MDIAT is necessary to capture a deeper understanding of a participant's implicit attitude toward persons with disabilities.

The results have important implications for researchers and practitioners. If using the IAT, one must consider participants' level of experience with the measurement technique. Consideration should be given to a participant's level of experience with the MDIAT specifically and with the IAT technique in general (Lane et al., 2007). Additionally, if using the IAT as a criterion score in a longitudinal design, concern must be given to the possibility that decreases in scores from baseline performance may reflect previous exposure to the technique rather than an improvement in implicit attitude from time one to time two. Given the results of the current study, important questions must be addressed before use of the IAT in workplace settings is warranted.

Future research should examine if the MDIAT exhibits predictive validity. A meta-analysis conducted by Greenwald et al. (in press) reported that in the area of stereotyping and prejudice, the IAT outperformed explicit measures in predictive ability for various behaviors. Additionally, future studies should attempt to further investigate the underlying causes for the experience effect found when using multiple administrations of the IAT within a single setting. The procedure might also benefit from further design modifications to the IAT procedure that would produce more robust and stable effect sizes regardless of previous exposure.

The current study intended to enhance the psychometric soundness of the measure developed by Thomas et al. (2007). It appears that the new scoring algorithm and

modifications implemented have not solved all of the problems found in the original implementation of the measure.

Limitations

The current study has several limitations related to the sample used. The investigation utilized undergraduate students in a laboratory setting. The sample was a convenience sample and some of the characteristics of the participant pool used in this sample are likely to be different than the composition found in a workforce setting. This sample was rather homogenous in age and predominately composed of female participants. In addition, the sample was homogenous in race/ethnicity. Being that these participants completed the study in exchange for extra credit, they may not have been as invested or motivated as employees in a workplace setting. For these reasons, the findings from this study may not directly generalize to a non-student population.

As Collela and Stone (2005) noted in their discussion of the current limitations in disability research, 30 of the 37 empirical studies that they found were conducted using laboratory settings with student or MBA manager participants. One of the barriers in conducting disability research is that companies are sometimes reluctant to allow disability research for fear of legal repercussions (Collela & Stone, 2005). In addition, in the early development of this measure, it is more practical to study this measure in a more controlled laboratory setting. After accumulating more validation information, it will be more useful to employ the MDIAT in an applied setting. Similar to other measurement validation studies (e.g., Pruett & Chan, 2006), the current study intended to be a preliminary investigation into the validity of the modified IAT as originally developed and investigated by Thomas and colleagues (Thomas et al., 2007). Notwithstanding the

limitations, this study advances the current application of the IAT as a potential method to assess attitudes toward persons with disabilities.

REFERENCES

Abroms, K. I., & Kodera, T. L. (1979). Acceptance hierarchy of handicaps: Validation of Kirk's statement 'Special education often begins where medicine stops'. *Journal of Learning Disabilities, 12,* 15-20.

Albertsons Inc. v. Kirkingburg, 1999. 143 F.32 1228, 119 S. Ct. 2162.

Alghazo, E. M., Dodeen, H., & Algaryouti, I. A. (2003). Attitudes of pre-service teachers towards persons with disabilities: Predicting for the success of inclusion. *College Student Journal, 37,* 515-522.

Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 42 U.S.C. 12101 *et seq.* (1990).

Antonak, R. F. (1981). *Development and psychometric analysis of the Scale of Attitudes Toward Disabled Persons* (Technical Report No. 1). Durham, NH: University of New Hampshire, Education Department.

Antonak, R. F. (1982). Development and psychometric analysis of the Scale of Attitudes Toward Disabled Persons. *The Journal of Applied Rehabilitation Counseling, 13,* 22-29.

Antonak, R. F., & Livneh, H. (1988). *The Measurement of Attitudes toward People with Disabilities : Methods, Psychometrics and Scales.* Springfield, IL : Charles C Thomas.

Antonak, R. F., & Livneh, H. (2000). Measurement of attitudes towards persons with disabilities. *Disability and Rehabilitation, 22,* 211-224.

Back, M. D., Schmukle, S. C., & Egloff, B. (2005). Measuring task-switching ability in the Implicit Association Test. *Experimental Psychology, 52,* 167-179.

Brehm, S. S., Kassin, S., & Fein, S. (2005). *Social psychology* (6th ed.). Boston: Houghton Mifflin Company.

Budig, M. J. (2002). Male advantage and the gender composition of jobs: Who rides the glass escalator. *Social Problems, 49,* 258-277.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45,* 16-28.

Cai, H., Sriram, N., & Greenwald, A. G. (2004). The Implicit Association Test's D measure can minimize a cognitive skill confound: Comment on McFarland and Crouch (2002*). Social Cognition, 22,* 673-684.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Cannon, B. J., & Szuhay, J. A. (1986). Faking can elevate scores on the Attitudes Toward Disabled Persons scale. *Rehabilitation Counseling Bulletin, 30,* 120-123.

Caruso, D. R., & Hodapp, R. M. (1988). Perceptions of mental retardation and mental illness. *American Journal of Mental Retardation, 93,* 118-124.

Cascio, W. F., & Aguinis, H. (2005). *Applied psychology in human resource management* (6th ed.). New Jersey: Pearson Prentice Hall.

Chism, M., & Satcher, J. (1997). Human resource management students' perceptions of employment and individuals with disabilities. *Journal of Applied Rehabilitation Counseling, 28*, 38-42.

Clain, S. H., & Leppel, K. (2001). An investigation into sexual orientation discrimination as an explanation for wage differences. *Applied Economics, 33,* 37-47.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Colella, A., & Stone, D. L. (2005). Workplace discrimination toward persons with disabilities: A call for some new research directions. In R. L. Dipboye & A. Colella (Eds.), *Discrimination at Work: The Psychological and Organizational Bases* (pp. 227-253). New Jersey: Lawrence Erlbaum Associates Publishers.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349-354.

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

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443-451.

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom health and life survey. *Psychology and Aging, 21*, 62-73.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt, Brace, Jovanovich.

EEOC ADA handbook, (2005). [On-line] Available: http://www.eeoc.gov.

E. F. Wonderlic Associates, Inc. (1983). *Wonderlic Personnel Test Manual.* Northfield, IL: Author.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297-327.

Feldman, D., Gordon, P. A., White, M. J., & Weber, C. (2002). The effects of people-first language and demographic variables on beliefs, attitudes, and behavioral intentions toward people with disabilities. *Journal of Applied Rehabilitation Counseling, 33,* 18-25.

Fichten, C. S., Schipper, F., & Cutler, N. (2005). Does volunteering with children affect attitudes toward adults with disabilities? A prospective study of unequal contact. *Rehabilitation Psychology, 50,* 2005.

Finkelstein, L. M., & Burke, M. J. (1998). Age stereotyping at work: The role of rater and contextual factors on evaluations of job applicants. *Journal of General Psychology, 125,* 317-345.

Geisinger, K. F. (2001). Review of the Wonderlic Personnel Test and Scholastic Level Exam. In B. S. Plake & J. C. Impara (Eds.), *The Fourteenth Mental Measurements Yearbook* (pp. 1360-1363). The University of Nebraska-Lincoln: Buros Institute of Mental Measurements.

Gething, L. (1994). The interaction with disabled persons scale. *Journal of Social Behavior and Personality, 9,* 23-42.

Gottlieb, J., Corman, L., & Curci, R. (1984). Attitudes toward mentally retarded children. In R. L. Jones (Ed.), *Attitude and Attitude Change in Special Education: Theory and Practice* (pp. 143-156). Reston, VA: Council for Exceptional Children.

Gough, H. G. (1960). The Adjective Check List as a personality assessment research scale. *Psychological Reports, 6,* 107-122.

Greenwald, A. G. (2004, January). *Revised top ten list of things wrong with the IAT*. Presentation at the Attitudes Preconference of the fifth annual meeting of the Society for Personality and Social Psychology, Austin, TX.

Greenwald, A. G. (2007, May). *Assessing the Validity of Implicit Association Test (IAT) Measures*. Presentation at the 19th annual meeting of the Association for Psychological Science, Washington, DC.

Greenwald, A. G. (2008). Landy is correct: Stereotyping can be moderated by individuating the out-group and by being accountable. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 430-435.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4-27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

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

Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48,* 85-93.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216.

Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the salience asymmetry interpretation of the Implicit Association Test: Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General, 134,* 420-425.

Haines, E. L., & Sumner, K. E. (2006). Implicit measurement of attitudes, stereotypes, and self-concepts in organizations: Teaching old dogmas new tricks. *Organizational Research Methods, 9,* 536-553.

Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate Data Analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hanges, P. J., & Ziegert, J. C. (2008). Stereotypes about stereotype research. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 436-438.

Hebl, M. R., Foster, J. B., Mannix, L. M., & Dovidio, J. F. (2002). Formal and interpersonal discrimination: A field study of bias toward homosexual applicants. *Personality and Social Psychology Bulletin, 28,* 815-825.

Hirschberger, G., Florian, V., & Mikulincer, M. (2005). Fear and compassion: A terror management analysis of emotional reactions to physical disability. *Rehabilitation Psychology, 50,* 246-257.

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*, 1369-1385.

Hosoda, M., Stone-Romero, E. F., & Coats, G. (2003). The effects of physical attractiveness on job-related outcomes: A meta-analysis of experimental studies. *Personnel Psychology, 56,* 431-462.

Hummert, M. L., Garstka, T. A., O'Brien, L. T., & Mellott, D. S. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology and Aging, 17*, 482-495.

Inquisit 3.0.1.1 [Computer Software] (2008). Seattle, WA: Millisecond Software.

Janicki, M. P. (1970). Attitudes of health professionals toward twelve disabilities. *Perceptual and Motor Skills, 30,* 77-78.

Jones, G. E., & Stone, D. L. (1995). Perceived discomfort associated with working with persons with varying disabilities. *Perceptual and Motor Skills, 81,* 911-919.

Kleck, R., Ono, H., & Hastorf, A. H. (1966). The effects of physical deviance upon face-to-face interaction. *Human Relations, 19,* 425-436.

Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. *Personality and Social Psychology Bulletin, 21,* 58-75.

Landy, F. J. (2008). Stereotypes, bias, and personnel decisions: Strange and stranger. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 379-392.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) (Pp. 59–102). In B. Wittenbrink & N. S. Schwarz (Eds.). *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford Press.

Lee, B. A. (1998). *A review of ADA jurisprudence in who should be accommodated.* Paper presented at the annual meeting of the Academy of Management, San Diego, CA.

Li, A. & Bagger, J. (2007). The Balanced Inventory of Desirable Responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement, 67*, 525-544.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*, 55.

Loo, R. (2001). Attitudes of management undergraduates towards person with disabilities: A need for change. *Rehabilitation Psychology, 46*, 288-295.

Loo, R. (2004). Attitudes toward employing persons with disabilities: A test of the sympathy-discomfort categories. *Journal of Applied Social Psychology, 34*, 2200-2214.

Lyons, M., & Hayes, R.. (1993). Student perceptions of persons with psychiatric and other disorders. *American Journal of Occupational Therapy, 47*, 541-547.

MacLean, D., & Gannon, P. M. (1995). Measuring attitudes toward disability: The Interaction with Disabled Persons Scale revisited. *Journal of Social Behavior & Personality, 10*, 791-806.

MacMillan, D. L., & Morrison, G. M. (1984). Sociometric research in special education. In R. L. Jones (Ed.), *Attitude and Attitude Change in Special Education: Theory and Practice* (pp. 70-92). Reston, VA: Council for Exceptional Children.

McCaughey, T. J., & Strohmer, D. C. (2005). Prototypes as an indirect measure of attitudes toward disability groups. *Rehabilitation Counseling Bulletin, 48*, 89-99.

McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20,* 483-510.

Merriam-Webster's Pocket Dictionary (2006). *Merriam-Webster's Pocket Dictionary.* Springfield, MA: Merriam-Webster, Inc.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85,* 1180-1192.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132,* 455-469.

Murphy v. United Parcel Service 1999. Service 141 F.3d 1185, 119 S.Ct. 1331.

Noonan, J. R., Barry, J. R., & Davis, H. C. (1970). Personality determinants in attitudes toward visible disability. *Journal of Personality, 39,* 1-15.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*, 565-584.

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

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

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54*, 14-29.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The Measurement of Meaning.* Urbana, IL: University of Illinois Press.

Papinchock, J. M. (2005). Title I of the Americans with Disabilities Act. In F. J. Landy (Ed.), *Employment discrimination litigation: Behavioral, quantitative, and legal perspectives* (pp. 294-335). San Fransisco, CA: Jossey-Bass.

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

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17-59). San Diego, CA: Academic Press, Inc.

Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6*. Unpublished manuscript, University of British Columbia, Vancouver, Canada.

Perry, E. L., & Finkelstein, L. M. (1999). Toward a broader view of age discrimination in employment-related decisions: A joint consideration of organizational factors and cognitive processes. *Human Resource Management Review, 9,* 21-49.

Perugini, M., O'Gorman, R., & Prestwich, A. (2007). An ontological test of the IAT: Self-activation can increase predictive validity. *Experimental Psychology, 54,* 134-147.

Pruett, S. R., & Chan, F. (2006) The development and psychometric validation of the Disability Attitude Implicit Association Test. *Rehabilitation Psychology, 51,* 202-213.

Ravaud, J., Madiot, B., & Ville, I. (1992). Discrimination towards disabled people seeking employment. *Social Science & Medicine, 35,* 951-958.

Rehabilitation Research and Training Center on Disability Demographics and Statistics (2005). *2004 Disability Status Reports*. Ithaca, NY: Cornell University.

Richardson, S. A., Goodman, N., Hastorf, A., & Dornbusch, S. (1961). Cultural uniformity in reaction to physical disabilities. *American Sociological Review, 26,* 241-247.

Rudman, L. A. (2008). The validity of the Implicit Association Test is a scientific certainty. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 426-429.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437-465.

Salend, S. J., & Giek, K. A. (1988). Independent living arrangements for individuals with mental retardation: The landlord's perspective. *Mental Retardation, 26,* 89-92.

Schmelkin, L. P. (1984). Hierarchy of preferences toward disabled groups: A reanalysis. *Perceptual and Motor Skills, 59,* 151-157.

Schneider, K. T., Swan, S., & Fitzgerald, L. F. (1997). Job-related and psychological effects of sexual harassment in the workplace: Empirical evidence from two organizations. *Journal of Applied Psychology, 82,* 401-415.

Shaver, P. R., & Scheibe, E. (1967). Transformation of social identity: A study of chronic mental patients and college volunteers in a summer camp setting. *Journal of Psychology, 66,* 19-37.

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

Steinmetz, E. (2006). *Americans with disabilities: 2002* (Current Population Rep. No. P70-107). Washington, DC: U.S. Census Bureau. [On-line]. Available: http://www.census.gov/prod/2006pubs/p70-107.pdf. Retrieved January 29, 2007.

Stephenson, W. (1953). *The Study of Behaviour: Q-Technique and its Methodology.* Chicago, IL: University of Chicago Press.

Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment, 78*, 370-389.

Stone, D. L., & Colella, A. (1996). A model of factors affecting the treatment of disabled individuals in organizations. *Academy of Management Review, 21,* 352-401.

Stone, E. F., Stone, D. L., & Dipboye, R. L. (1992). Stigmas in organizations: Race, handicaps, and physical attractiveness. In K. Kelley (Ed.), *Issues, theory, and research in industrial and organizational psychology* (pp. 385-457). New York: Elsevier Science Publishers.

Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 28,* 191-193.

Sutton v. United Airlines (1999). 130 F.3d 893, 119S. Ct. 2139.

Taylor, H. (October 14, 1998). Americans with disabilities still persuasively disadvantaged on a broad range of indicators. *The Harris Poll, 56*. [On-line]. Available: http://www.harrisinteractive.com/harris_poll/index.asp?PID=152. Retrieved January 29, 2007.

Taylor, H. (October 7, 2000). Conflicting trends in employment of people with disabilities 1986-2000. *The Harris Poll, 59*. [On-line]. Available: http://www.harrisinteractive.com/harris_poll/index.asp?PID=121. Retrieved January 29, 2007.

Tervo, R. C., Palmer, G., & Redinius, P. (2004). Health professional student attitudes towards people with disability. *Clinical Rehabilitation, 18,* 908-915.

The American Heritage Dictionary (2004). *The American Heritage Dictionary of the English Language* (4th ed.). Boston: Houghton Mifflin Company.

Thomas, A. (2000). Stability of Tringo's hierarchy of preference toward disability groups: 30 years later. *Psychological Reports, 86,* 1155-1156.

Thomas, A. (2001). The multidimensional character of biased perceptions of individuals with disabilities. *Journal of Rehabilitation, 67,* 3-9.

Thomas, A., Doyle, A., & Vaughn, D. (2007). Implementation of a computer based Implicit Association Test as a measure of attitudes toward individuals with disabilities. *Journal of Rehabilitation, 73,* 3-14.

Thomas, A., Palmer, J. K., Coker-Juneau, C. J., & Williams, D. J. (2003). Factor structure and construct validity of the Interaction with Disabled Persons scale. *Educational and Psychological Measurement, 63,* 465-483.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33,* 529-554.

Tringo, J. L. (1970). The hierarchy of preference toward disability groups. *The Journal of Special Education, 4,* 295-306.

Wallymahmed, A. H., McKay-Moffat, S. F., & Cunningham, C. C. (2007). The Interaction with Disabled Persons scale: A validation with UK midwives. *Social Behavior and Personality, 35*, 1049-1060.

Wells, S. J. (2001). Is the ADA working? *HR Magazine, 46,* 38-47.

White, M. J., Gordon, P., & Jackson, V. (2006). Implicit and explicit attitudes toward athletes with disabilities. *Journal of Rehabilitation, 72,* 33-40.

Wittenbrink, B., Judd, C., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262-74.

Yuker, H. E., & Block, J. R. (1986). *Research with the Attitude Toward Disabled People Scales: 1960-1985.* Hampstead, NY: Hofstra University Center for the Study of Attitudes Toward People with Disabilities.

Yuker, H. E., Block, J. R., & Campbell, W. J. (1960). *A scale to measure attitudes toward disabled persons* (Human Resources Study No. 5). Albertson, NY: Human Resources Center.

Yuker, H. E., Block, J. R., & Younng, J. H. (1966). *The measurement of attitudes toward disabled persons*. (Human Resources Study No. 7). Albertson, NY: Human Resources Center.

APPENDIX A

List of Tables

Table 1

Schematic Overview of the Conventional Algorithm of the Implicit Association Test (Adopted from Greenwald, Nosek, & Banaji, 2003)

Block	No. of trials	Function	Items assigned to left-key response	Items assigned to right-key response
1	20	Practice	Black American name	White American name
2	20	Practice	Pleasant	Unpleasant
3	20	Practice	Black and Pleasant	White and Unpleasant
4	40	Test	Black and Pleasant	White and Unpleasant
5	20	Practice	White American Names	Black American Names
6	20	Practice	White and Pleasant	Black and Unpleasant
7	40	Test	White and Pleasant	Black and Unpleasant

Note: As noted in Greenwald et al. (2003), the standard procedure is to present half of the participants with the above block order and the other half with the substitution of blocks one, three, and four with blocks five, six, and seven respectively. Moving these blocks allows for counterbalancing of order effects.

Table 2

Schematic Overview of the Improved Algorithm of the Implicit Association Test (Adopted from Greenwald et al., 2003)

Block	No. of trials	Function	Items assigned to left-key response	Items assigned to right-key response
1	20	Practice	Person WITH Cancer	Person WITHOUT Cancer
2	20	Practice	Pleasant	Unpleasant
3	20	Test Block 1A	Person WITH Cancer and Pleasant	Person WITHOUT Cancer and Unpleasant
4	40	Test Block 2A	Person WITH Cancer and Pleasant	Person WITHOUT Cancer and Unpleasant
5	40	Practice	Person WITHOUT Cancer	Person WITH Cancer
6	20	Test Block 1B	Person WITHOUT Cancer and Pleasant	Person WITH Cancer and Unpleasant
7	40	Test Block 2B	Person WITHOUT Cancer and Pleasant	Person WITH Cancer and Unpleasant

Table 3

Stimulus Words Selected for Disability and Non-Disability Conditions

Person with Alcoholism	**Person without Alcoholism**
Addicted	Sober
Drunk	Control
Compulsive	Stable
Unhealthy	Healthy
Person with Cancer	**Person without Cancer**
Sick	Healthy
Weak	Strong
Tired	Normal
Frail	Lucky
Person with Mental Illness	**Person without Mental Illness**
Confused	Normal
Different	Capable
Crazy	Healthy
Dependent	Independent
Person with Paraplegia	**Person without Paraplegia**
Immobile	Mobile
Restricted	Active
Disabled	Healthy
Challenged	Functional

Table 4

Stimulus Words Selected for Attribute Categorization

Pleasant	**Unpleasant**
Vacation	Hatred
Sunrise	Devil
Diploma	Bomb
Love	Poison

Table 5

Randomly Selected Order Conditions

Order 1	Order 2	Order 3	Order 4
Mental Illness	Mental Illness	Paraplegia	Cancer
Alcoholism	Paraplegia	Alcoholism	Paraplegia
Paraplegia	Cancer	Mental Illness	Alcoholism
Cancer	Alcoholism	Cancer	Mental Illness

Table 6

Means, Standard Deviations, and Intercorrelations for Study Variables

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Composite MDIAT *D*	.65	.24	--												
2. Alcoholism *D*	.60	.35	.71**	--											
3. Cancer *D*	.68	.39	.63**	.21**	--										
4. Mental Illness *D*	.65	.39	.65**	.39**	.11	--									
5. Paraplegia *D*	.65	.34	.66**	.32**	.30**	.17**	--								
6. IDP Total	72.8	11.0	.09	.16*	.04	.02	.02	--							
7. IDP Social Discomfort	12.0	4.5	.08	.12	.07	.02	.00	.79**	--						
8. IDP Empathy	23.2	2.5	.00	.04	.04	-.06	-.02	.61**	.27**	--					
9. IDP Fear of Having the Disability	17.4	3.6	.07	.09	.02	.04	.03	.74**	.42**	.41**	--				
10. BIDR Total	159.2	22.5	.00	.04	-.05	.07	-.06	-.18**	-.19**	-.03	-.11	--			
11. BIDR SDE	84.5	11.0	.02	-.02	.03	.04	.00	-.25**	-.27**	-.05	-.15*	.73**	--		
12. BIDR IM	74.7	16.2	-.01	.07	-.09	.08	-.08	-.08	-.08	-.00	-.04	.89**	.34**	--	
13. WPT-R	23.8	5.0	.12	.04	.10	.01	.18**	.02	.02	-.11	.04	-.14*	-.09	-.13*	--

Note. * p < .05; ** p < .01; *n* = 245. MDIAT = Multiple Disability Implicit Association Test. Variables 2-5 represent the four separate IAT administrations composing the MDIAT. IDP = Interaction with Disabled Persons Scale. Variables 7-9 represent the three factors identified in the IDP by Thomas et al., 2003. BIDR = Balanced Inventory of Desirable Responding; SDE = self-deceptive enhancement subscale; IM = impression management subscale. WPT-R = Wonderlic Personnel Test-Revised.

Table 7

Means and Standard Deviations of MDIAT Outcome Scores

Variable	M	SD
Alcoholism IAT	.60	.35
Cancer IAT	.68	.39
Mental Illness IAT	.65	.39
Paraplegia IAT	.65	.34
Composite MDIAT	.65	.24

Table 8

Hierarchical Regression Analysis Summary for Individual Difference Variables predicting MDIAT Composite Scores

Variable	B	SEB	β	R^2	ΔR^2
Step 1				.02	
WPT-R	0.01	.00	.12		
BIDR Total	0.00	.00	.02		
Step 2				.02	.00
Interaction	0.00	.00	.05		

Note. WPT-R = Wonderlic Personnel Test-Revised; BIDR Total = Balanced Inventory of Desirable Responding Total Score; Interaction = WPT-R x BIDR Total.
$n = 245$; $^*p < .05$.

Table 9

Descriptive Statistics of MDIAT Composite Scores by Gender and Race

	Caucasian		African American	
IAT Administration	Male[a]	Female[b]	Male[c]	Female[d]
Composite *D*				
M	.64	.64	.84	.66
SD	.28	.23	.13	.23

[a]$n = 49$. [b]$n = 152$. [c]$n = 7$. [d]$n = 19$.

Table 10

Descriptive Statistics of IDP by Gender and Race

	Caucasian		African American	
	Male[a]	Female[b]	Male[c]	Female[d]
IDP Total Score				
M	70.2	74.8	68.5	68.7
SD	10.9	10.5	5.5	13.3
IDP Factor 1: Social Dis.				
M	11.7	12.5	9.1	9.8
SD	4.6	4.5	3.3	4.2
IDP Factor 2: Empathy				
M	21.9	23.8	22.7	22.9
SD	2.8	2.2	1.8	3.0
IDP Factor 3: Fear				
M	17.5	17.7	18.4	15.7
SD	3.5	3.5	2.9	3.6

[a]$n = 49$. [b]$n = 152$. [c]$n = 7$. [d]$n = 19$.

Note. Factor 1 = Social discomfort; Factor 2 = Empathy; Factor 3 = Fear of having the disability.

Table 11

Descriptive Statistics of MDIAT Composite Scores by Procedural Order Variables

	Congruent First		Incongruent First	
IAT Administration	Imp First[a]	Exp First[b]	Imp First[c]	Exp First[d]
Composite *D*				
M	.70	.72	.64	.53
SD	.26	.21	.20	.26

Note. Imp First = Implicit measure was completed first. Exp First = Explicit measure was completed first.

[a]$n = 60$. [b]$n = 61$. [c]$n = 62$. [d]$n = 62$.

Table 12

Descriptive Statistics of IDP by Procedural Order Variables

	Congruent First		Incongruent First	
	Imp First[a]	Exp First[b]	Imp First[c]	Exp First[d]
IDP Total Score				
M	72.8	72.9	74.9	70.4
SD	12.1	10.7	11.0	9.9
IDP Factor 1: Social Dis.				
M	12.0	11.9	13.1	10.9
SD	4.8	4.6	4.5	3.9
IDP Factor 2: Empathy				
M	23.3	22.9	23.6	22.8
SD	2.4	2.8	2.4	2.6
IDP Factor 3: Fear				
M	17.5	17.7	18.4	15.7
SD	3.5	3.5	2.9	3.6

[a]$n = 60$. [b]$n = 61$. [c]$n = 62$. [d]$n = 62$.

Note. Imp First = Implicit measure was completed first. Exp First = Explicit measure was completed first.

Factor 1 = Social discomfort; Factor 2 = Empathy; Factor 3 = Fear of having the disability.

APPENDIX B

Pilot Study Informed Consent

(NOTE: DO NOT SIGN THIS DOCUMENT UNLESS AN IRB APPROVAL STAMP WITH CURRENT DATES HAS BEEN APPLIED TO THIS DOCUMENT.)

INFORMED CONSENT
for a Research Study entitled
"The Creation and Rating of Words Representing Persons with Disabilities"

You have been invited to participate in a research study that will be used in the investigation of attitudes toward persons with disabilities. You were invited to participate because you are an undergraduate student currently enrolled at Auburn University. The study is being conducted by Daly Vaughn under the direction of Dr. Adrian Thomas, an Associate Professor of Psychology at Auburn University.

What will be involved if you participate? If you decide to participate in this research study, you will complete a demographics sheet. In addition, you will complete one questionnaire asking you to write down words you believe represent persons with and persons without various disabilities as well as words representing the attribute pleasant and the attribute unpleasant. Next, you will rate words previously created for these categories on their representativeness. You may complete this study in the laboratory space provided for the study located in Thach 102A or Thach 102B.

Are there any risks or discomforts? The risks associated with participating in this study are minimal. You may find creating or rating words about persons with disabilities uncomfortable or distressing.

Are there any benefits to yourself or others? There are no personal benefits to you for participating. However, you will be helping disability researchers to better understand attitudes toward persons with disabilities.

Will you receive compensation for participating? You will earn thirty minutes of extra credit in a psychology class for the completion of this study. Your performance during this study does not affect the amount of extra credit you receive. You may use SONA systems to assign the extra credit to a class.

Are there any costs? If you decide to participate, you will lose only about fifteen to thirty minutes of your time today.

Participant's initials _____ Page 1 of 2

If you change your mind about participating, you can withdraw at any time during the study. Your participation is completely voluntary. If you choose to withdraw, your data can be withdrawn as long as it is identifiable. Your decision about whether or not to participate or to stop participating will not jeopardize your future relations with Auburn University, the Department of Psychology, or the researchers involved.

Your privacy will be protected. Any information obtained in connection with this study will remain anonymous and linked only by an assigned identification number. The ID number will be stored and this number will be used for data analysis purposes only. Information obtained through your participation may be used to fulfill an educational requirement, published in a professional journal, and/or presented at a professional meeting. If so, none of your identifiable information will be presented.

If you have questions about this study, *please ask them now or* contact Daly Vaughn at vaughed@auburn.edu or Adrian Thomas at thomaa6@auburn.edu. A copy of this document will be given to you to keep.

If you have questions about your rights as a research participant, you may contact the Auburn University Office of Human Subjects Research or the Institutional Review Board by phone (334)-844-5966 or e-mail at hsubjec@auburn.edu or IRBChair@auburn.edu.

HAVING READ THE INFORMATION PROVIDED, YOU MUST DECIDE WHETHER OR NOT YOU WISH TO PARTICIPATE IN THIS RESEARCH STUDY. YOUR SIGNATURE INDICATES YOUR WILLINGNESS TO PARTICIPATE.

_____		_____	
Participant's signature	Date	Investigator obtaining consent	Date
_____		_____	
Printed Name		Printed Name	

APPENDIX C

List of Definitions for Various Disability Conditions

List of Definitions for Various Disability Conditions[5]

1) Person with alcoholism: person with addiction to alcoholic beverages.

2) Person with cancer: person with a malignant tumor that tends to spread.

3) Person with mental illness: person with any of various conditions characterized by impairment of the person's normal cognitive, emotional, or behavioral functioning.

4) Person with paraplegia: person that cannot use their lower trunk or legs.

5) Person without the disability: any person not considered as having the disability

[5] Definitions one, two, and four were modified from *Merriam-Webster Pocket Dictionary's* (2006) definitions of the disability. Definition three was modified from *The American Heritage Dictionary's* (2004) definition of mental illness. Definition five was created as a counterpart to each of the disability conditions for the purpose of clarification in the present study.

APPENDIX D

Pilot Study Material

DEMOGRAPHIC DATA **ID # _____**

Please Answer the Following Questions. Keep In Mind that Your Answers Are Completely Confidential.
This Information Will Be Used For Data Analysis Purposes Only.

1. Gender: Male

   Female

2. Race: Caucasian

   African American

   Native American

   Hispanic/Latino

   Asian American

   Other

3. Age: _____ years

INSTRUCTIONS

For the following conditions, please list one-word adjectives/descriptors that you feel represent or are strongly associated with a person with this condition. You should write as many words as come to mind about the condition (try to write down at least five words for each condition). We are interested in studying perceptions of persons with disabilities. Your honest, forthright answers will help us understand what adjectives and descriptors are most strongly associated and best represent persons with and without the various disabilities, so that we might be able to better help employers, rehabilitation psychologists, and persons with disabilities. Your participation in this activity is completely voluntary. All responses will be collected anonymously and you may choose to withdraw at any time.

Person WITH Cancer	**Person WITHOUT Cancer**
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

Continued on Next Page

Person WITH Paraplegia	**Person WITHOUT Paraplegia**

Continued on Next Page

Person WITH Mental Illness	**Person WITHOUT Mental Illness**

Continued on Next Page

Person WITH Alcoholism	**Person WITHOUT Alcoholism**

Continued on Next Page

INSTRUCTIONS: Now please write down descriptors or adjectives that you believe represent or are strongly associated with the attributes of pleasant and unpleasant. These words do not need to be related in any way to persons with disabilities or persons without disabilities. Write down as many words as you can think of that strongly represent pleasant and unpleasant (try to write down at least five words for each attribute).

Pleasant	**Unpleasant**
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

After completing this portion of the study, please inform the research assistant and you will be given the next packet to be completed.

INSTRUCTIONS **ID # _____**

Thank you for completing the previous packet. For the following conditions, please rate the one-word adjectives/descriptors listed under each disability condition on how strongly associated or representative you think it is for the following conditions. We are interested in studying perceptions of persons with disabilities. Your honest, forthright answers will help us understand what adjectives and descriptors are most strongly associated or representative of persons with and persons without the various disabilities, so that we might be able to better help employers, rehabilitation psychologists, and persons with disabilities. Your participation in this activity is completely voluntary. All responses will be collected anonymously and you may choose to withdraw at any time.

Please use the following scale:
1 = Very Unrepresentative
2 = Moderately Unrepresentative
3 = Slightly Unrepresentative
4 = Neutral
5 = Slightly Representative
6 = Moderately Representative
7 = Very Representative

Person with Cancer

	Very Unrepresentative	Moderately Unrepresentative	Slightly Unrepresentative	Neutral	Slightly Representative	Moderately Representative	Very Representative
1. Weak	1	2	3	4	5	6	7
2. Frail	1	2	3	4	5	6	7
3. Sick	1	2	3	4	5	6	7
4. Survivor	1	2	3	4	5	6	7
5. Unfortunate	1	2	3	4	5	6	7
6. Terminal	1	2	3	4	5	6	7

Person without Cancer

1. Healthy	1	2	3	4	5	6	7
2. Happy	1	2	3	4	5	6	7
3. Strong	1	2	3	4	5	6	7
4. Lucky	1	2	3	4	5	6	7
5. Vibrant	1	2	3	4	5	6	7
6. Fortunate	1	2	3	4	5	6	7

(Continued on next page)

Please use the following scale:

1 = Very Unrepresentative
2 = Moderately Unrepresentative
3 = Slightly Unrepresentative
4 = Neutral
5 = Slightly Representative
6 = Moderately Representative
7 = Very Representative

	Very Unrepresentative	Moderately Unrepresentative	Slightly Unrepresentative	Neutral	Slightly Representative	Moderately Representative	Very Representative
Person <u>with</u> Paraplegia							
1. Immobile	1	2	3	4	5	6	7
2. Confined	1	2	3	4	5	6	7
3. Dependent	1	2	3	4	5	6	7
4. Challenged	1	2	3	4	5	6	7
5. Impaired	1	2	3	4	5	6	7
6. Restricted	1	2	3	4	5	6	7
Person <u>without</u> Paraplegia							
1. Mobile	1	2	3	4	5	6	7
2. Independent	1	2	3	4	5	6	7
3. Lucky	1	2	3	4	5	6	7
4. Capable	1	2	3	4	5	6	7
5. Freedom	1	2	3	4	5	6	7
6. Functional	1	2	3	4	5	6	7
Person <u>with</u> Mental Illness							
1. Crazy	1	2	3	4	5	6	7
2. Incapable	1	2	3	4	5	6	7
3. Troubled	1	2	3	4	5	6	7
4. Different	1	2	3	4	5	6	7
5. Confused	1	2	3	4	5	6	7
6. Isolated	1	2	3	4	5	6	7
Person <u>without</u> Mental Illness							
1. Strong	1	2	3	4	5	6	7
2. Clarity	1	2	3	4	5	6	7
3. Functional	1	2	3	4	5	6	7
4. Capable	1	2	3	4	5	6	7
5. Stable	1	2	3	4	5	6	7
6. Adjusted	1	2	3	4	5	6	7

(Continued on next page)

Please use the following scale:

1 = Very Unrepresentative
2 = Moderately Unrepresentative
3 = Slightly Unrepresentative
4 = Neutral
5 = Slightly Representative
6 = Moderately Representative
7 = Very Representative

Person <u>with</u> Alcoholism

	Very Unrepresentative	Moderately Unrepresentative	Slightly Unrepresentative	Neutral	Slightly Representative	Moderately Representative	Very Representative
1. Weak	1	2	3	4	5	6	7
2. Abusive	1	2	3	4	5	6	7
3. Addicted	1	2	3	4	5	6	7
4. Careless	1	2	3	4	5	6	7
5. Immature	1	2	3	4	5	6	7
6. Compulsive	1	2	3	4	5	6	7

Person <u>without</u> Alcoholism

1. Sober	1	2	3	4	5	6	7
2. Stable	1	2	3	4	5	6	7
3. Control	1	2	3	4	5	6	7
4. Dependable	1	2	3	4	5	6	7
5. Disciplined	1	2	3	4	5	6	7
6. Responsible	1	2	3	4	5	6	7

(Continued on next page)

INSTRUCTIONS: Ratings of Attribute Words: Now please rate the following words as to how representative or strongly associated they are to the attributes of pleasant and unpleasant.

Please use the following scale:
1 = Very Unrepresentative
2 = Moderately Unrepresentative
3 = Slightly Unrepresentative
4 = Neutral
5 = Slightly Representative
6 = Moderately Representative
7 = Very Representative

Pleasant

	Very Unrepresentative	Moderately Unrepresentative	Slightly Unrepresentative	Neutral	Slightly Representative	Moderately Representative	Very Representative
1. Champion	1	2	3	4	5	6	7
2. Diamond	1	2	3	4	5	6	7
3. Diploma	1	2	3	4	5	6	7
4. Rainbow	1	2	3	4	5	6	7
5. Sunrise	1	2	3	4	5	6	7
6. Vacation	1	2	3	4	5	6	7

Unpleasant

1. Bomb	1	2	3	4	5	6	7
2. Devil	1	2	3	4	5	6	7
3. Hatred	1	2	3	4	5	6	7
4. Pollute	1	2	3	4	5	6	7
5. Slime	1	2	3	4	5	6	7
6. Poison	1	2	3	4	5	6	7

Thank you for your time! This completes the study.

APPENDIX E

Primary Study Informed Consent

(NOTE: DO NOT SIGN THIS DOCUMENT UNLESS AN IRB APPROVAL STAMP WITH CURRENT DATES HAS BEEN APPLIED TO THIS DOCUMENT.)

INFORMED CONSENT
for a Research Study entitled
"An Investigation of Attitudes toward Persons with Disabilities"

You have been invited to participate in a research study on attitudes toward persons with disabilities. You were invited to participate because you are an undergraduate student currently enrolled at Auburn University. The study is being conducted by Daly Vaughn under the direction of Dr. Adrian Thomas, an Associate Professor of Psychology at Auburn University.

What will be involved if you participate? If you decide to participate in this research study, you will complete several questionnaires and computer based measures, including a demographics form, exposure to persons with disabilities form, two attitude measures, and two other measures investigating variables of interest. You may complete this study in the laboratory space provided for the study located in Thach 102A or Thach 102B.

Are there any risks or discomforts? The risks associated with participating in this study are minimal. You may find questions related to attitudes toward persons with disabilities uncomfortable or distressing. To minimize these risks, we have reserved the right to withhold scores on the two attitude measures from participants.

Are there any benefits to yourself or others? There are no personal benefits to you for participating. However, you will be helping disability researchers to better understand attitudes toward persons with disabilities.

Will you receive compensation for participating? You will earn one hour and thirty minutes of extra credit in a psychology class for the completion of this study. Your performance during this study does not affect the amount of extra credit you receive. You may use SONA systems to assign the extra credit to a class.

Are there any costs? If you decide to participate, you will lose only about one hour and fifteen to thirty minutes of your time today.

Participant's initials _____ Page 1 of 2

If you change your mind about participating, you can withdraw at any time during the study. Your participation is completely voluntary. If you choose to withdraw, your data can be withdrawn as long as it is identifiable. Your decision about whether or not to participate or to stop participating will not jeopardize your future relations with Auburn University, the Department of Psychology, or the researchers involved.

Your privacy will be protected. Any information obtained in connection with this study will remain anonymous and linked only by an assigned identification number. The ID number will be stored and this number will be used for data analysis purposes only. Information obtained through your participation may be used to fulfill an educational requirement, published in a professional journal, and/or presented at a professional meeting. If so, none of your identifiable information will be presented.

If you have questions about this study, *please ask them now or* contact Daly Vaughn at vaughed@auburn.edu or Adrian Thomas at thomaa6@auburn.edu. A copy of this document will be given to you to keep.

If you have questions about your rights as a research participant, you may contact the Auburn University Office of Human Subjects Research or the Institutional Review Board by phone (334)-844-5966 or e-mail at hsubjec@auburn.edu or IRBChair@auburn.edu.

HAVING READ THE INFORMATION PROVIDED, YOU MUST DECIDE WHETHER OR NOT YOU WISH TO PARTICIPATE IN THIS RESEARCH STUDY. YOUR SIGNATURE INDICATES YOUR WILLINGNESS TO PARTICIPATE.

_____		_____	
Participant's signature	Date	Investigator obtaining consent	Date
_____		_____	
Printed Name		Printed Name	

APPENDIX F

Demographics Questionnaire

Demographics Questionnaire ID: _____

DEMOGRAPHIC DATA

Please Answer the Following Questions. Keep In Mind that Your Answers Are Completely Confidential.
This Information Will Be Used For Data Analysis Purposes Only.

1. Gender: Male
   Female

2. Race: Caucasian
   African American
   Native American
   Hispanic/Latino
   Asian American
   Other

3. Age: _____ years

APPENDIX G

Interaction with Disabled Persons Scale

Interaction with Disabled Persons Scale

Here is a list of statements that some people have said describe how they feel when they have contact with a person with a disability. Of course, how we respond to people depends on how well we know them as individuals. However, we would like to know how you feel in general when you meet a person with a disability. Please read each statement carefully and circle how much it describes how you feel.

Please use the following scale:
1 = I disagree very much
2 = I disagree somewhat
3 = I disagree a little
4 = I agree a little
5 = I agree somewhat
6 = I agree very much

	I disagree very much	I disagree somewhat	I disagree a little	I agree a little	I agree somewhat	I agree very much
1. It is rewarding when I am able to help.	1	2	3	4	5	6
2. It hurts me when they want to do something and can't.	1	2	3	4	5	6
3. I feel frustrated because I don't know how to help.	1	2	3	4	5	6
4. Contact with a disabled person reminds me of my own vulnerability.	1	2	3	4	5	6
5. I wonder how I would feel if I had this disability.	1	2	3	4	5	6
6. I feel ignorant about disabled people.	1	2	3	4	5	6
7. I am grateful that I do not have such a burden.	1	2	3	4	5	6
8. I try to act normally and to ignore the disability.	1	2	3	4	5	6
9. I feel uncomfortable and find it hard to relax.	1	2	3	4	5	6
10. I am aware of the problems that disabled people face.	1	2	3	4	5	6

Interaction with Disabled Persons Scale (Continued)

Here is a list of statements that some people have said describe how they feel when they have contact with a person with a disability. Of course, how we respond to people depends on how well we know them as individuals. However, we would like to know how you feel in general when you meet a person with a disability. Please read each statement carefully and circle how much it describes how you feel.

Please use the following scale:
1 = I disagree very much
2 = I disagree somewhat
3 = I disagree a little
4 = I agree a little
5 = I agree somewhat
6 = I agree very much

	I disagree very much	I disagree somewhat	I disagree a little	I agree a little	I agree somewhat	I agree very much
11. I can't help staring at them.	1	2	3	4	5	6
12. I feel unsure because I don't know how to behave.	1	2	3	4	5	6
13. I admire their ability to cope.	1	2	3	4	5	6
14. I don't pity them.	1	2	3	4	5	6
15. After frequent contact, I find I just notice the person not the disability.	1	2	3	4	5	6
16. I feel overwhelmed with discomfort about my lack of a disability.	1	2	3	4	5	6
17. I am afraid to look at the person straight in the face.	1	2	3	4	5	6
18. I tend to make contacts only brief and finish them as quickly as possible.	1	2	3	4	5	6
19. I feel better with disabled people after I have discussed their disability with them.	1	2	3	4	5	6
20. I dread the thought that I could eventually end up like them.	1	2	3	4	5	6

APPENDIX H

Balanced Inventory of Desirable Responding

BIDR V. 6

Using the scale below as a guide, write a number beside each statement to indicate how much you agree with it.

1 _____ 2 _____ 3 _____ 4 _____ 5 _____ 6 _____ 7

NOT TRUE SOMEWHAT TRUE VERY TRUE

_____ 1. My first impressions of people usually turn out to be right.

_____ 2. It would be hard for me to break any of my bad habits.

_____ 3. I don't care to know what other people really think of me.

_____ 4. I have not always been honest with myself.

_____ 5. I always know why I like things.

_____ 6. When my emotions are aroused, it biases my thinking.

_____ 7. Once I've made up my mind, other people can seldom change my opinion.

_____ 8. I am not a safe driver when I exceed the speed limit.

_____ 9. I am fully in control of my own fate.

_____ 10. It's hard for me to shut off a disturbing thought.

_____ 11. I never regret my decisions.

_____ 12. I sometimes lose out on things because I can't make up my mind soon enough.

_____ 13. The reason I vote is because my vote can make a difference.

_____ 14. My parents were not always fair when they punished me.

_____ 15. I am a completely rational person.

_____ 16. I rarely appreciate criticism.

_____ 17. I am very confident of my judgments.

_____ 18. I have sometimes doubted my ability as a lover.

_____ 19. It's all right with me if some people happen to dislike me.

_____ 20. I don't always know the reasons why I do the things I do.

Using the scale below as a guide, write a number beside each statement to indicate how much you agree with it.

1 2 3 4 5 6 7

NOT TRUE SOMEWHAT TRUE VERY TRUE

_____ 21. I sometimes tell lies if I have to.

_____ 22. I never cover up my mistakes.

_____ 23. There have been occasions when I have taken advantage of someone.

_____ 24. I never swear.

_____ 25. I sometimes try to get even rather than forgive and forget.

_____ 26. I always obey laws, even if I'm unlikely to get caught.

_____ 27. I have said something bad about a friend behind his or her back.

_____ 28. When I hear people talking privately, I avoid listening.

_____ 29. I have received too much change from a salesperson without telling him or her.

_____ 30. I always declare everything at customs.

_____ 31. When I was young I sometimes stole things.

_____ 32. I have never dropped litter on the street.

_____ 33. I sometimes drive faster than the speed limit.

_____ 34. I never read sexy books or magazines.

_____ 35. I have done things that I don't tell other people about.

_____ 36. I never take things that don't belong to me.

_____ 37. I have taken sick-leave from work or school even though I wasn't really sick.

_____ 38. I have never damaged a library book or store merchandise without reporting it.

_____ 39. I have some pretty awful habits.

_____ 40. I don't gossip about other people's business.

APPENDIX I

List of Figures

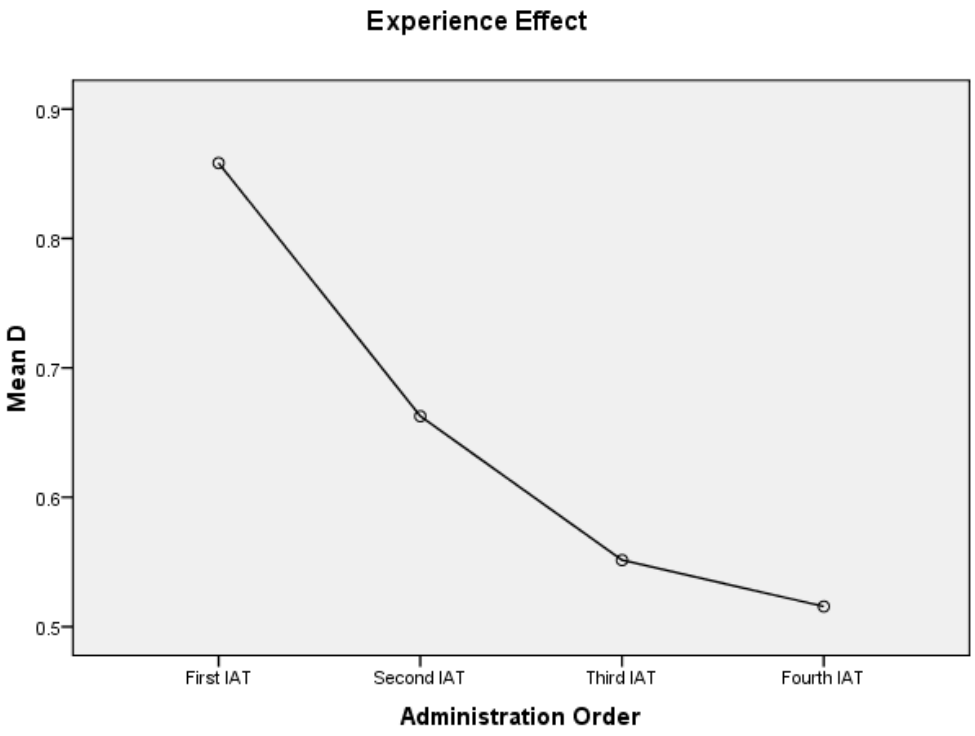

Figure 1. Means plot for individual IAT scores by administration order.

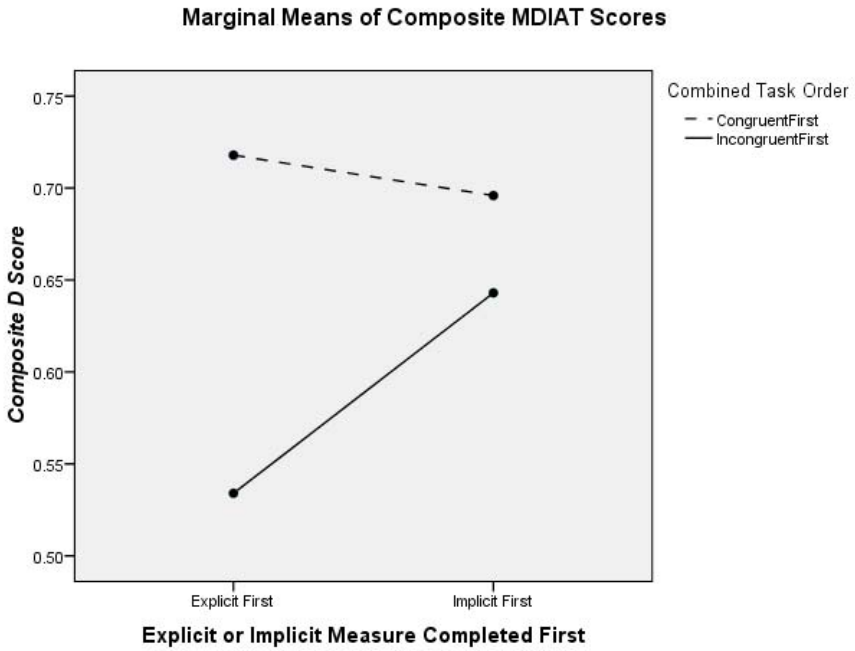

Figure 2. Interaction plot of procedural variables on composite MDIAT scores.