2011

CONSTRUCT VALIDITY OF A LABORATORY AGGRESSION PARADIGM: A MULTITRAIT-MULTIMETHOD APPROACH

Joshua Parker Phillips
University of Kentucky, joshua.p.phillips@gmail.com

ABSTRACT OF DISSERTATION

Joshua Parker Phillips

The Graduate School

University of Kentucky

2011

CONSTRUCT VALIDITY OF A LABORATORY AGGRESSION PARADIGM:
A MULTITRAIT-MULTIMETHOD APPROACH

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Arts and Sciences
at the University of Kentucky

By
Joshua Parker Phillips

Lexington, Kentucky

Director: Dr. Peter R. Giancola, Professor of Psychology

Lexington, Kentucky

2011

ABSTRACT OF DISSERTATION

CONSTRUCT VALIDITY OF A LABORATORY AGGRESSION PARADIGM:
A MULTITRAIT-MULTIMETHOD APPROACH

There continues to be doubt regarding the validity of laboratory aggression paradigms. This paper provides an investigation of the construct validity of one prominent aggression task, the Taylor Aggression Paradigm (TAP), within a Multitrait Multimethod Matrix (MTMM) methodology. Participants consisted of 151 male undergraduate psychology students with a median age of 19 years old (*M*=19.45, *SD* = 2.03). Participants completed self-report and behavioral measures of aggresion, impulsivity, and pro-social behavior which were analyzed using a Correlated Trait – Correlated Method Confirmatory Factor Analysis model. Results supported the construct validity of the MTMM model and the TAP. This study provides one of the only a priori tests of construct validity for the TAP and provides a basis for aditional validation studies using this methodology.

KEYWORDS: Aggression, Construct Validity, Multitrait Multimethod Matrix, Factor Analysis, Taylor Aggression Paradigm

Joshua Parker Phillips
Student's Signature

January 31, 2011
Date

CONSTRUCT VALIDITY OF A LABORATORY AGGRESSION PARADIGM:
A MULTITRAIT-MULTIMETHOD APPROACH

By

Joshua Parker Phillips

Peter R. Giancola, Ph.D.
Director of Dissertation

David T. R. Berry, Ph.D.
Director of Graduate Studies

January 31, 2011

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted , but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Joshua Parker Phillips

The Graduate School

University of Kentucky

2011

CONSTRUCT VALIDITY OF A LABORATORY AGGRESSION PARADIGM:
A MULTITRAIT-MULTIMETHOD APPROACH

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Joshua Parker Phillips

Lexington, Kentucky

Director: Dr. Peter R. Giancola, Professor of Psychology

Lexington, Kentucky

2011

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One: Introduction

Background

Over the past 50 years there has been an ongoing debate in the literature regarding the validity of laboratory aggression paradigms. It has been argued that such tasks do not have sufficient ecological validity compared with observational or naturalistic methodologies (Kane, Joseph, & Tedeschi, 1976; Orne, 1962; Ritter & Eslea, 2005; Tedeschi & Quigley, 1996, 2000). However, there exist no data to support this claim (Anderson & Bushman, 1997). Despite numerous attempts to resolve these arguments, they have persisted throughout the history of this literature. Critics continue to contest the ecological validity of laboratory aggression paradigms (Ferguson, 2007; Gottfredson & Hirschi, 1993; Ritter & Eslea, 2005; Tedeschi & Quigley, 1996, 2000). Although advocates of these tasks have provided data to support their construct validity (for reviews see Anderson & Bushman, 1997; Giancola & Chermack, 1998), the literature is lacking in well-constructed and rigorous investigations of the construct validity of laboratory aggression paradigms.

Defining Aggression

In order to establish construct validity it is important to determine that the behavior assessed is consistent with an established definition of aggression. Therefore, Baron and Richardson's (1994) longstanding definition will be utilized (originally published by Baron in 1977), which stipulates that aggression is: *"any form of behavior directed toward the goal of harming or injuring another living being who is motivated to avoid such treatment"* (Baron & Richardson, 1994; p.7). This definition draws upon historical conceptualizations of aggression (e.g. Buss, 1961) as well as addressing limitations of previous definitions (e.g. definitions of aggression that disregard the perpetrator's intent to harm or the victim's wish to avoid such harm). It is also in keeping with definitions put forth by other contemporary aggression theorists (Anderson & Bushman, 2002; Berkowitz, 1993; Geen, 2001).

The conceptualization and assessment of aggression is often determined by a specific research question or theory in question. Lack of agreement on basic theoretical models has certainly contributed to the diversity of aggression measures (Barrett, Stanford, Kent, & Felthous, 1997). Given the diversity of theoretical models and range of behaviors that can be categorized as aggressive, many researchers have attempted to establish taxonomies of aggression. Early work by Buss (1961) categorized aggressive behaviors amongst three dimensions: (1) direct - indirect, (2) active - passive, and (3) physical - verbal. More recently, several researchers have suggested that subtypes of aggressive behaviors can be described within a more parsimonious taxonomic system that classifies aggression using only two of these dimensions: (1) direct - indirect and (2) active - passive (Parrott & Giancola, 2007; Richardson & Green, 2003). Although there is still no consensus about the use of these terms, Parrott and Giancola (2007) posited that direct aggression encompasses any direct route of aggression that results in the victim's ability to identify the perpetrator (i.e., punching someone), whereas indirect aggression is typified by harm resulting from circuitous actions (e.g., gossip, spreading rumors, ordering a shooting) where the perpetrator is known or unknown to the victim. Moreover, active aggression occurs when there is a behavior directed toward harming whereas passive aggression is defined as a lack of action that results in harm. Other researchers have created alternative taxonomies including: affective/hostile - instrumental (Bandura, 1973; Bushman & Anderson, 2001; Feshbach, 1964; Geen, 1989), impulsive - premeditated (Barratt & Slaughter, 1998; Houston, Stanford, Villemarette-Pittman, Conklin, & Helfritz, 2003), proactive - reactive (Crick & Dodge, 1994; Dodge & Coie, 1987), conscious - impulsive (Berkowitz, 1993), adaptive and pro-social - maladaptive and antisocial (Rosenzeweig, 1941), and overt - covert (Loeber & Schmaling, 1985). The use of a taxonomic system is potentially beneficial in that it provides conceptual clarity in the identification of the type of aggression that laboratory paradigms actually assess. However, there are inconsistencies within the aggression literature when categorizing different types of aggression. A violent behavior could be labeled in numerous ways depending on the theoretical dimensions used to conceptualize the specific behavior. In the absence of agreement on a specific taxonomic system, research on aggression has largely continued without a theoretical delineation of various subtypes of aggression. Instead, researchers have simply labeled various subtypes of interest based on function of the behavior or mechanism of action. Aggression can be conceptualized as a multifaceted construct that can be assessed as a range of behaviors or a specific subtype of those behaviors. As such, numerous subtypes have been identified within the literature (e.g., physical, verbal, relational, manipulative, proactive, reactive, overt, covert, social, hostile, instrumental, etc.) (Berkowitz, 1969; Buss, 1971; Crick & Dodge, 1994; Feshback, 1964; Geen, 1989). Further, animal researchers have identified subtypes commonly found in animal behavior (e.g. predatory, inter-male, fear induced, irritable, territorial, maternal, instrumental, sex-related, etc.) (Archer, 1988;

Moyer, 1976). Given the numerous subtypes aggressive behaviors, it is not surprising that there is no clear consensus on the best method for categorizing these subtypes.

Laboratory Assessment of Aggression

It is important to briefly address the history of the assessment of aggression within psychological laboratories. Early theorists attempted to explain all aggressive behavior as an outcome of frustration. Specifically, that aggression was always the result of frustration (Dollard, Doob, Miller, Mowrer, & Sears, 1939). This theory defined aggression as any *"…sequence of behavior, the goal-response to which is the injury of the person towards whom it is directed"* (Dollard et al., 1939, p.9). One investigation attempted to elicit aggression within the laboratory by inducing frustration using sleep deprivation, periods of enforced silence, and withholding food (Sears, Hovland, & Miller, 1940). Aggression was assessed using self-report as well as observer ratings. It is interesting to note that these early researchers felt that *"The utter failure of the several paper and pencil tests to reflect this aggression in any consistent and objectively measureable way leaves little hope that the problem of momentary aggression can be studied with their aid"* (Sears, Hovland, & Miller, 1940; p.293). However, the aggression observed in these types of studies was verbal in nature; most likely due to social factors discouraging the expression, as well as the measurement of physical aggression. The difficulties in eliciting frustration and violent behavior using this type of methodology prompted investigators to develop alternative means of assessing physical aggression.

The inherent dangers in provoking aggression (i.e. attempted retaliation or retribution by the participant) led scientists to search for safe and reliable ways in which to assess physical aggression within the laboratory. One of the first laboratory aggression paradigms developed as a safe and reliable measure of physical aggression was the *Teacher/Learner Paradigm* (also known as the *Buss Aggression Machine*; Buss, 1961). It can be argued that this paradigm directly impacted all other laboratory aggression tasks that followed. This task required the participant to take the role of a teacher who helped their "partner" (actually a confederate) learn some arithmetic problems. The participant and partner were seated in different rooms. For each trial of the task, the partner had to answer an arithmetic question. If the answer was correct, the participant would then reinforce the partner by illuminating a light on a console. However, if answer was incorrect, the participant would then administer an electric shock to this person as a punishment. Throughout this task, the confederate's responses were predetermined.

Aggression was operationalized as the intensity of the electric shocks administered to the partner. The primary criticism of this task is that responses may have been influenced by altruistic intentions, as the purpose of the task was to "help" one's partner to best answer the arithmetic problems (Baron & Eggleston, 1972). Additionally, inasmuch as the participants never received any electric shocks during the task, they were never provoked to behave in an aggressive manner.

Taylor (1967) then developed the Taylor Aggression Paradigm (TAP), in part, to address the above concerns regarding the *Buss Aggression Machine.* The TAP requires participants to compete against a fictitious opponent on a reaction-time task. They are told that their opponent is seated in a nearby room. Prior to each reaction time-trial, participants select 1 of 10 shock intensities ("1" = low to "10" = high) that they wish to administer to their opponent. A reaction-time trial then follows. In the event that the participant wins the trial, his/her opponent ostensibly receives the selected shock and the winner is provided with feedback indicating which shock his/her opponent had selected for him/her. In the event that the participant loses the trial, s/he receives a shock ostensibly from his/her opponent. In this task, no opponent exists and the sequence of wins and losses as well as the intensity of the shocks set by the "opponent" are predetermined. In order to assess the effect of provocation on aggressive responding, the intensity of shocks received by the participant increases throughout the task. The TAP operationalizes aggression as the average shock intensity selected over trials. Since the development of the TAP, numerous modifications have emerged. Specifically, alterations have been included to allow participants to select an aggressive response immediately after winning a reaction time trial (Giancola & Zeichner, 1995a), rather than prior to either winning or losing. Additional modifications have allowed for a choice to refrain from any retaliation (Zeichner, Frey, Parrott, & Butryn, 1999), the use of a "0" or a no shock option (McCloskey & Berman, 2003), and the measurement of frequency (Berkowitz & LePage, 1967) as well as time duration (Zeichner & Pihl, 1979) of shocks selected. Others have even substituted the use of noxious tone blasts in place of electric shocks (Bushman, Baumeister, & Phillips, 2001). These variants allow researchers to build upon the theoretical framework for the measurment of aggression established by the TAP.

Numerous other laboratory aggression tasks do exist. Two of the other most prominent paradigms include the Point Subtraction Aggression Paradigm (PSAP) and the Hot Sauce Task

(HST). The PSAP was developed by Cherek (1981) and requires participants to either press one button to earn points or a second button to subtract points from a fictitious opponent. The dependent measure of aggression is the number of times the point subtraction button is pressed. The HST (Lieberman et al., 1999) is a newly developed, and increasingly used (Ayduk, Gyurak, & Luerssen, 2008; Evers, Fischer, Mosquera, & Manstead, 2005; Lieberman et al., 1999; McGregor et al., 1998; Warburton, Williams, & Cairns, 2006), aggression task. This task takes place under the guise of a food tasting study. After tasting and rating a food sample, they are informed that they will prepare a food sample for a fictitious participant that has previously provoked them. Aggression is operationalized as the amount of hot sauce the participant chooses to allocate to his/her ostensible provocateur. These tasks are mentioned in order to place the TAP in the proper context. The purpose of this paper is not to examine the construct validity of all laboratory aggression paradigms, but focus on the TAP in particular. The continued use of the TAP throughout the aggression literature suggests that the psychological community would significantly benefit from a systematic examination of its construct validity. As such, this paper will review the existing evidence for the construct validity of the TAP and attempt to provide a unique contribution to that literature.

Construct Validity

Construct validity refers to the process of theory evaluation and testing. Specifically, psychological constructs must be inferred from the method used to evaluate them. The primary concern of testing construct validity is the extent to which the method of assessment reflects the psychological constructs it is intended to measure. Early researchers noted that psychological constructs are by nature unobservable mechanisms that influence the observable (MacCorquodale & Meehl, 1948). Therefore, there was a need to establish a method for some objective way of measuring unobservable constructs. One of the first attempts was criterion-related validity. Criterion-related validity refers to the degree to which a measure is able to predict a criterion (Cureton, 1950). Ultimately criterion-related validity was limited by the validity of the assessment of the criterion. The measure could only be valid to the extent that the criterion was valid.

Early works on criterion-related validity lead to Cronbach & Meehl's (1955) seminal paper which presented a methodology to accumulate evidence for the validity of a measure. They postulated that constructs are developed when there is no quantifiable measure that can capture the totality of the behavior in question. Cronbach and Meehl (1955) emphasized that construct validity was established through predictions of the theoretical relations between psychological constructs. That is, construct validity is developed through the degree to which the measure relates to other measures of psychological constructs in a theoretically-consistent manner. However, the construct validity of a given measure is always open to criticism and reevaluation. New research allows for continuing support or nonsupport of construct validity. Recently, Smith (2005) presented a five-step model for assessing construct validity that emphasizes the need for theory-based inferences and careful articulation of those inferences. This model stipulates that the scientist must: 1) clearly specify the theoretical construct in question, 2) describe how theory leads to an informative hypothesis, 3) select an appropriate research design/approach, 4) articulate how observations relate to predictions, and 5) revise the theory as well as the construct as necessary. This model allows for a reevaluation of construct validity in the event that the theoretical basis for the construct changes over time. Further, this approach also stresses the fact that establishing construct validity is a process.

In an attempt to provide a comprehensive framework for establishing construct validity, Campbell and Fiske (1959) introduced the multitrait multimethod matrix (MTMM). This paper had a significant impact on validation research primarily through its emphasis on establishing convergent and discriminant validity. Convergent validity can be defined generally as the degree to which a measurement instrument correlates with other measures of the same construct. Discriminant validity refers to the lack of a significant relation between the instrument and other constructs that are theoretically unrelated or distinct. Both convergent and discriminant validity is required to validate a psychological measure, following from Chronbach & Meehl's (1955) assertion that validity of a measure is established through consistency in relations between theoretical constructs . Evidence accumulates over time to demonstrate a pattern of relations between measures. If this pattern of relations is consistent with the theorized pattern between the psychological constructs each measure represents, the measure is determined to have good construct validity.

Construct Validity Evidence

This section will review the existing evidence for the construct validity of the TAP. Although there have been relatively few direct tests, there is substantial indirect evidence for the construct validity of the TAP. Indirect evidence of construct validity can be accumulated by demonstrating that well established relations between key variables and aggression in real-world settings can be extended into the laboratory. For example, it is well known that acute alcohol intoxication is related to interpersonal violence (e.g., Fals-Stewart, 2003; Murphy, Winters, O'Farrell, Fals-Stewart, & Murphy, 2005). Consequently, showing that alcohol also increases aggression in the laboratory will help establish the construct validity of the TAP. Below, we will review how a number of variables, known to be related to aggression in real-world settings, are related to performance on the TAP.

Face Validity Data

Although there are some limitations with the face validity of button presses on the TAP, there is some evidence for the face validity of the use of electric shocks in instances of real world aggression. Specifically, the use of a stun gun has been documented in cases of child abuse (Frechette & Rimsza, 1992), infant homicide (Turner & Jumbelic, 2003), and workplace violence (Feldmann & Johnson, 1998). While the intensity of the shocks administered within the TAP vary, some versions of the TAP employ a #20 button as a response option. Participants are informed that this shock intensity will physically injure their opponent. The use of a #20 button increases the face validity of the task by providing participants with the ability to administer a shock intensity that corresponds to those used in real world instances of aggression. In addition, unpublished anecdotal reports suggest that participant behavior during the TAP is often characterized by: angry yelling, use of obscenities, and rude gestures.

Convergent Validity Data

Convergent validity for laboratory aggression tasks has been demonstrated via positive correlations with self-report measures of aggression. Specifically, shock selections on the TAP have been significantly related to self-report measures of dispositional aggressivity, including physical assault (Anderson, Buckley, & Carnagey, 2008; Bushman, 1995; Giancola & Zeichner, 1995a; Giancola & Parrott, 2008; Terrell, Hill, & Nagoshi, 2008), scores on the combined Assault, Indirect, Verbal, and Negativism subscales of the Buss Durkee Hostility Inventory (BDHI) (an index of behavioral hostility) (Hammock & Richardson, 1992), and verbal aggression (Giancola & Parrott, 2008). In addition to self reported aggression, convergent validity for the TAP has also been demonstrated by positive associations with measures of permissive beliefs about aggression (Giancola & Parrott, 2008), propensity to commit rape (Malamuth & Ceniti, 1986), and sexual arousal towards rape (Malamuth, 1983). Aggression as measured by the TAP in early adulthood is accounted for by self-report frequency of physical violence since age 15 (Phillips & Giancola, 2009). Further, self-report alcohol consumption during those incidents of physical violence predicted aggressive responses on the TAP for intoxicated but not sober participants.

Discriminant Validity Data

Discriminant validity for the TAP has been established via a lack of relations between shock intensity selections and measures of guilt, suspicion, resentment, inwardly directed anger, and indirect hostility (Giancola & Zeichner, 1995a). Furthermore, TAP shock selections have been found to be unrelated to pro-social behaviors like helping or competition (Bernstein, Richardson, & Hammock, 1987).

Known-Groups Comparisons

Another approach that can be used to establish the construct validity of laboratory aggression tasks is to determine whether they can differentiate between violent and non-violent individuals. Others have shown that violent incarcerated offenders exhibited greater levels of aggression on the TAP than their non-violent incarcerated counterparts (Wolfe & Baron, 1971). Due to the fact that parolees and prison inmates were used in these latter studies, they provide strong evidence that the TAP assess an intrinsic component of "violence" (i.e., violent parolees and violent inmates) rather than, or in addition to, a more general trait of "non-violent criminality." Finally, a recent study found that men with a genetic marker associated with violence were more aggressive on the TAP than men without this marker (Verona, Joiner, Johnson, & Bender, 2006).

Gender Differences

Crime statistics indicate that men engage in more violent behavior than women (U.S. Department of Justice, 2008). However, empirical investigations suggest that gender differences in aggression may be smaller than expected and not always consistent across studies (for reviews see Bettencourt & Miller, 1996; Eagly & Steffen, 1986; Frodi, Macaulay, & Thome, 1977; Harris, 1996; Hyde, 1984; Richardson & Hammock, 2007; Verona & Vitale, 2006). A recent meta-analytic review found that young adult males engage in significantly more aggressive behaviors than females (Archer, 2004). However, Richardson (2005) argues that gender differences in aggression have been overestimated. Recent research indicates rising rates of violence among adolescent girls (Odgers et al., 2007) and declining gender differences in serious violent behavior over time (Graves, 2007). Although the exact nature of gender differences in overall aggressive behavior is unclear, there do appear to be male-female differences in the type of aggression employed. Men exhibit greater levels of direct physical aggression compared with women whereas there appear to be no gender differences with respect to indirect aggression (reviewed in Richardson & Hammock, 2007). Similarly, females are more likely than males to engage in relational aggression, whereby a person attacks another via damaging their peer relationships (Crick & Grotpeter, 1995; Vaillancourt, Miller, Fagbemi, Côté, Tremblay, 2007). A similar pattern of findings has been replicated in several studies using the TAP. Although men typically display greater levels of aggression on laboratory tasks under low provocation, there appear to be fewer, if any, differences under high provocation (Bettencourt & Miller, 1996). In addition, men have been found to administer more frequent and greater shock intensities whereas women were more likely to wait longer before administering a shock as well as shocking for longer durations (Giancola et al., 2002; Giancola & Zeichner, 1995b; Zeichner, Parrott, & Frey, 2003). Males have also been found to behave more aggressively towards other males whereas females have been shown to display similar levels of aggression toward both genders (Giancola & Zeichner, 1995b). Similarly, gender differences in aggression on a modified version of the TAP appear to be moderated by whether or not the participant anticipated meeting their competitor (Terrell, Hill, & Nagoshi, 2008). In this particular study women displayed less aggression when they believed they would later meet their opponent. Overall, these findings are consistent with meta-analytic reviews of gender differences in aggression demonstrating that on the TAP males and female both exhibit aggressive behavior, but vary in the type of aggression most often employed.

Neuropsychological Functioning

Brain damage to the prefrontal cortex has been associated with increased aggression (reviewed in Hawkins & Trobst, 2000; Morgan & Lillienfeld, 2000; Siever, 2008). Persons who have suffered damage to the prefrontal cortex are more aggressive than persons without such damage or persons with brain injuries in different locations (Grafman et al., 1996). Furthermore, performance on neuropsychological measures of prefrontal cortical functioning has been negatively correlated with aggression in psychopaths (Kiehl et al., 2001), prison inmates

(Coccaro & Siever, 2005), and delinquent children (Boes, Tranel, Anderson, & Nopoulos, 2008). Given these data, the fact that poor performance on neuropsychological measures of prefrontal functioning in normal individuals has been associated with higher TAP shock selections (Giancola, 2004; Giancola & Zeichner, 1994; Hoaken, Assaad, & Pihl, 1998; Lau, Pihl, & Peterson, 1995) adds to the construct validity of this particular task. Recent neuroimaging studies have also demonstrated activation of the prefrontal cortex during participation in a modified version of the TAP (Lotze, Veit, Anders, & Birbaumer, 2007; Veit et al., 2010).

Neurochemical Effects

Research has also suggested an inverse relation between serotonin levels in the central nervous system (CNS) and impulsive aggression (Coccaro, 1989, 1998). Serotonin is hypothesized to exert inhibitory control over aggressive behavior. Individuals with a violent history have been shown to have lower levels of blood plasma and cerebrospinal fluid serotonin metabolites than individuals without violent histories (Brown, Goodwin, Ballenger, Goyer, & Major, 1979; Linnoila, Virkkunen, & Higley, 1993; Virkkunen, Nuutila, Goodwin, & Linnoila, 1987; Zhou et al., 2006). Serotonin levels are also negatively correlated with cortisol, a hormone associated with stress response (Deakin & Graeff, 1991). Cortisol levels increase in situations typically found to be stressful such as public speaking, anticipation of the death of a family member, or surgery (Parker, 1989). Cortisol serves to decrease the availability of tryptophan, the amino acid precursor to serotonin, thereby reducing CNS serotonin levels. Populations high in aggression have been found to exhibit low cortisol levels (Coccaro & Siever, 2005). Some researchers have theorized that individuals that frequently engage in aggression become less stressed during these instances, leading to a blunted cortisol response (van Goozen, Fairchild, Snoek, & Harold, 2007). Further, individuals that report higher levels of aggression display a blunted cortisol response, even following the administration of a serotonergic agonist (Buydens-Branchey, Branchey, Fergeson, Hudson, & McKernin, 1997; Coccaro, Kavoussi, Cooper, & Hauger, 1997; Coccaro, Kavoussi, & Hauger, 1995). There is evidence to suggest that in individuals that are more likely to be more reactive to situational stressors, high cortisol levels have been significantly associated with increased aggression as measured by a version of the TAP (Berman, Gladue, & Taylor, 1993). In addition, provocation and subsequent aggression on a modified version of the TAP is significantly related to an increase in cortisol levels (Bohnke, Bertsch, Kruk, & Naumann, 2010). Studies using laboratory aggression paradigms have found a similar relation between serotonin functioning and aggression. Lowered serotonin functioning manipulated via dietary depletion of tryptophan has been shown to lead to increased aggressive behavior as measured by modified versions of the TAP (Pihl, Young, Harden, & Plotnick, 1995).

Testosterone Effects

A large body of literature documents a positive relation between testosterone levels and aggression in both humans (reviewed in Archer, 1991) and animals (reviewed in Ellis, 1986). Researchers have found elevated testosterone levels in individuals with a history of violent behavior (Virkkunen, Rawlings, Tokola, & Poland, 1994), psychopathic traits (Stålenheim, Eriksson, von Knorring, & Wide, 1998), and a clinical diagnosis of Antisocial Personality Disorder (Aromäki, Lindman, & Eriksson, 2002). Research utilizing laboratory aggression tasks has found that individuals with higher levels of saliva testosterone behaved more aggressively than those with lower levels on a modified version of the TAP (Berman, Gladue, & Taylor, 1993).

Effects of Psychoactive Drugs

Research indicates that although there is no direct pharmacological link between substance use and aggression, there is a consistent association between substance use and aggressive behaviors (reviewed in Fagan, 1990). However, not all psychoactive drugs effect aggression in the same way. For example, alcohol has been implicated in approximately 50% of violent crimes worldwide (Murdoch, Pihl, & Ross, 1990; Pernanen, 1991). There have been numerous investigations of the relation between acute alcohol consumption and shock selections on the TAP. Findings from these studies overwhelmingly demonstrate that acute alcohol intoxication increases aggression compared with placebo or non-alcohol beverages (reviewed in Bushman & Cooper, 1990; Chermack & Giancola, 1997; Kelley & Cherek, 1993; Taylor & Chermack, 1993). Similarly, self-report data suggest that frequency of cocaine use is associated with a greater likelihood of engaging in aggressive behaviors (e.g. Macdonald, Erickson, Wells, Hathaway, & Pakula, 2008). In addition, a recent meta-analytic review suggests that cocaine has a stronger relation to intimate partner violence than any other illicit substance (Moore et al., 2007). An investigation of the effects of orally administered cocaine on aggression using the TAP concluded that participants that received a high dose of cocaine behaved more aggressively than those that received a placebo or low dose (Licata, Taylor, Berman, & Cranston, 1993). In contrast to alcohol and cocaine, reviews of the literature indicate that the use of marijuana does not lead to subsequent aggressive behaviors (Abel, 1977; Zimmer & Morgan, 1997). In fact, studies utilizing the TAP have demonstrated that administration of Δ^9-tetrahydrocannabinol (THC), the active compound in marijuana, suppressed aggressive behavior (Myerscough & Taylor, 1985; Taylor et al., 1976). Myerscough and Taylor (1985) noted that under intense provocation, participants who received moderate or high doses of THC responded non-aggressively throughout the entire TAP protocol. It is hypothesized that the sedative effects of THC elicit feelings of passivity which inhibits aggressive behavior. It should be noted that CNS depressants often used as sedatives can increase aggressive behavior. Besides alcohol, benzodiazepines are CNS depressants (Julian, 1978) that have been shown to elicit aggression in the laboratory. Benzodiazepines are most commonly prescribed as an anxiolytic, but also have the potential to disinhibit behavior and impair judgment (Taylor & Chermack, 1993). These properties are hypothesized to increase the likelihood of aggressive behavior. Laboratory investigations utilizing the TAP have demonstrated that both triazolam and diazepam increase aggressive responding compared to placebo (Ben-Porath & Taylor, 2002; Berman & Taylor, 1995). However, not all CNS depressants increase aggressive behavior. It has been hypothesized that depressant drugs that do not have anxiolytic effects are less likely to increase aggressive behavior (Taylor & Hulsizer, 1998). Barbiturates are a class of CNS depressants that do not reduce anxiety. Investigations using the TAP indicate that two common barbiturates (secobarbital and phetobarbital), do not affect aggressive behavior, even at relatively high doses (Chermack & Taylor, 1993).

Effects of Provocation

Provocation is arguably one of the most powerful elicitors of aggressive behavior (Anderson & Bushman, 2002; Bettencourt & Miller, 1996, Geen, 2001). One advantage of laboratory aggression paradigms is their ability to manipulate provocation. The TAP and its modifications manipulate provocation by increasing or decreasing the intensity of the shocks administered to the participant. Studies using the TAP have demonstrated that provocation has a greater aggression-potentiating effect than alcohol or gender (Giancola et al., 2002; Giancola and Zeichner, 1995b; Hoaken and Pihl, 2000; Hoaken, Campbell, Stewart, & Pihl, 2003).

Summary of Validity Evidence

Taken together, there is significant evidence for the construct validity of the TAP. However, the primary evidence for the construct validity of the TAP comes in the form of convergent validity data. Given that the discriminant validity data is limited (e.g. Bernstein, Richardson, & Hammock, 1987; Giancola & Zeichner, 1995a), the evidence for the construct validity of the TAP would benefit from further research in this area. In terms of developing the construct validity of the TAP, the clear strength is the well established relations to other variables of interest. However, the numerous studies that provide this indirect support also indicated that there have been relatively few studies designed to specifically test the construct validity of the TAP.

Multitrait Multimethod Matrix

As mentioned above, careful specification of theoretical constructs is a crucial step in establishing construct validity. Within the literature, the conceptualization and definition of aggression has been refined over time (see Anderson & Bushman, 2002; Baron & Richardson, 1994; Berkowitz, 1993; Geen, 2001). However, specification of a theoretical construct alone is not sufficient to demonstrate that an instrument assesses the hypothesized construct. In addition to convergent and discriminant validity, Campbell and Fiske (1959) also noted the importance of variance in measurement. They postulated scores on a psychological measure not only reflect variance in the construct of interest but also variance due to measurement method and error. Error variance is present in any measure of an unobservable psychological construct. Method variance is any variance present in an observed score that is not due to error or part of the construct being measured. Instead, method variance is variance due to the way in which the construct is assessed. For example, method variance can be conceptualized as the degree to which two self-report measures are correlated due to similarity of processes involved in assessment (such as the metric of responses or format of questions). As such, validation studies should account for method variance. In order to account for method variance Campbell and Fiske (1959) introduced the Multitrait Multimethod Matrix (MTMM) methodology. MTMM methodology consists of concurrent examination of two or more traits assessed with at least two different methods. The MTMM methodology has the advantage of assessing convergent validity, discriminant validity, and method variance simultaneously. Researchers have noted that potential sources of method variance include both mode (e.g. self-report, observation, interview, etc…) and setting (e.g. laboratory, inpatient, outpatient, etc…) (Burns & Hayes, 2006). As such, the MTMM methodology could provide evidence for convergent and discriminant validity, as well as evaluate the impact of method variance on aggression measured within the laboratory.

Traditionally, MTMM data has been analyzed using a correlation matrix. However, researchers have noted that there are significant problems with relying on the traditional correlation comparisons to establish convergent and discriminant validity (see Reichardt & Coleman, 1995). There are several Structural Equation Modeling (SEM) approaches that can be utilized to analyze MTMM data. Reviews of this literature have noted that Confirmatory Factor Analysis (CFA) approaches are most consistent with the theoretical basis for the MTMM correlation matrix (see Kenny & Kashy, 1992; Lance, Noble, & Scullen, 2002; Schmitt & Stults, 1986; Widaman, 1985; 1992). Of the various CFA approaches to modeling MTMM data, the Correlated Trait – Correlated Method (CT-CM) model is an accurate representation of Campbell and Fiske's (1959) MTMM matrix (Marsh & Grayson, 1995; Widaman, 1985). The CT-CM model consists of the hypothesized number of trait and method factors. All of the trait factors and all of the method factors are allowed to correlate amongst themselves. However, the CT-CM model assumes that the trait and method factors are uncorrelated. As such, the CT-CM model makes it possible to estimate the proportion of variance in an observed variable due to trait, method, and residual components. This permits a direct and unbiased estimation of trait and method variance components. Although it is possible for the CT-CM model to return inadmissible solutions (see Kenny & Kashy, 1992), the advantages of the CT-CM model have lead to the recommendation that the CT-CM CFA approach be used as the preferred analysis method for MTMM data (i.e. Becker & Cote, 1994; Lance, Noble, & Scullen, 2002).

Some investigators have raised concerns regarding how aggression as measured by laboratory aggression tasks corresponds to "real world" aggression (Ritter & Eslea, 2005; Tedeschi & Quigley, 1996). Specifically, it has been claimed that the behaviors observed on the TAP do not generalize to actual violent behaviors. As such, it is important to address this issue within the construct validation process. The MTMM methodology is particularly well suited to help address this criticism. This claim is an empirical question that can be investigated by evaluating the construct validity of the task. Moreover, the possibility that aggression as measured by the TAP is substantively different than instances of "real world" aggression can be evaluated by the magnitude of method variance present in TAP scores.

Therefore the purpose of this study is to evaluate the construct validity of the TAP within the MTMM methodology. As MTMM methodology requires the assessment of two or more constructs using two or more measures, two related constructs; impulsivity and pro-social behavior will also be assessed. Impulsivity has been hypothesized to be a possible underlying process of aggression and is positively correlated with violent behaviors (Bowman, 1997; Brown, Kent, Bryant, Gevedon, 1989; Horesh, Gothelf, Ofek, Weizman, & Apter, 1997; Plutchik & Van Praag, 1989). Unsurprisingly, pro-social behavior and related constructs have been shown to be negatively correlated with aggression (Eron & Huesmann, 1984; Feshbach & Feshbach, 1969; Romano, Tremblay, Boulerice, & Swisher, 2005). Aggression, impulsivity, and pro-social behavior will be assessed using self-report measures and behavioral tasks. This methodology provides means to evaluate the presence and magnitude of method variance in the TAP. Additionally, it also presents an opportunity to establish additional evidence of both convergent and discriminant validity.

It is hypothesized that the TAP will demonstrated evidence for both convergent and discriminant validity. Specifically, the traditional MTMM correlation matrix is hypothesized to reveal evidence of the convergent validity of the TAP through significant positive correlations between the TAP and self-report measures of aggression. Further, the SEM analytic techniques are predicted to provide evidence of the convergent validity of the TAP through significantly better fit indices of the hypothesized CT-CM CFA model than alternative models. It is also predicted that the factor loadings on the observed TAP scores will provide additional evidence of convergent validity. The TAP is also hypothesized to demonstrate evidence of discriminant validity through negative correlations with self-report and behavioral measures of pro-social behavior. Similarly, the TAP is predicted to display positive correlations with self-report and behavioral measures of impulsivity. More importantly, it is predicted that the hypothesized CT-CM CFA model will display evidence of discriminant validity through comparisons with alternative models and the correlations between aggression, impulsivity, and pro-social traits. Finally, it is predicted that the SEM analysis will reveal the presence of shared method effects for both self-report and behavioral measures. However, it is predicted that these method effects will differ between self-report and behavioral measures due to the differences in mode of assessment.

Chapter Two: Method

Participants

Participants consisted of 151 male undergraduate psychology students. Participants were recruited through PSY 100 courses at the University of Kentucky. Participants received course credit as compensation for participation in this study. Given that this study is one of the first direct investigations of the construct validity of the TAP, this study examined men only. Since research suggests that men exhibit greater levels of direct physical aggression compared with women (reviewed in Richardson & Hammock, 2007), it was hypothesized that the TAP would demonstrate the largest relations to impulsivity and pro-social behavior for male participants. However, it should be noted that the literature suggests that the TAP demonstrates the same relation to other variables with women, only to a lesser degree (e.g. Bettencourt & Miller, 1996). The median age of participants was 19 years old (*M*=19.45, *SD* = 2.03). The majority, 88.1%, of participants identifed as White or Caucasian, 7.9% identified as African-American, 0.7% identified as Hispanic, 2.6% idenitified as Asian, and 0.7% identified as an other race. Other demographic indicators revealed that all of the participants were never married, 99% had a high-school degree and were working on a bachelor's degree. Furthermore, 18% supported themselves financially and earned on average approximately $20,000 per year; the remainder were supported by a parent or family member.

Measures

Self-Report Aggression Questionnaires.

The *Aggression Questionnaire (AQ;* Buss & Perry, 1992*)* was utilized to assess self-report aggressive behavior. The AQ consists of 29 items rated on a five-point Likert-Type scale ranging from 1 (never or hardly applies to me) to 5 (very often applies to me). The AQ is comprised of four subscales; physical aggression, verbal aggression, anger, and hostility. The anger subscale assesses the emotional or affective component of aggressive behavior. The hostility subscale assesses the cognitive component of aggressive behavior. The physical and verbal subscales assess the instrumental or motor component of aggressive behavior. As such only the physical and verbal subscales were included in the final analysis. The AQ is well validated in college and adult samples (Bernstein & Gesn, 1997; Bryant & Smith, 2001; Buss & Perry, 1992; Felsten & Hill, 1999; Gerevich, Bácskai, & Czobor, 2007; Harris, 1995; Harris & Knight-Bohnhoff, 1996). Previous reliability estimates for the four subscales range from .72 to .80 (Buss & Perry, 1992). The *Richardson Conflict Response Questionnaire* (RCRQ; Richardson & Green, 2003) was also utilized to assess self-report aggressive behavior. The RCRQ consists of 28 items rated on a 5-point Likert-Type scale that assess the frequency of engaging in specific behaviors. The RCRQ is comprised of two subscales; direct and indirect aggression. The direct aggression subscale assesses aggressive behaviors where the target of the violent behavior can identify the perpetrator. In contrast, the indirect aggression subscale assesses aggressive actions that effect the target without knowledge of the perpetrator. The RCRQ has displayed good internal validity, with Cronbach alphas ranging between .77 - .91 for direct aggression and .80 - .84 for indirect aggression (Richardson & Green, 2003). In order to maintain consistency in the specific type of aggression examined in this study, only the direct aggression subscale was utilized in the data analyses.

Behavioral Aggression Task.

A modified version of the *Taylor Aggression Paradigm* (TAP; Taylor, 1967) was used to measure aggressive behavior. This task placed participants in a situation where electric shocks were administered to, and received from, a fictitious opponent under the guise of a competitive reaction-time task. Physical aggression was operationalized as the shock intensities selected by the participants. Participants were seated at a table in a small room. This room contained a table with a computer monitor and keyboard. White adhesive labels marked "1" through "10" were attached to the number keys running across the top of the keyboard. The labels "low," "medium," and "high" were placed above keys "1," "5," and "10," respectively, to indicate the subjective levels of shock corresponding to the number keys. Following a winning trial and pressing a shock button, participants were able to view their shock selection on a specially designed "volt meter" on the computer screen and by the illumination of one of 10 "shock lights" on the computer screen. Both of these indicators displayed readings that correspond with the selected shock level. These images were used to reinforce participants' belief that they are actually administering shocks. Upon losing a trial, participants received a shock and received feedback regarding the level of that shock in the form of a signal on the volt meter and the illumination of one of the 10 "shock lights" on the computer screen. Participants were informed that they have a choice of 10 different shock intensities to administer at the end of each winning trial for a duration of their choosing.

The TAP consisted of a total of 34 trials. All shocks delivered to the participants were of a one second duration. In actuality, reaction-times were not measured; the competitive task was used to lead participants to believe that they are engaged in an adversarial interaction with another individual. The win/loss sequence was predetermined and controlled by the computer program that executes the task. The sequence was presented in a fixed-random order with no more than three consecutive wins or losses. The trials were interspersed by five second intervals. The initiation of trials, administration of shocks to the participants, and the recording of their responses were controlled by a computer. The experimenter, other electronic equipment, and the computer that controls the task were located in an adjacent room out of the participant's view.

Self-Report Pro-social Questionnaires.

The *Pro-social Tendencies Measure (PTM;* Carlo & Randall, 2002*)* was utilized to assess self-report pro-social behavior. The PTM was designed for use in college age populations and consists of 23 items rated on a 5-point Likert-Type scale ranging from 1 (does not describe me at all) to 5 (describes me greatly). The PTM was developed to assess pro-social behavior based on items from previously developed pro-social disposition and behavioral scales. Initial validation of the PTM resulted in adequate model fit coefficients using confirmatory factor analysis as well as evidence of convergent validity through correlations with other measures of pro-social behavior (Carlo & Randall, 2002). The PTM assess six subtypes of pro-social behavior: public, anonymous, dire, emotional, compliant, and altruism. Previous reliability estimates for the six subscales range from .75 to .80 (Carlo & Randall, 2002). For the purposes of this study, the anonymous and altruistic subscales were included in the data analyses. The *Self-Report Altruism Scale* (AS; Rushton, Chrisjohn, & Fekken, 1981) was also utilized to assess self-report pro-social behavior. The AS consists of 20 items where responders were asked to rate how often they engaged in various pro-social behaviors ranging from 1(not at all) to 5 (very often). The AS has demonstrated validity through correlations with measures of social responsibility and emotional empathy, as well as reliability estimates of .89 (Rushton, Chrisjohn, & Fekken, 1981).

Behavioral Pro-social Task.

The *Prisoner's Dilemma Game* (PDG; Rapoport & Chammah, 1965) is a laboratory measure of pro-social behavior wherein participants play a non-zero sum game with two response options: cooperation or defection. Participants were informed that they would take part in a strategy game with another male participant. In actuality, participants were not playing with an actual person, but rather with a computer program. In this task, choosing to cooperate increases the possibility of exploitation but maximizes mutual gain. In contrast, choosing to defect protects against exploitation and maximizes the possible individual gain. However, if both "players" chose to defect, they both lost. Participants were informed that the strategy game consisted of four possible payouts determined by each participant's choice. The payout matrix was as follows: when both chose to cooperate, each participant received four points; when both chose to defect, each participant lost two points; when one participant chose to defect and the other chose to cooperate, the defector received eight points and the cooperator lost five points. Participants played a total of ten trials with the computer. In order to ensure variability in responding, the computer was set to defect on the first, fifth, and ninth trials. For all other trials, the computer utilized a tit-for-tat response strategy of mimicking the player's response from the previous trial. Pro-social behavior was operationalized as number of times the participant chooses to cooperate.

Self-Report Impulsivity Questionnaire.

The *Revised UPPS Impulsive Behavior scale (UPPS-P;* Cyders & Smith, 2007*)* was utilized to assess self-report impulsive behavior. The UPPS-P is a revised version of the original UPPS impulsive behavior scale developed by Whiteside and Lynam (2001). The UPPS-P consists of 58 items rated on a four-point Likert-Type scale ranging from 1 (agree strongly) to 4 (disagree strongly). The UPPS-P assesses five personality pathways to impulsive behaviors. These pathways are: Negative Urgency, Positive Urgency, Lack of Premeditation, Sensation Seeking, and Lack of Perseverance. These five subscales have demonstrated good convergent and discriminate validity (Cyders & Smith, 2007; Smith et al., 2007). Previous reliability estimates for the five subscales range from .80 to .94 (Cyders & Smith, 2007). Based on the conceptualization of impulsivity as measured on the task described below several subscales were excluded from the analyses. Due to the lack of strong positive and negative emotions in the laboratory, and the short time period, only the Sensation Seeking and Lack of Premeditation were included in the final analysis. The *Barratt Impulsiveness Scale-11* (BIS-11; Patton, Stanford, & Barratt, 1995) was utilized to assess self-report impulsive behavior. The BIS-11 consists of 30 items on a 4-point Likert-Type scale that ranges from 1 (rarely/never) to 4 (almost always/always) This scale has three second-order factors: attentional impulsivity, motor impulsivity, and non-planning impulsivity. These factors correspond to inability to tolerate cognitive complexity, tendency to act impetuously, and a lack of sense of the future respectively. The BIS-11 has yielded reliability estimates in this population of .82 (Patton, Stanford, & Barratt, 1995).

Behavioral Impulsivity Task.

Participants completed a computerized version of the *Go/No-Go* task (Newman & Kosson, 1986). Participants were informed that a series of numbers would be presented, one at a time, in the center of a computer screen. Participants were informed that they have an opportunity to win points on the basis of their performance on the task. Participants were informed that each time a number appeared on the screen, they would have to choose whether or not they are going to press the spacebar on the keyboard. Further, they were told that their choice would result in either winning or losing points. They were given no further instructions. Prior to beginning the task, participants were informed that they would begin with 20 points. They were informed that each time they won or lost a trial, the experimenter would respectively give or take away a point from the participant. Participants did not win or lose money if they made no response at all. The task consisted of a total of 85 trials. A total of 10 numbers were used. Five numbers are "winners" (37, 96, 78, 53, 29) and five are "losers" (43, 82, 64, 73, 31). The numbers were presented on the computer screen for 2 s with an inter-trial interval of one second. The first 5 trials consisted of all winning numbers (to establish a dominant response set), and the remainder of the trials were randomly ordered with no consecutive win or loss sequence exceeding three trials. Participants had to learn, by trial and error, when to respond and when to not respond. Trials were presented in eight continuous blocks of 10, excluding the first 5. Impulsivity was operationalized as the total errors of commission (i.e., pressing the spacebar when incorrect) for the last 40 trials of the task. Such errors reflect an inability to inhibit incorrect responding under circumstances involving sustained attention (Newman & Kosson, 1986).

Procedure

Upon arriving at the laboratory, the procedure of the study was explained to the participants and informed consent was obtained. In order to disguise the constructs being assessed participants were given a fictitious cover story. Participants were informed that the purpose of the investigation is to assess the impact of personality on ability and accuracy of performance on a variety of computerized tasks. Next, demographic information was collected. Participants then completed the self-report questionnaires. All self-report data was collected using computerized versions of each measure. Following the completion of the self-report measures, participants completed the computerized behavioral tasks. Since each behavioral task is designed to assess a different construct, it is unlikely that the order participants complete the tasks would impact responses. However, in order to control for any potential order effects, the sequence that the tasks were completed in was counterbalanced. Participants were randomly assigned to one of six counterbalanced sequences. After completing the behavioral tasks, participants were debriefed and compensated with course credit for their participation.

Chapter Three: Results

Prior to analysis, the data were screened to ensure that they did not violate the assumptions of SEM. There were no cases of missing data. However, eight cases of multivariate outliers were identified and deleted (Mahalanobis Distance values; D^2, p<.001), resulting in a sample size of 151. Data was also screened to ensure the deception manipulation was effective. Participants were asked about their subjective perceptions of their opponent, whether their opponent tried hard to win, whether they thought the task was a good measure of reaction-time, and how well they believed they performed on the task, etc. Typical responses included that the opponent was competitive, did well, played fairly, tried hard, did better or about the same as them, and that the task was a good test of reaction-time. Their descriptions indicated that the deception manipulation was successful.

Correlation Data

The correlations for relevant self-report subscales and behavioral tasks are provided in Table 3.1. In terms of Campbell and Fiske's (1959) MTMM coefficients, these correlations provide mixed evidence for the validity of TAP shock intensity selections as a measure of aggression. Examination of the correlations between self-report measures of aggression and TAP responses (Monotrait-Heteromethod coefficients), suggests limited evidence of convergent validity. Although the correlation between self-report direct aggression and TAP shock selection is significantly different from zero and positive (r=.217, p<.007) it is small in magnitude.

Correlations between TAP responses and other behavioral measures (Heterotrait-Monomethod coefficients) indicate some evidence of discriminant validity. Specifically, the correlation between the behavioral measure of pro-social behavior and TAP measured aggression is significantly different from zero and negative, as hypothesized (r=-.342, p<.001). Additionally, the majority of the correlations between TAP responses and self-report measures of impulsivity and pro-social behavior (Heterotrait-Heteromethod coefficients) are not significantly different from zero, also providing evidence of the discriminant validity of the TAP. Despite the small magnitude of these correlations, the average Monotrait-Heteromethod coefficient (convergent validity indicator; r=.165) is larger than the average Heterotrait-Heteromethod coefficient (discriminant validity indicator; r=-.092). The difference between the average of these two coefficients is also significant (*Z*=2.22; *p*=0.01), in accordance with Campbell and Fiske's (1959) criteria for establishing discriminant validity.

Structural Equation Modeling Data

The construct validity of the TAP was further examined utilizing a Correlated Trait – Correlated Method (CT-CM) CFA model. The hypothesized CT-CM model consists of three correlated trait factors (aggression, impulsivity, and pro-social behavior) and two correlated method factors (self-report and behavioral measures). The CT-CM model returned an admissible solution. Multiple fit indices were considered when evaluating model fit. These fit indices include: a model chi-square test, the Goodness of Fit Index (GFI), the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Residual (SRMR). Model chi-square tests are problematic as they are sensitive to sample size. Therefore, significant chi-square tests are not necessarily indicative of poor model fit. However, model chi-square tests are included due to its usefulness in comparing hierarchical models within the CT-CM model framework. Based on Kline's (2005) recommendations the following values are indicative of good model fit: GFI > .90, CFI > .90, RMSEA < .05 [upper bound of 90% CI < .10], SRMR <.10.

Model fit indices for the CT-CM model indicates a good fit ($\chi^2(60)=78.946$, $p=.051$; GFI=.930; CFI=.937; RMSEA=.046[.000-.072]; SRMR=.0635). This model was then compared to alternative models to test specific hypotheses. The CT-CM model was first evaluated in terms of improvement in fit over the most parsimonious models, the Null and Single Factor models. The Null model consists of the indicator variables only with no modeled trait or method factors (see Figure 3.2). The Single Factor model is comprised of one single factor that accounts for the relation between all indicators (see Figure 3.3). Fit indices for the Null and Single Factor model indicate poor fit (see Table 3.2 for model fit indices). Analysis of the difference in χ^2 values demonstrates that the CT-CM model has significant improvement in fit over the Null model (χ^2 dif(31)=311.813, $p<.001$) and the Single Factor model (χ^2 dif(17)=163.833, $p<.001$).

In order to test the convergent validity of the hypothesized CT-CM model, an alternative model was fit with the trait factors removed. This model consisted of the two correlated method factors only (see Figure 3.4). Model fit indices of the method factor only model denote poor fit (see Table 3.2). The inclusion of the three trait factors in the CT-CM model provides significant improvement in fit compared to the correlated methods model (χ^2 dif(17)=153.548, *p*<.001). The significant improvement in model fit provides evidence of convergent validity for the CT-CM model.

The discriminant validity of the CT-CM model was also examined through comparison with an alternative model that contained a single trait factor. In this alternative model, the two correlated method factors remain but the three trait factors were replaced with a single factor (see Figure 3.5). Model fit indices indicate a poor model fit (see Table 3.2). The specification of three distinct trait factors provides significant improvement in fit over a single unidimensional trait (χ^2 dif(4)=114.142, *p*<.001). This significant difference in model fit provides evidence for the discriminant validity of the CT-CM model.

The CT-CM model was also compared to an alternative model with method factors removed to test for the presence of method effects. The alternative model consisted of the three correlated trait factors only (see Figure 3.6). Model fit indices suggest poor model fit (see Table 3.2). The addition of the two method factors resulted in a significant improvement in model fit (χ^2 dif(15)=128.88, *p*<.001). This significant difference in model fit suggests the presence of method effects in this model. That is, a significant proportion of variance in this model is due to method effects. Constraining the correlation between method factors to zero does not result in a significant change in model fit (χ^2 dif(1)=0.087, *p*=.768). Thus, the method factors are not correlated as the correlation between them is not significantly different from zero.

To further establish the convergent validity of TAP shock selections as a measure of aggression, the aggression trait factor loadings were examined. In general, high factor loadings suggest evidence of convergent validity. The unstandardized factor loading for the Aggression trait factor on TAP shock selections is 6.897, *p*=.029. That is, for every one unit increase in the Aggression factor, TAP shock selections increase by 6.897. This significant factor loading suggests additional evidence for the convergent validity of the TAP as a measure of aggression. However, it is also important to examine the standardized factor loadings when comparing multiple factor loadings. When evaluating measurement CFA models (e.g. models where indicators are specified to measure multiple factors) standardized factor loadings are interpreted as standardized regression coefficients (Kline, 2005). The standardized factor loading for the Aggression trait factor on TAP is .20, while the aggression trait factor loadings for the self-report aggression measures are markedly higher in magnitude (.54-.74). Together, this pattern of factor loadings does suggest limited evidence for the convergent validity of the TAP.

The correlations between the aggression and other trait factors were also examined as an additional indicator of discriminant validity (see Figure 3.1). In order to test whether the correlations between the trait factors were significantly different from zero, the CT-CM model was compared to the nested models with the correlations between one set of trait factors removed (i.e. constrained to zero). Chi squared difference tests demonstrated that the correlation between the aggression and impulsivity trait factors (r=-.16) was not significantly different from zero (χ^2 dif(1)=1.972, p=.1602). However, this correlation was hypothesized to be positive and significantly different from zero. In addition, the correlation between aggression and pro-social behavior trait factors (r=-.53) was significantly different (χ^2 dif(1)=16.6, p<.001). These results support the discriminant validity of the aggression trait factor as there is no significant correlation between aggression and impulsivity but there is a significant negative correlation with pro-social behavior as hypothesized.

The behavioral task method factor loadings on TAP shock selections were also evaluated to examine the potential for common method effects. High method factor loadings are indicative of shared method effects. While this factor loading is significant (β=-.33, p<.001), it is a negative factor loading. This negative factor loading indicates that for every one standard deviation increase in the behavioral task method factor, TAP shock responses decrease by .33 standard deviations. In general this can be interpreted as the variance in responses on the TAP not due to aggression being negatively related to method variance in other behavioral measures. Specifically, the method variance present in the Prisoner's Dilemma Task is negative correlated with the method variance in the TAP given the lack of factor loading for the impulsivity behavioral task (β=.00). This suggests that there is significant variance in TAP shock selections not explained by the trait aggression factor. However, the shared method variance of the computerized behavioral tasks utilized in this experiment is not a homogenous factor.

	Aggression				Impulsivity						Prosocial Behavior			
	BPAQ-P	*BPAQ-V*	*RCRQ-D*	*TAP*	*BIS-A*	*BIS-M*	*BIS-N*	*UPPS-LP*	*UPPS-SS*	*GNG*	*PTM-A*	*PTM-AL*	SRAS	PD
BPAQ- Physical	-													
BPAQ- Verbal	.528	-												
RCRQ-Direct	.577	.428	-											
TAP-Total Shock	.112	-.047	.217	-										
BIS-Attentional Impulsivity	-.116	-.012	-.102	-.152	-									
BIS-Motor Impulsivity	-.068	-.110	.044	.025	-.065	-								
BIS-Nonplanning Impulsivity	-.054	-.033	-.080	-.156	.510	.049	-							
UPPSP-Lack of Planning	.208	.291	.107	.006	.191	.197	.175	-						
UPPSP-Sensation Seeking	-.050	-.098	-.164	-.167	.021	.174	.176	.203	-					
Go/NoGo- Commission Errors	.102	-.047	.000	-.010	.012	.096	.012	.020	.038	-				
PTM-Anonymous	-.013	-.063	-.244	-.101	.072	.129	.167	.462	.031	-.063	-			
PTM-Altruism	-.199	-.228	-.296	-.150	.023	-.113	.085	-.338	.362	-.082	.106	-		
SRAS	.200	.133	.148	-.042	.028	.069	.080	.254	-.015	.037	.173	-.037	-	
PD-Cooperation	-.045	.066	-.043	-.342	.054	-.023	.049	.061	.100	-.007	.125	.061	.125	-

Table 3.1. Correlations Between Observed Variables

Table 3.2. Model Fit Indices for Hypothesized and Alternative Structural Equation Models

Model	χ^2	df	*p*	GFI	CFI	RMSEA	CI	SRMR
Correlated Traits + Correlated Methods	78.946	60	.051	.930	.937	.046	.000-.072	.0635
Null Model	390.795	91	<.001	.716	.000	.148	.133-.163	.1626
Single Factor Model	246.779	77	<.001	.808	.434	.121	.104-.138	.1153
Correlated Method Factors Only	232.494	77	<.001	.820	.481	.116	.099-.133	.1136
Single Trait + Correlated Methods	193.088	64	<.001	.847	.569	.116	.097-.135	.1006
Correlated Trait Factors Only	207.826	75	<.001	.838	.557	.109	.091-.126	.1139

Figure 3.1. Hypothesized correlated trait – correlated method model that includes observed variables, three correlated latent trait factors, and two correlated latent method factors with standardized factor loadings and correlations.

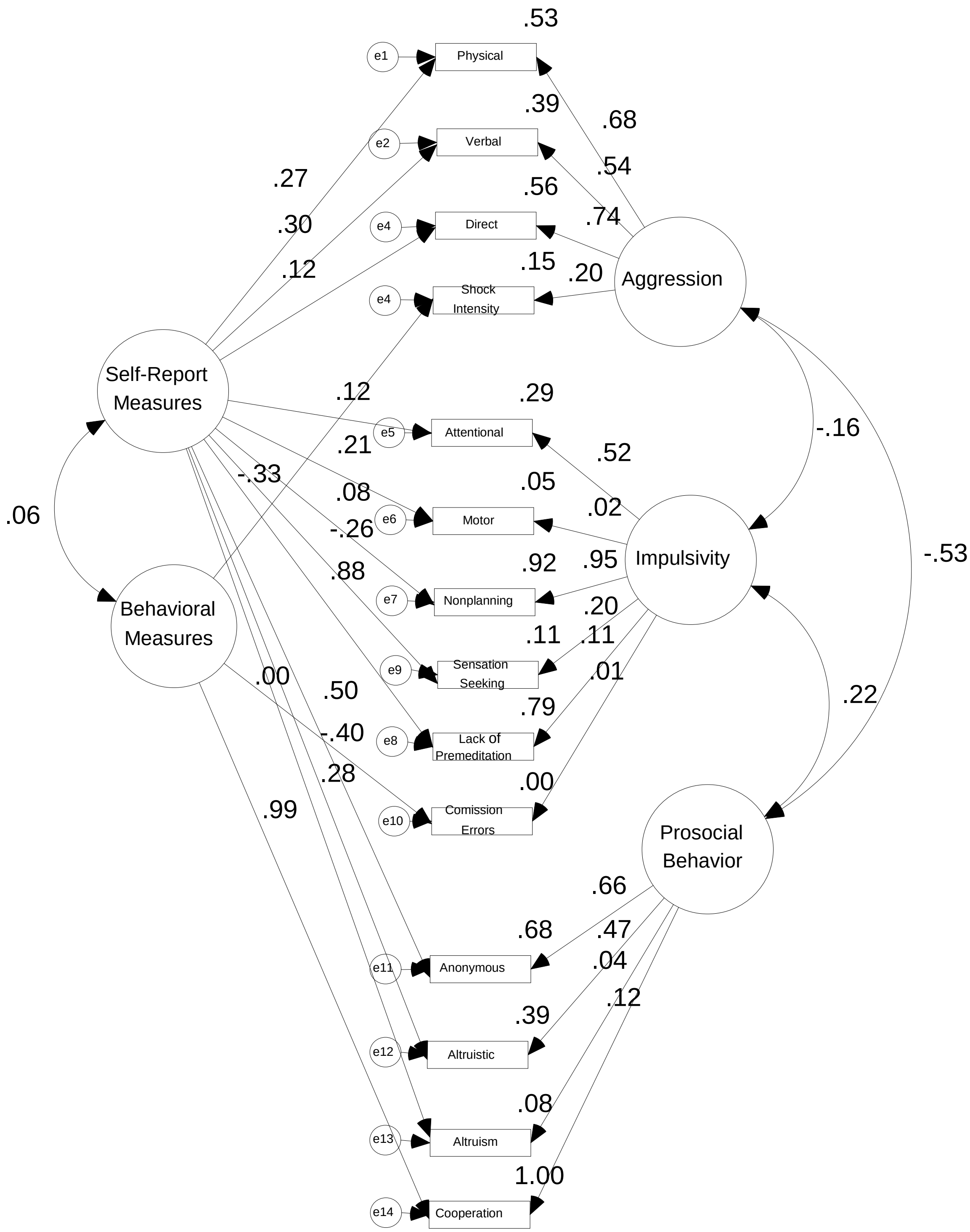


Figure 3.2. Null hypothesis model that includes observed variables only and no latent variables with standardized factor loadings.

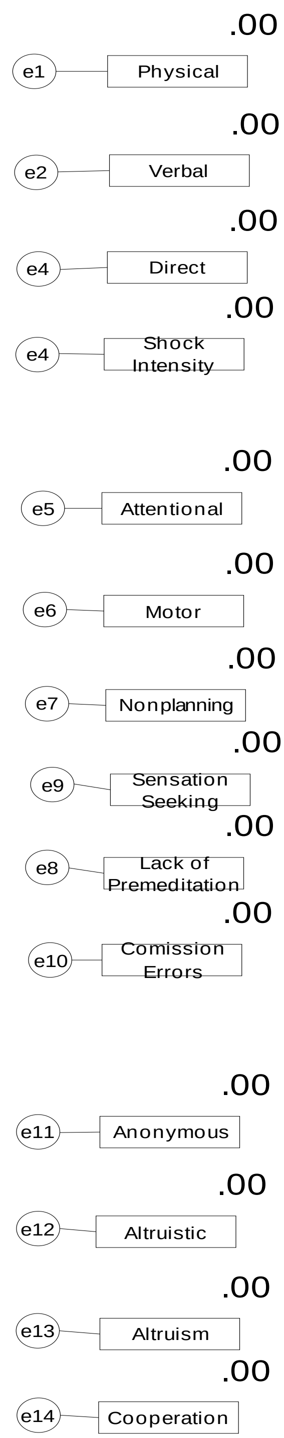


Figure 3.3. Single factor model that includes observed variables and one latent variable with standardized factor loadings.

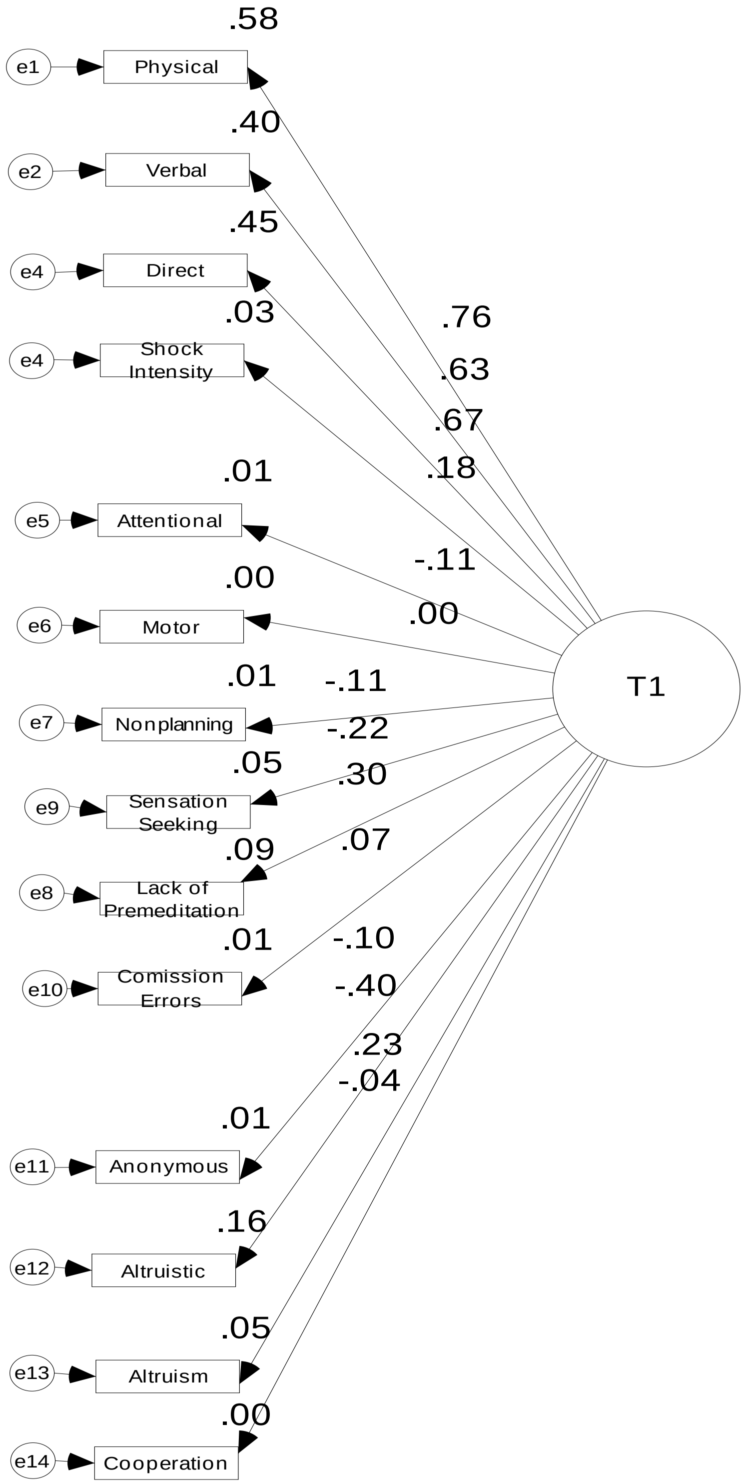


Figure 3.4. Correlated method model that includes observed variables and two correlated latent method factors with standardized factor loadings and correlations.

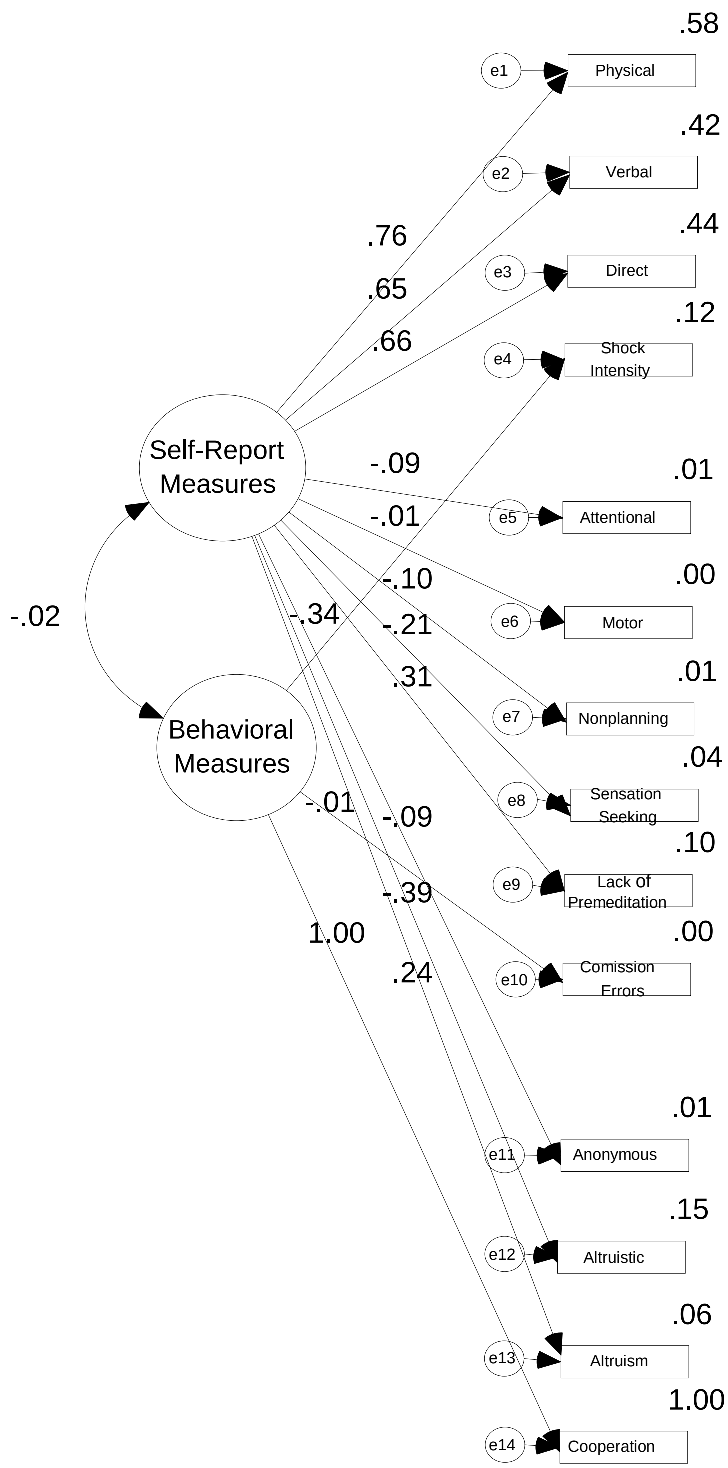


Figure 3.5. Single trait and correlated method model that includes observed variables, one latent trait factor, and two correlated latent method factors with standardized factor loadings and correlations.

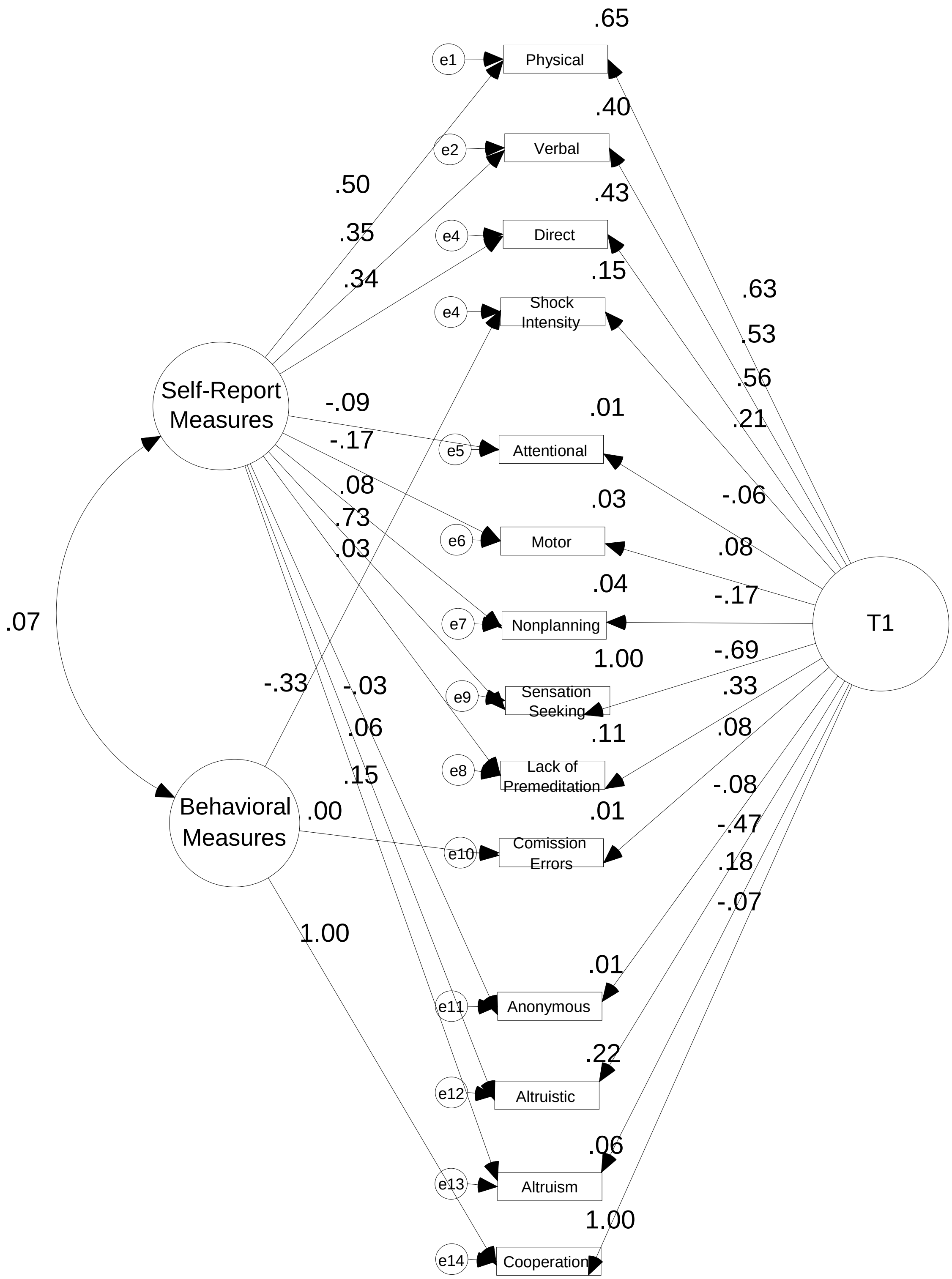
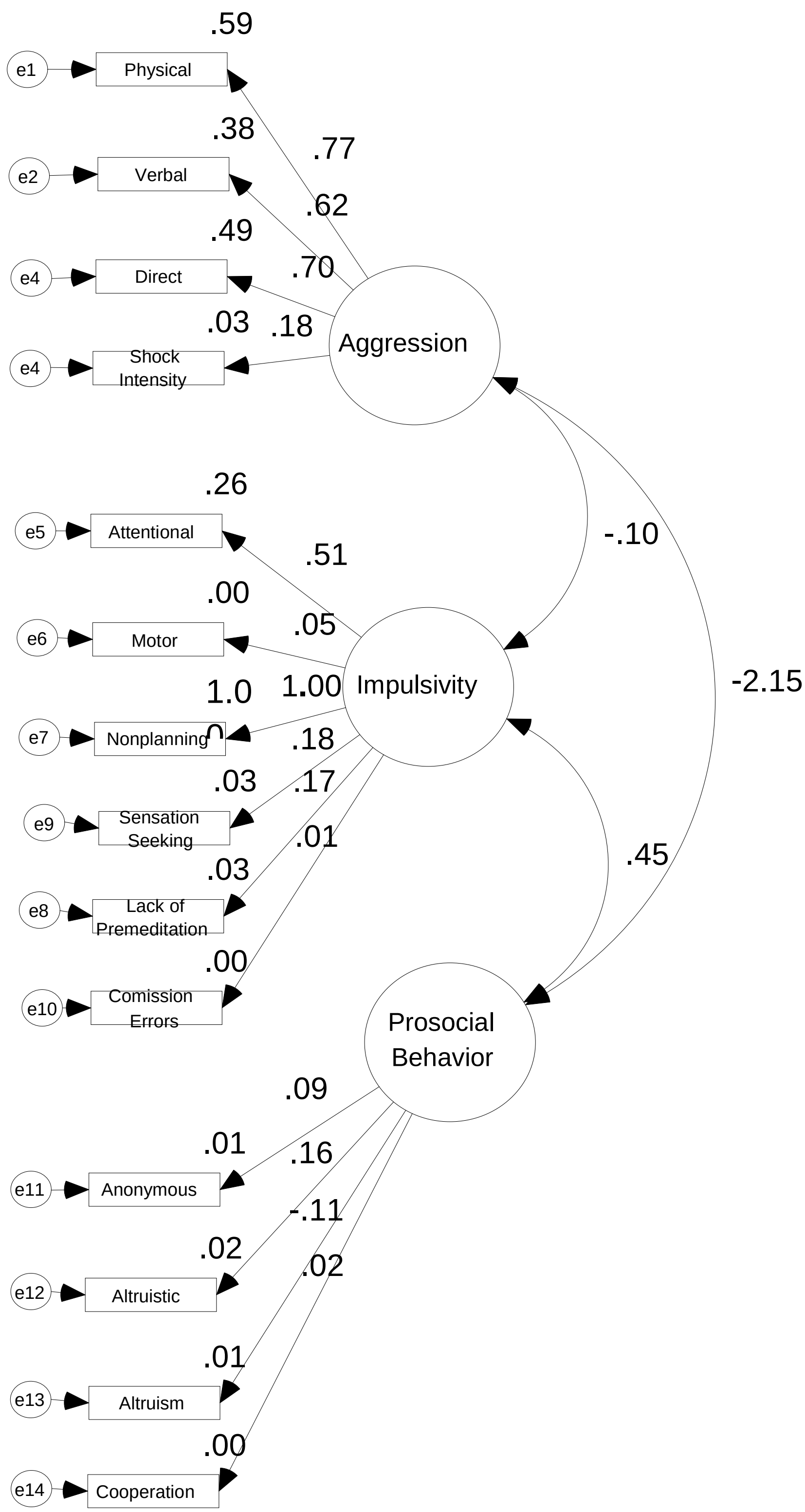


Figure 3.6. Correlated trait model that includes observed variables and three correlated latent trait factors with standardized factor loadings and correlations.



Chapter Four: Discussion

The results from this study provided several methods by which the construct validity of the TAP may be examined. Campbell and Fiske's (1959) criteria for evaluating correlation matrices are perhaps the most basic method for evaluating the convergent and discriminant validity of the TAP. In MTMM correlation matrices, the criteria for convergent validity is that monotrait-heteromethod correlations demonstrate a statistically significant different from zero and that the correlation is "large enough to encourage further examination of validity" (Campbell & Fiske, 1959, p. 82). In the above data, not all the monotrait-heteromethod correlations for the TAP are statistically significant. The correlation between the TAP and the RCRQ direct aggression subscale is significant. However, the correlations between the TAP and the two BPAQ subscales (verbal and physical aggression) are not significantly different from zero. This could be interpreted as limited evidence supporting the hypothesis that there would be significant positive correlations between the TAP and self-report measures of aggression as an indicator of convergent validity. The non-significant correlations could be interpreted as evidence that the RCRQ direct aggression subscale is a more accurate measure of aggression as measured by the TAP, compared to the BPAQ verbal and physical aggression subscales. It is possible that the type of aggression measured by the two BPAQ subscales are not as closely related to the type of aggression measured by the TAP as the RCRQ direct aggression subscale. In fact, the BPAQ verbal aggression subscale is slightly negatively correlated with TAP responses. As noted above, there are numerous hypothesized taxonomic systems that classify aggression using these dimensions (i.e. direct - indirect or verbal - physical). It is possible that the correlation observed between the TAP and the RCRQ direct aggression subscale is evidence of convergent validity for the TAP as a measure of direct aggression. It is unclear if the magnitude of this correlation would be sufficiently by Campbell and Fiske's standards. However, it is clear that the lack of statistically significant and sufficiently large correlations does not bode well for establishing the convergent validity of the TAP.

Campbell and Fiske's (1959) criteria for evaluating discriminant validity using a MTMM correlation matrix can be generally described as establishing evidence that the monotrait-heteromethod correlations are larger in magnitude than the heterotrait-heteromethod and heterotrait-monomethod correlations. There is evidence that the average correlations between the TAP and self-report measures of impulsivity and pro-social behavior (heterotrait-heteromethod) are significantly smaller in magnitude than the average correlation between the TAP and self-report measures of aggression. As hypothezised, there is also evidence that the correlation between the TAP and the pro-social behavioral task (heterotrait-monomethod) was statistically significant and in the predicted direction. However, the other heterotrait-monomethod correlation, between the TAP and the behavioral impulsivity task, was not significantly different from zero nor in the predicted direction. Nevertheless, it is smaller in magnitude than the average correlation between the TAP and self-report measures of aggression (monotrait-heteromethod correlations).

At best, evaluation of the MTMM correlation matrix provides only weak support for the convergent validity of the TAP. However, there are considerable issues with utilizing Campbell and Fiske's (1959) criteria for evaluating MTMM matrices. One significant issue is that their criteria treat convergent and discriminant validity as dichotomous variables. With the variability in correlations observed in the present study, it is difficult to interpret the construct validity of the TAP in terms of a present or absent dichotomy. In contrast, it has been noted that utilizing a SEM approach in analyzing MTMM data provides more information on construct validity than the original correlation matrix criteria (see Reichardt & Coleman, 1995). SEM techniques have the advantage of conceptualizing the degree to which a measure has convergent and discriminant validity. In previous research on the TAP, researchers have provided numerous different pieces of supporting evidence of its validity. Each study that provides evidence supporting the validity of the TAP contributes to an overarching body of knowledge on the degree to which responses on the TAP are related to a variety of other constructs. The advantage of the SEM analytic procedures used in this study is that it provides a means to evaluate the degree of convergent and discriminant validity evidence present. Therefore, SEM approaches are preferable to the correlation matrix method of analyzing MTMM data, particularly when there is adequate model fit (Reichardt & Coleman, 1995).

As noted above, the data was fit to a Correlated Trait-Correlated Method (CT-CM) CFA model. This model consisted of observed variables from the self-report measures chosen based on consistently with the construct and each behavioral task. Each behavioral task yielded one observed variable. In the TAP this was the average of all shock intensity selections from the task. For example, the RCRQ direct, BPAQ verbal, and BPAQ physical aggression subscales were chosen as the self-report subscales that closest matched the definition that aggression is *"any* *form of behavior directed toward the goal of harming or injuring another living being who is motivated to avoid such treatment"* (Baron & Richardson, 1994; p.7). Similarly the BPAQ hostility and BPAQ anger subscales were excluded as they are conceptualized as the cognitive and affective components of aggression that do not assess a specific aggressive behavior. The RCRQ indirect aggression subscale was excluded due to its theoretical differences from the type aggression measured by the TAP and the other included self-report subscales. Specifically, it is conceptualized as a measure of aggression that occurs through a circuitous route rather than an action directed at a specific target. Latent variables were modeled for aggression, impulsivity, and pro-social traits as well as behavioral task and self-report methods.

Analysis of the CT-CM model resulted in good model fit indices, indicating that the proposed model fit the observed data well. Subsequent comparisons to alternative CFA models provided additional evidence that the CT-CM model was the best model to explain the observed data. The model fit indices (provided in Table 3.2) clearly show that the CT-CM model has better fit values than any of the alternative models, supporting the hypothesis that this model would provide evidence of convergent validity. In addition, since the alternative models tested are nested models, chi-square difference tests provide a means for establishing whether the difference in model fit is statistically significant. The first model comparison between the Null and the CT-CM model simply provides evidence that there is shared variance between some of the observed variables in the model. While this finding is not surprising, it does establish that specifying no relation between the hypothesized underlying constructs assessed by these measures is a poor fit of the observed data. Further, comparison of the single trait model and the CT-CM model indicates that the shared variance in the observed data is not adequately explained by a single primary factor.

Systematic evaluations of nested alterative models provide solid evidence for the convergent and discriminant validity of the CT-CM model. First, comparison of the hypothesized CT-CM model to the nested model with trait factors removed and correlated method factors (0T-CM) demonstrates that the method factors alone are not sufficient to adequately fit the observed data. The inclusion of the hypothesized trait factors significantly improves model fit. The hypothesized latent trait variables explain a significant amount of variance in the observed variables and provide supporting evidence for the convergent validity of the CT-CM model.

Next, comparison of the CT-CM model to the nested model with a single trait factor and correlated method factors (1T-CM) provides evidence that the three hypothesized trait factors fit the data significantly better than a single trait factor. The aggression, impulsivity, and pro-social trait indicators share some variance after removing the variance shared due to measurement method. This comparison provides evidence that the CT-CM model has discriminant validity in that the three separate trait factors significantly better explain the shared variance than a single trait factor when shared method variance is removed. Finally, the comparison of the CT-CM model to a nested model with the hypothesized three correlated trait factors and method factors removed (CT-0M) show significant evidence of the presence of method effects in the CT-CM model, supporting the hypothesis that there exist shared method effects for both self-report and behavioral measures. The removal of the two method factors results in a significant decrease in model fit. Therefore, the hypothesized method effects are present in this data as evidenced by a significant increase in the variance accounted for in the observed data with the inclusion of method factors. As predicted additional analysis revealed that the shared method effects for self- report and behavioral measures were not correlated.

These findings have several interesting implications. First, SEM analyses provide evidence that the data are best explained by the hypothesized CT-CM model. This model is consistent with the theorized relations between self-report and behavioral measures as well as the constructs of aggression, impulsivity, and pro-social traits. The latent variables loaded onto the observed variables as predicted resulting in the three trait and two method factors. Second, the model has strong evidence of convergent and discriminant validity based on the magnitude of change and statistical significance in the difference between model fit data. The model shows convergent validity in that the self-report and behavioral measures are assessing the same construct for aggression, impulsivity, and pro-social traits respectively. The model has evidence of discriminant validity in that the observed data is assessing three distinct traits. This is particularly important as previously, there have been very few demonstrations of the discriminant validity of the TAP. Third, there are method effects present in the observed data. It is also noteworthy that the two method effects are uncorrelated. Not only does the method of measurement account for a significant amount of variance found in the observed data, but as hypothesized, the method of measurement also results in variance that is unique to the mode of measurement.

As hypothesized, examination of the factor loadings for the observed TAP responses provides additional evidence of construct validity. As noted above, the factor loadings indicate that for every one unit increase in the aggression trait factor, TAP shock selections increase by 6.897. The aggression trait factor does account for a significant portion of the variance observed in TAP shock selections. In addition, the aggression trait factor accounted for significant portions of the variance observed in the self-report measures of aggression. These factor loadings suggest that the TAP does have good convergent validity. However, it should also be noted that the standardized factor loadings for the self-report aggression measures are markedly larger than that of the TAP. Further, there is a significant amount of variance in the observed TAP shock selections that is attributed to unique error variance in this model. Nevertheless, the aggression trait factor loadings provide further evidence for the convergent validity of the TAP.

Despite the good model fit for the hypothesized model, there are a few unexpected results that should be addressed. First, a negative correlation between aggression and impulsivity trait factors was observed (see Table 3.1). It was predicted that aggression and impulsivity would be positively related. Instead, data indicate that correlations between these traits exhibited a degree of inconsistency across measures. Very few of the correlations were significant or even in the predicted direction. In examining this relation we also see that within the hypothesized SEM model, the correlation between the aggression and impulsivity trait factors was negative. One possible explanation is that the predicted relation was not found due to the characteristics of the sample. In previous research on this particular relation, clinical populations characterized by these traits (e.g. borderline personality disorder, intermittent explosive disorder, and paranoid personality disorder) have demonstrated large positive associations between aggression and impulsivity (McCloskey, Lee, Berman, Noblett, & Coccaro., 2008; McCloskey et al., 2009; Moeller et al., 1997). However, the relation appears to be inconsistent in populations with other clinical features or normal controls (McCloskey et al., 2009). In addition, the correlations between behavioral and self-report impulsivity measures in this study are non-significant, in keeping with previous research using non-clinical samples (Reynolds et al., 2006; Richards et al., 1999). It is possible that the relation is different or difficult to detect in populations with low base rates of impulsive or aggressive behavior. Given the lack of consistent correlations between measures of impulsivity in this sample, inclusion of a clinically relevant sample would likely result in the predicted correlations between the aggression and impulsivity trait factors. It is also possible that the model would have improved fit indices if significant correlations between self-report and behavioral measures of impulsivity would result in higher impulsivity trait factor loadings, and account for more variance in the model. However, it should still be noted that model fit was good with this sample and indicative of good construct validity of the model.

A second unexpected finding was the presence of some negative factor loadings within the CT-CM model. These negative factor loadings were present for both the self-report and behavioral method factors. The self-report method factor loaded negatively onto the UPPS+P sensation seeking subscale and the PTM altruism subscale. Overall, the magnitude of self-report method factor loadings are largely consistent across observed variables with the exception of these two subscales. It seems unlikely that the factor would be associated with lower scores on any self-report measure. However, when examining the individual items that make up these subscales, one difference between them and other self-report subscales stands out. Almost all items that comprise these subscales are reverse scored. In fact, all items are reverse scored on the sensation seeking subscale and five of seven items on the altruism subscale are reverse scored. The consistently negative language on these two scales may be a contributing factor. It could be argued that these subscales were originally written to assess a lack of sensation seeking and a lack of altruism scale. As such the shared method variance reflects their original wording in that decreased scores on sensation seeking or altruism (i.e. an increase in lack of sensation seeking or lack of altruism) is associated with an increase in the self-report method variance.

Much more problematic is the negative factor loading of the behavioral measure factor on the TAP. The factor loading indicates that increased scores on the TAP are associated with a decrease in the behavioral method factor. Alternatively, this could be interpreted as a positive association between the method factor and non-aggressive responding on the TAP. Comparing the behavioral method factor loadings, it is apparent that they are far less consistent than the self-report method factor loadings. One interpretation of this finding is that the behavioral tasks utilized in this study do not have a common source of shared variance. In fact, the factor loadings for the three behavioral measures display a large difference in amount of variance accounted for. There are also vast discrepancies in how these factor loadings are interpreted. Out of the three behavioral measures, the variance accounted for by the Prisoner’s dilemma task is perhaps the easiest to interpret as it was set as a reference indicator. As such, there is a large amount of variance accounted for by the behavioral task method factor. However, the Go/No Go task has a very small loading with almost no variance accounted for by the behavioral task method factor. It could be argued that these factor loadings provide evidence that the method effect for the behavioral tasks is inconsistent. It is also possible that differences in responding on these tasks result in much more variance than one might expect. All three tasks were computerized, involved a decision to press a button for a possible positive or negative outcome, and appear to be very similar modes of measurement. However, there are certainly discrepancies in the number of possible responses, length of task, cover story, and other aspects of these tasks. It is possible that the negative factor loading is due to the fact that the response options on the other tasks were coded zero and one where responses on the TAP ranged from one to ten. Therefore, lower responses on the TAP shared more variance with either response on the other two behavioral tasks. Regardless, it should be noted that the method factor loadings are an indication that there is less variance due to shared measurement effects found in the behavioral tasks than the self-report measures. Due to the inconsistencies in the loadings for the behavioral task method factor, it is difficult to speak to the impact of method variance on the ecological validity of the TAP. It is possible that there is some variance due to method that impacts the genrealizability of TAP aggression scores. However, it is also of note that the magnitude of method factor loadings for the TAP and self-report measures of aggression are comparable, suggesting the same interpretation for self-report measures of aggression.

A third unexpected finding is that the Go/No Go task has extremely low factor loadings. As noted above, the behavioral measurement factor did not load onto errors of commission committed during the Go/No Go task. In examining the commission error observed variable, it did not violate any of the assumptions of SEM. It displayed adequate normality in that it was not skewed or kurtotic and there were no extreme outliers. However, the impulsivity trait factor loading was also extremely low indicating that the impulsivity trait factor does not explain a significant amount of the variance in the observed errors of commission. One very definite possibility is the lack of significant correlations to self-report measures of impulsivity typically found in non-clinical samples mentioned earlier (Mitchell, 1999; Reynolds, Ortengren, Richards, & de Wit, 2006; Richards, Zhang, Mitchell, & de Wit, 1999). It is possible that due to the nature of the sample used in this study, the Go/ No Go task did not share enough variance with the self-report impulsivity measures to load onto the impulsivity trait factor. Therefore, one explanation may be the general finding that on this task specific populations display higher levels of impulsivity and demonstrate larger effect sizes and correlations with self-report impulsivity. Unfortunately, this means that the model failed to account for almost all of the variance observed in the Go/No Go measure of impulsivity. Within this model, the vast majority of the variance in the task is attributed to unique error variance. The alternative is that the variance is due to some other factor not included in this model. It seems unlikely that almost all of the variability would be due to error. Recent research utilizing this task indicates that while it is indeed a measure of impulsivity, that impulsivity may be attributed to several cognitive factors. These cognitive factors include attention to gains rather than losses, learning rate, and erratic or random responding (Yechiam et al., 2006). When an individual is attending to the positive reinforcement and not the negative present in this task, they increase the likelihood that they will commit an error of commission. Similarly, the learning rate of the individual may explain the variability in responding. For example, an individual with a very good learning rate is unlikely to make errors of commission throughout the majority of the task, where an individual with a very slow learning rate may commit errors of commission throughout the task. Finally, it is possible that some individuals did not attend to the positive and negative reinforcement and simply responded randomly throughout the task. *A*ny or all of these factors could account for the variability observed in this task.

It is important to note that there are several limitations inherent in this study. First, the sample utilized is a potential limitation due to the largely homogenous properties. All participants were male and most were approximately the same age and had similar social backgrounds. Further, this sample had no significant history of any clinically relevant features such as known psychopathology or history of criminally violent behavior. In addition, this sample was a limited size. A larger more diverse sample would allow for not only greater generalizability, but would likely provide clarification of some of the counterintuitive relations noted above. Based on the findings from this study and previous research, it appears possible that the relation between the variables measured in this study can vary significantly in different populations. The sample utilized here was successful at demonstrating the construct validity of the hypothesized model and the TAP. However, if relations between self-report and behavioral measures tend to be more pronounced in clinically relevant populations, then use of such a sample within this methodology would provide more robust evidence of construct validity. One possible option for future research in this area would be to recruit a larger sample more representative of the general population. This sample could include a larger range of ages, both men and women, and a range of clinical features. Utilization of such a sample would be able to provide perhaps a still more comprehensive way of examining the construct validity of the TAP. It may also demonstrate more accurate model fit and factor loading estimates. It would also be possible to compare group differences in model fit indices to those from the present study. Further separating a larger sample into groups based on sex, age, or other characteristics could allow for incremental validity estimates of the utility of the TAP. This type of experiment could also provide a direct test between a general community sample and the more traditional college convenience sample.

Another alternative avenue of future research would be to carry out the same experiment with a sample of clinical interest (e.g. violent offenders). As previous research has demonstrated that the TAP is capable of differentiating between these populations and controls (Verona et al., 2006; Wolfe & Baron, 1971), it would be valuable to examine the difference in construct validity of the TAP between these two groups. In addition, it appears that behavioral measures may indeed perform better in populations where the variable being assessed occurs at a higher rate and magnitude (McCloskey et al., 2009). This type of research would allow for direct comparisons of model estimates and factor loadings from the present study. Significant differences in model fit indices would provide valuable information about the specific populations in which the hypothesized model is most accurate. Despite the good model fit indices found in this study, it is still possible that these indices could be significantly better or worse in other samples. Provided that the researchers utilized the same measures and fit the data to the same model hypothesized in this study, it would be possible to examine differences due to the sample characteristics. Such a difference would provide compelling evidence that a clinically relevant feature such as violent history or psychopathology can moderate the relation between aggression, impulsivity, and pro-social behavior. Comparisons between different populations would allow for a more precise way to describe the construct of interest. This information would be essential in determining clinical usefulness as well as interpretations of scores on behavioral measures. This line of research could also directly address the issue of ecological validity and the ability to generalize between TAP responses and actual violent behaviors. Significant differences would not invalidate the findings of this study and previous reports of the construct validity of the TAP, but rather suggest that the TAP has predictive utility in identifying particular populations of interest. If such an experiment revealed no significant differences between samples, it would provide evidence additional evidence of construct validity and greater confidence that aggression as measured by the TAP is an indication of the actual level of aggressive behavior present in a given individual.

Summary

In general, this study represents a significant step for future research in the construct validity of the TAP. It provides a clear framework that can easily be tested using other populations and provides a comparison sample to be tested against. The presence of good model fit indices in this study is indicative of the construct validity of the entire model. Additionally, the data suggest differences in method variance between self-report and behavioral measures. The consistency of shared method effects in self-report data continues to be a significant issue in psychological research. Future research in this vein may at least be confident that any inflationary impact on relations between variables that is typically attributed to shared method variance may be accounted for within this methodology. In addition, researchers interested in the other self-report or behavioral measures used in this experiment may also build upon this design to provide a more direct examination of those measures or underlying constructs. In studies with significantly greater sample size, it would be possible to expand the number of constructs, observed variables, and the resultant nomological network. While this type of study would preclude direct tests of statistical differences from the current data, it would expand the understanding of relations between constructs of interests and the construct validity of the measures involved.

This study provides a significant advancement to the existing literature on the construct validity of the TAP. In addition to being one of the only existing a priori tests of construct validity, it is one of the most methodologically advanced means of examining this evidence. Despite the limited and homogenous sample, this study found supporting evidence for the convergent and discriminant validity of the TAP. As noted above, this is one of the few studies that have provided direct evidence of the discriminant validity of the TAP. In addition, it provides a framework and data that could be used for future comparisons and a starting point for research utilizing different populations. This study also provides evidence that in contrast to the shared method variance typically found in self-report measures, the TAP and other behavioral measures do not appear to display the same type or degree of shared variance. Finally, this study also provides evidence for the validity of the relation of the TAP and the construct of aggression to other theoretically related constructs.

References

Abel, E. (1977). The relationship between cannabis and violence: A review. *Psychological Bulletin, 84,* 193-211.

Anderson, C., Buckley, K, & Carnagey, N. (2008). Creating your own hostile environment: A laboratory examination of trait aggressiveness and the violence escalation cycle. *Personality and Social Psychology Bulletin, 34 (4*), 462-473.

Anderson, C., & Bushman, B. (1997). External validity of 'trivial' experiments: The case of laboratory aggression. *Review of General Psychology*, *1*(1), 19-41.

Anderson, C., & Bushman, B. (2002). Human aggression. *Annual Review of Psychology*, *53*(1), 27-51.

Archer, J. (1988). *The Behavioural Biology of Aggression.* Cambridge University Press, Cambridge.

Archer, J. (1991). The influence of testosterone on human aggression. *British Journal of Psychology*, *82*(1), 1-28.

Archer, J. (2000). Sex differences in aggression between heterosexual partners: A meta-analytic review*. Psychological Bulletin, 126(5),* 651-680.

Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology, 8*, 291-322.

Aromäki, A., Lindman, R., & Eriksson, C. (2002). Testosterone, sexuality and antisocial personality in rapists and child molesters: A pilot study. *Psychiatry Research*, *110*(3), 239-247.

Ayduk, O., Gyurak, A., & Luerssen A. (2008). Individual differences in the rejection–aggression link in the hot sauce paradigm: The case of rejection sensitivity, *Journal of Experimental Social Psychology, 44,*775-782.

Baron, R., & Eggleston, R. (1972). Performance on the 'aggression machine': Motivation to help or harm?. *Psychonomic Science*, *26*(6), 321-322.

Baron, R., & Richardson, D. (1994). *Human aggression (2nd ed.)*. New York, NY, US: Plenum Press.

Barratt, E. & Slaughter, L. (1998). Defining, measuring, and predicting impulsive aggression: A heuristic model. *Behavioral Science and Law*, 16, 285-302.

Barratt, E., Stanford, M., Kent, T., & Felthous, A. (1997). Neuropsychological and cognitive psychophysiological substrates of impulsive aggression. *Biological Psychiatry*, 41, 1045-1061.

Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Becker, T., & Cote, J. (1994). Additive and multiplicative method effects in applied psychological research: An empirical assessment of three models. *Journal of Management*, *20*(3), 625-641.

Ben-Porath, D. & Taylor, S. (2002). The effects of diazepam (Valium) and aggressive disposition on human aggression: An experimental investigation. *Addictive Behaviors, 27,* 167 – 177.

Berkowitz, L. (1969). The frustration-aggression hypothesis revisited. In L. Berkowitz (Ed.), *Roots of aggression: A re-examination of the frustration-aggression hypothesis* (pp. 1-28). New York: Atherton Press.

Berkowitz, L. (1993). Towards a general theory of anger and emotional aggression: Implications of the cognitive-neoassociationistic perspective for the analysis of anger and other emotions. *Perspectives on anger and emotion* (pp. 1-46).

Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology*, *7*(2), 202-207.

Berman, M., Gladue, B., & Taylor, S. (1993). The effects of hormones, Type A behavior pattern, and provocation on aggression in men. *Motivation and Emotion*, *17*(2), 125-138.

Berman, M. & Taylor, S. (1995). The effects of triazolam on aggression in men. *Experimental and Clinical Psychopharmacology, 3 (4),* 411-416.

Bernstein, I. & Gesn, P. (1997). On the dimensionality of the Buss/ Perry Aggression Questionnaire. *Behavioral Research and Therapy*, 35, 563–568.

Bernstein, S., Richardson, D., & Hammock, G. (1987). Convergent and discriminant validity of the Taylor and Buss measures of physical aggression. *Aggressive Behavior*, *13*(1), 15-24.

Bettencourt, B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin*, *119*(3), 422-447.

Björkqvist, K., Lagerspetz, K., & Kaukiainen, A. (1992). Do girls manipulate and boys fight? Developmental trends in regard to direct and indirect aggression. *Aggressive Behavior*, *18*(2), 117-127.

Boes, A., Tranel, D., Anderson, S., & Nopoulos, P. (2008). Right anterior cingulate: A neuroanatomical correlate of aggression and defiance in boys. *Behavioral Neuroscience*, *122*(3), 677-684.

Böhnke, R., Bertsch, K., Kruk, M., & Naumann, E. (2010). The relationship between basal and acute HPA axis activity and aggressive behavior in adults. *Journal of Neural Transmission, 117(5)*, 629-637.

Bowman, M. (1997). Brain impairment in impulsive violence. In: C. Webster & M. Jackson (Eds.) *Impulsivity: Theory, Assessment, and Treatment*, pp 116–141. New York: Guilford Press.

Brown, C., Kent, T., Bryant, S., & Gevedon, R. (1989). Blood platelet uptake of serotonin in episodic aggression. *Psychiatry Research*, *27*(1), 5-12.

Brown, G., Goodwin, F., Ballenger, J., Goyer, P., & Major, L. (1979). Aggression in humans correlates with cerebrospinal fluid amine metabolites*. Psychiatry Research, 1*(2), 131-139.

Bryant, F. & Smith, B. (2001). Refining the architecture of aggression: A measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138–167.

Burns, G. & Haynes, S. (2006). Clinical psychology: Construct validation with multiple sources of information and multiple settings. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 401-418). American Psychological Association, Washington, D.C.

Bushman, B. (1995). Moderating role of trait aggressiveness in the effects of violent media on aggression. Journal of Personality and Social Psychology, 69(5), 950-960.

Bushman, J. & Anderson, C. (2001). Is it time to pull the plug on hostile versus instrumental aggression dichotomy? *Psychological Review*, 108, 273-279.

Bushman, B., Baumeister, R., & Phillips, C. (2001). Do people aggress to improve their mood? Catharsis beliefs, affect regulation opportunity, and aggressive responding. *Journal of Personality and Social Psychology*, *81*(1), 17-32.

Bushman, B., & Cooper, H. (1990). Effects of alcohol on human aggression: An integrative research review. *Psychological Bulletin*, *107*(3), 341-354.

Buss, A. (1961). Investigating Aggression in the Laboratory. *The psychology of aggression* (pp. 35-52). Hoboken, NJ, US: John Wiley & Sons Inc.

Buss A. (1971). Aggression plays. In: Singer J.L. (Ed.). The *Control of Aggression and Violence*. Academic Press, New York.

Buss, A. & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63,* 452-459.

Buydens-Branchey, L., Branchey, M., Fergeson, P., Hudson, J., & McKernin, C. (1997). The meta-chlorophenylpiperazine challenge test in cocaine addicts: Hormonal and psychological responses*. Biological Psychiatry, 41*(11), 1071 – 1086.

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

Campbell, A., & Muncer, S. (2008). Intent to harm or injure? Gender and the expression of anger. *Aggressive Behavior*, *34*(3), 282-293.

Carlo, G. & Randall, B. (2002). The development of a measure of pro-social behaviors for late adolescents. *Journal of Youth and Adolescence*, *31*, 31-44.

Cherek, D. (1981). Effects of smoking different doses of nicotine on human aggressive behavior. *Psychopharmacology*, *75*(4), 339-345.

Chermack, S., & Giancola, P. (1997). The relation between alcohol and aggression: An integrated biopsychosocial conceptualization. *Clinical Psychology Review*, *17*(6), 621-649.

Chermack, S. & Taylor, S. (1993). Barbiturates and human physical aggression. *Journal of Research in Personality, 27,* 315–327.

Coccaro, E. (1989). Central serotonin and impulsive aggression. *British Journal of Psychiatry*, *155*(8), 52-62.

Coccaro, E. (1998). Impulsive aggression: A behavior in search of clinical definition. *Harvard Review of Psychiatry*, *5*(6), 336-339.

Coccaro, E., Kavoussi, R., Cooper, T., & Hauger, R. (1997). Central serotonin activity and aggression: Inverse relationship with prolactin response to d-fenfluramine, but not CSF 5-HIAA concentration, in human subjects. *American Journal of Psychiatry*, *154*(10), 1430-1435.

Coccaro, E., Kavoussi, R., & Hauger, R. (1995). Physiological responses to d-fenfluramine and ipsapirone challenge correlate with indices of aggression in males with personality disorder. *International Clinical Psychopharmacology*, *10*(3), 177-179.

Coccaro, E., & Siever, L. (2005). Neurobiology. *The American Psychiatric Publishing textbook of personality disorders* (pp. 155-169). Washington, DC, US: American Psychiatric Publishing, Inc.

Crick, N., & Dodge, K. (1994). A review and reformulation of social information- processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74-101.

Crick, N. & Grotpeter, J.(2005). Relational aggression, gender, and social-psychological adjustment*.* *Child Development*, 66, 710-722.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302.

Cureton, E. (1950). Validity. In E. Lingquist (Ed.) *Educational Measurement* pp. 621-694*.* American Council on Education, Washington, D.C.

Cyders, M., & Smith, G. (2007). Mood-based rash action and its components: Positive and negative urgency. *Personality and Individual Differences, 43*, 839–850.

Deakin, J. & Graeff, F. (1991). 5-HT mechanisms of defense. *Journal of Psychopharmacology*, 5, 305-315.

Dodge, K. & Coie, J. (1987). Social information processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology*, *53*, 1146-1158.

Dollard, J., Doob, L., Miller, N., Mowrer, O., & Sears, R. (1939). *Frustration and aggression*. New Haven: Yale University Freer.

Eagly, A. H, & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*, 309-330.

Ellis, L. (1986). Evidence of neuroandrogenic etiology of sex roles from a combined analysis of human, nonhuman primate, and nonprimate mammalian studies. *Personality and Individual Differences, 7,* 519-552.

Eron, L. & Huesmann, L. (1984)*.* The relation of pro-social behavior to the development of aggression and psychopathology. *Aggressive Behavior, 10*, 201–211.

Evers, C., Fischer, A., Rodriguez Mosquera, P., & Manstead, A. (2005). Anger and social appraisal: A 'spicy' sex difference?. *Emotion*, *5*(3), 258-266.

Fagan, J. (1990). Intoxication and aggression. Crime and Justice, ed. M. Tonry and J. Wilson, pp. 241–320. Volume 13 of *Crime and Justice: A Review of Research*. Chicago: University of Chicago Press.

Fals-Stewart, W. (2003). The occurrence of partner physical aggression on days of alcohol consumption: A longitudinal diary study. *Journal of Consulting and Clinical Psychologist,* 71(1), 41-52.

Feldmann, T. & Johnson, P. (1998). Work place violence: A new form of lethal aggression. In *Lethal Violence: A source book on fatal domestic, acquaintance, and stranger violence* (Ed. Hall, H.). CRC Press: Boca Raton, FL.

Felsten, G., & Hill, V. (1999). Aggression Questionnaire hostility scale predicts anger in response to mistreatment. *Behavior Research and Therapy, 37*, 87–97.

Ferguson, C. (2007). Evidence for publication bias in video game violence effects literature: A meta-analytic review. *Aggression and Violent Behavior*, *12*(4), 470-482.

Feshbach, S. (1964). The function of aggression and the regulation of aggressive drive. *Psychological Review, 71*, 257-272.

Feshbach, N. & Feshbach, S. (1969). The relationship between empathy and aggression in two age groups. *Developmental Psychology, 1*, 102-107.

Frechette, A. & Rimsza, M. (1992). Stun gun injury: A new presentation of the battered child syndrome. *Pediatrics, 89 (5),* 898-901.

Frodi, A., Macaulay, J., & Thome, P. (1977). Are women always less aggressive than men? A review of the experimental literature. *Psychological Bulletin*, *84*(4), 634-660.

Geen, R. (1989). *Human aggression.* Monterey, CA: Brooks-Cole.

Geen, R. (2001). *Human aggression*. Philadelphia: Open University Press.

Gerevich, J., Bácskai, E., & Czobor, P. (2007). The generalizability of the Buss-Perry Aggression Questionnaire. *International Journal of Methods in Psychiatric Research*, *16*(3), 124-136.

Giancola, P. (2004). Executive Functioning and Alcohol-Related Aggression. *Journal of Abnormal Psychology*, *113*(4), 541-555.

Giancola, P. & Chermack, S. (1998). Construct validity of laboratory aggression paradigms: A response to Tedeschi and Quigley (1996). *Aggression and Violent Behavior*, *3*(3), 237-253.

Giancola, P., Helton, E., Osborne, A., Terry, M., Fuss, A., & Westerfield, J. (2002). The effects of alcohol and provocation on aggressive behavior in men and women. *Journal of Studies on Alcohol*, *63*(1), 64-73.

Giancola, P., & Parrott, D. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior*, *34*(2), 214-229.

Giancola, P., & Zeichner, A. (1994). Neuropsychological performance on tests of frontal-lobe functioning and aggressive behavior in men. *Journal of Abnormal Psychology*, *103*(4), 832-835.

Giancola, P., & Zeichner, A. (1995a). Construct validity of a competitive reaction-time aggression paradigm. *Aggressive Behavior*, *21*(3), 199-204.

Giancola, P., & Zeichner, A. (1995b). An investigation of gender differences in alcohol-related aggression. *Journal of Studies on Alcohol*, *56*(5), 573-579.

Gottfredson, M., & Hirschi, T. (1993). A control theory interpretation of psychological research on aggression. *Aggression and violence: Social interactionist perspectives* (pp. 47-68). Washington, DC, US: American Psychological Association.

Grafman, J., Schwab, K., Warden, D., Pridgen, A., Brown, H. R., & Salazar, A. M.: (1996). Frontal lobe injuries, violence and aggression: A report of the Vietnam Head Injury Study. *Neurology 46,* 1231- 1238.

Graves, K. (2007). Not always sugar and spice: Expanding theoretical and functional explanations for why females aggress. *Aggression and Violent Behavior*, *12*, 131-140.

Hammock, G., & Richardson, D. (1992). Predictors of aggressive behavior. *Aggressive Behavior*, *18*(3), 219-229.

Harris, M. (1996). Aggression, gender, and ethnicity. *Aggression and Violent Behavior, 1,* 123-146.

Hyde, J. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology*, *20*, 722-736.

Harris, J. (1995). Confirmatory factor analysis of the Aggression Questionnaire. *Behaviour Research and Therapy, 33,* 991–993.

Harris, M. & Knight-Bohnhoff, K. (1996). Gender and aggression: II. Personal aggressiveness. *Sex Roles*, 35 (1-2), 27-42.

Hawkins, K. & Trobst, K. (2000). Frontal lobe dysfunction and aggression: Conceptual issues and research findings. *Aggression and Violent Behavior*, *5*(2), 147-157.

Hoaken, P., Assaad, J., & Pihl, R. (1998). Cognitive functioning and the inhibition of alcohol-induced aggression. *Journal of Studies on Alcohol*, *59*(5), 599-607.

Hoaken, P. & Pihl, R. (2000). The effects of alcohol intoxication on aggressive responses in men and women. *Alcohol and Alcoholism, 35*, 4714-477.

Hoaken, P., Campbell, T., Stewart, S., & Pihl, R. (2003). Effects of alcohol on cardiovascular reactivity and the mediation of aggressive behavior in adult men and women. *Alcohol and Alcoholism, 38*, 84-92.

Horesh, N., Gothelf, D., Ofek, H., Weizman, T., & Apter, A. (1999). Impulsivity as a correlate of suicidal behavior in adolescent psychiatric inpatients. *Crisis, 20*, 8–14.

Houston, R., Stanford, M., Villemarette-Pittman, N., Conklin, S., Helfritz, L. (2003). Neurobiological correlates and clinical implications of aggressive subtypes. *Journal of Forensic Neuropsychology, 3*, 67-87.

Kane, T., Joseph, J., & Tedeschi, J. (1976). Person perception and the Berkowitz paradigm for the study of aggression. *Journal of Personality and Social Psychology, 33*, 663-673.

Kelly, T. & Cherek, D. (1993). The effects of alcohol on free-operant aggressive behavior. *Journal of Studies on Alcohol, 11,* 40-52.

Kenny, D. & Kashy, D. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112*, 165-172.

Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Bring, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry, 50*, 677 – 684.

Kline, R. (2005). *Principles and Practice of Structural Equation Modeling (2nd Ed.).* New York: Guiliford Press

Lagerspetz, K., Björkqvist, K., & Peltonen, T. (1988). Is indirect aggression typical of females? Gender differences in aggressiveness in 11- to 12-year-old children. *Aggressive Behavior*, *14*(6), 403-414.

Lance, C., Noble, C., & Scullen, S. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. *Psychological Methods*, *7*(2), 228-244.

Lau, M., Pihl, R., & Peterson, J. (1995). Provocation, acute alcohol intoxication, cognitive performance, and aggression. *Journal of Abnormal Psychology*, *104*(1), 150-155.

Licata, A., Taylor, S., Berman, M., & Cranston, J. (1993). Effects of cocaine on human aggression. *Pharmacology, Biochemistry and Behavior, 45,* 549-552.

Lieberman, J., Solomon, S., Greenberg, J., & McGregor, H. (1999). A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior*, *25*(5), 331-348.

Linnoila, M., Virkkunen, M., & Higley, G. (1993). Impulse control disorders. *International Clinical Psychopharmacology, 8*(Suppl. 1), 53-56.

Loeber, R. & Schmaling, K. (1985). Empirical evidence for overt and covert antisocial conduct problems: A meta-analysis. *Journal of Abnormal and Child Psychology.* 13, 337-352.

Lotze, M., Veit, R., Anders, S., & Birbaumer, N. (2007). Evidence for a different role of the ventral and dorsal medial prefrontal cortex for social reactive aggression: An interactive fMRI study. *NeuroImage, 34*, 470 –478.

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences*. Stanford University Press.

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55*, 95-107.

Macdonald, S., Erickson, P., Wells, S. Hathaway, A., & Pakula, B. (2008). Predicting violence among cocaine, cannabis, and alcohol treatment clients. *Addictive Behaviors, 33(1*), 201-205.

Malamuth, N. (1983). Factors associated with rape as predictors of laboratory aggression against women. *Journal of Personality and Social Psychology*, *45*(2), 432-442.

Malamuth, N., & Ceniti, J. (1986). Repeated exposure to violent and nonviolent pornography: Likelihood of raping ratings and laboratory aggression against women. *Aggressive Behavior*, *12*(2), 129-137.

Marsh, H., & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. *Structural equation modeling: Concepts, issues, and applications* (pp. 177-198). Thousand Oaks, CA US: Sage Publications, Inc.

McCloskey, M., & Berman, M. (2003). Alcohol intoxication and self-aggressive behavior. *Journal of Abnormal Psychology*, *112*(2), 306-311.

McCloskey, M., Lee, R., Berman, M., Noblett, K., & Coccaro, E. (2008). The relationship between impulsive verbal aggression and intermittent explosive disorder. *Aggressive Behavior*, *34*(1), 51-60.

McCloskey, M., New, A., Siever, L., Goodman, M., Koenigsberg, H., Flory, J., et al. (2009). Evaluation of behavioral impulsivity and aggression tasks as endophenotypes for borderline personality disorder. *Journal of Psychiatric Research*, *43*(12), 1036-1048.

McGregor, H., Lieberman, J., Greenberg, J., Solomon, S., Arndt, J., Simon, L., & Pyszczynski, T. (1998). Terror management and aggression: Evidence that mortality salience motivates aggression against worldview-threatening others. *Journal of Personality and Social Psychology*, 74(3), 590-605.

Moeller, F., Dougherty, D., Rustin, T., Swann, A., Allen, T., Shah, N., et al. (1997). Antisocial personality disorder and aggression in recently abstinent cocaine dependent subjects. *Drug and Alcohol Dependence*, *44*(2-3), 175-182.

Moore, T., Stuart, G., Meehan, J., Rhatigan, D., Hellmuth, J., & Keen, S. (2008). Drug abuse and aggression between intimate partners: A meta-analytic review. *Clinical Psychology Review*, *28*(2), 247-274.

Morgan, A.B. & Lillienfeld, S.O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review, 20*, 113-136.

Moyer, K. (1976). *The Psychobiology of Aggression.* Harper & Row, New York.

Murdoch, D., Pihl, R., & Ross, D. (1990). Alcohol and crimes of violence: Present issues. *International Journal of the Addictions*, *25*(9), 1065-1081.

Murphy, C., Winters, J., O'Farrell, T., Fals-Stewart, W., & Murphy, M. (2005). Alcohol consumption and intimate partner violence by alcoholic men: Comparing violent and nonviolent conflicts. *Psychology of Addictive Behaviors, 19,* 35–42.

Myerscough, R., & Taylor, S. (1985). The effects of marijuana on human physical aggression. *Journal of Personality and Social Psychology, 49,* 1541-1546.

Newman, J., & Kosson, D. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*, *95*, 252–256.

Odgers, C., Moretti, M., Burnette, M., Chauhan, P., Waite, D., & Reppucci, N. D. (2007). A latent variable modeling approach to identifying subtypes of serious and violent female juvenile offenders. *Aggressive Behavior*, *33*, 339-352.

Orne, M. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776-783.

Parker, L. (1989). Adrenal androgen metabolism. In L. N. Parker, Adrenal androgens in clinical medicine (pp. 3–28). San Diego, CA: Academic Press.

Parrott, D., & Giancola, P. (2007). Addressing 'The criterion problem' in the assessment of aggressive behavior: Development of a new taxonomic system. *Aggression and Violent Behavior*, *12*(3), 280-299.

Patton, J., Stanford, M., & Barratt, E. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, *51*(6), 768-774.

Pernanen, K. (1991). *Alcohol in human violence*. New York: Guilford Press.

Pihl, R., Young, S., Harden, P., & Plotnick, S. (1995). Acute effect of altered tryptophan levels and alcohol on aggression in normal human males. *Psychopharmacology*, *119*(4), 353-360.

Phillips, J. & Giancola, P. (2009, March). Lifetime history of alcohol and non-alcohol related violence: Examining the validity of a laboratory aggression paradigm. Poster session presented at the annual Spring Academic Conference of the Kentucky Psychological Association, Lexington, KY.

Plutchik, R. & Van Praag, H. (1989). The measurement of suicidality, aggressivity and impulsivity. *Progress in Neuro-Psychopharmacological and Biological Psychiatry,* 13(Suppl), S23–S34.

Rapoport, A. & Chammah, A. (1965). *Prisoner's Dilemma.* University of Michigan Press, Ann Arbor, MI.

Reichardt, C. & Coleman, S. (1995). The criteria for convergent and discriminant validity in a multitrait-multimethod matrix. *Multivariate Behavioral Research, 30 (4),* 513-538.

Reynolds, B., Ortengren, A., Richards, J., & de Wit, H. (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences*, *40*(2), 305-315.

Richards, J., Zhang, L., Mitchell, S., & de Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior: Effect of alcohol. *Journal of the Experimental Analysis of Behavior*, *71*(2), 121-143.

Richardson, D. (2005). The myth of female passivity: Thirty years of revelations about female aggression. *Psychology of Women Quarterly*, *29*, 238-247.

Richardson, D., & Green, L. (2003). Defining direct and indirect aggression: The Richardson Conflict Response Questionnaire. *Revue Internationale de Psychologie Sociale*, *16*(3), 11-30.

Richardson, D. & Hammock, G. (2007). Social context of human aggression: Are we paying too much attention to gender?. *Aggression and Violent Behavior*, *12*, 417-426.

Ritter, D., & Eslea, M. (2005). Hot Sauce, Toy Guns, and Graffiti: A Critical Account of Current Laboratory Aggression Paradigms. *Aggressive Behavior*, *31*(5), 407-419.

Romano, E., Tremblay, R., Boulerice, B., & Swisher, R. (2005). Multilevel correlates of childhood physical aggression and pro-social behavior. *Journal of Abnormal Child Psychology, 33(5)*, 565–578.

Rosenzweig, S. (1941). Need-persistive and ego-defensive reactions to frustration as demonstrated by an experiment on repression. *Psychological Review*, 48, 347-349.

Rushton, J., Chrisjohn, R., & Fekken, G. (1981). The altruistic personality and the Self-Report Altruism Scale. *Personality and Individual Differences*, *2*(4), 293-302.Rushton, Chrisjohn, & Fekken, 1981

Schmitt, N., & Stults, D. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement*, *10*(1), 1-22.

Sears, R., Hovland, I., & Miller, N. (1940). Minor studies of aggression: I. Measurement of aggressive behavior. *Journal of Psychology: Interdisciplinary and Applied*, *9*, 275-294.

Siever, L. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry*, *165*(4), 429-442.

Smith, G. (2005). On construct validity: Issues of method and measurement*. Psychological Assessment, 17(4),* 396-408.

Smith, G., Fischer, S., Cyders, M., Annus, A., Spillane, N., & McCarthy, D. (2007). On the validity of discriminating among impulsivity-like traits. *Assessment, 14*, 155–170.

Stålenheim, E., Eriksson, E., von Knorring, L., & Wide, L. (1998). Testosterone as a biological marker in psychopathy and alcoholism. *Psychiatry Research*, *77*(2), 79-88.

Taylor S. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality, 35*, 297-310.

Taylor, S., & Chermack, S. (1993). Alcohol, drugs and human physical aggression. *Journal of Studies on Alcohol*, *Suppl. 11*, 77-88.

Taylor, S. & Hulsizer, M. (1998). Psychoactive drugs and human aggression. In: R. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy,* pp. 139–165, Academic Press, San Diego, CA.

Taylor, S., Vardaris, R., Rawtich, A., Gammon, C., Cranston, J., & Lubetkin, A. (1976). The effects of alcohol and delta-9-tetrahydrocannabinol on human physical aggression. *Aggressive Behavior, 2,* 153-161.

Tedeschi, J., & Quigley, B. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression and Violent Behavior, 1*, 163-177.

Tedeschi, J., & Quigley, B. (2000). A further comment on the construct validity of laboratory aggression paradigms: A response to Giancola and Chermack. Aggression and Violent Behavior, 5(2), 127-136.

Terrell, H., Hill, E., & Nagoshi, C. (2008). Gender differences in aggression: The role of status and personality in competitive interactions. *Sex Roles, 59*, 814-826.

Turner, M. & Jumbelic, M. (2003). Stun gun injuries in the abuse and death of a seven-month-old infant*. Journal of Forensic Sciences, 48 (1),* 180-182.

U.S. Department of Justice (2008). *Criminal victimization in the United States, 2006 Statistical Tables.* Washington, DC: Bureau of Justice Statistics.

Vaillancourt, T., Miller, J. L., Fagbemi, J., Côté, S., & Tremblay, R. E. (2007). Trajectories and predictors of indirect aggression: Results from a nationally representative longitudinal study of Canadian children aged 2-10. *Aggressive Behavior*, *33*, 314-326.

van Goozen, S., Fairchild, G., Snoek, H., & Harold, G. (2007). The evidence for a neurobiological model of childhood antisocial behavior. *Psychological Bulletin, 133(1)*, 149-182.

Veit, R., Lotze, M., Sewing, S., Missenhardt, H., Gaber, T., & Birbaumer, N. (2010). Aberrant social and cerebral responding in a competitive reaction time paradigm in criminal psychopaths. *NeuroImage, 49(4)*, 3365-3372.

Verona, E., Joiner, T., Johnson, F., & Bender, T. (2006). Gender specific gene-environment interactions on laboratory-assessed aggression. *Biological Psychology*, *71*(1), 33-41.

Verona, E., & Vitale, J. (2006). Psychopathy in women: Assessment, manifestations, and etiology. *Handbook of the psychopathy* (pp. 415-436). New York, NY: Guilford Press.

Virkkunen, M., Nuutila, A., Goodwin, F., & Linnoila, M. (1987). Cerebrospinal fluid monoamine metabolite levels in male arsonists. *Archives of General Psychiatry*, *44*(3), 241-247.

Virkkunen, M., Rawlings, R., Tokola, R., & Poland, R. (1994). CSF biochemistries, glucose metabolism, and diurnal activity rhythms in alcoholic, violent offenders, fire setters, and healthy volunteers. *Archives of General Psychiatry*, *51*(1), 20-27.

Warbuton, W., Williams, K., & Cairns, D. (2006). When ostracism leads to aggression: The moderating effects of control deprivation. *Journal of Experimental Social Psychology*, 42, 213-220.

Whiteside, S. & Lynam, D. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30,* 669–689.

Widaman, K. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, *9*(1), 1-26.

Widaman, K. (1992). Multitrait-multimethod models in aging research. *Experimental Aging Research*, *18*(3-4), 185-201.

Wolfe, B., & Baron, R. (1971). Laboratory aggression related to aggression in naturalistic social situations: Effects of an aggressive model on the behavior of college student and prisoner observers. *Psychonomic Science*, *24*(4), 193-194.

Yechiam, E., Goodnight, J., Bates, J., Busemeyer, J., Dodge, K., Pettit, G., Newman, J. (2006). A formal cognitive model of the go/no-go discrimination task: Evaluation and implications. *Psychological Assessment*, *18*(3), 239-249.

Zeichner, A., Frey, F., Parrott, D., & Butryn, M. (1999). Measurement of laboratory aggression: A new response-choice paradigm. *Psychological Reports*, *85*(3), 1229-1237.

Zeichner, A., Parrott, D., & Frey, F. (2003). Gender differences in laboratory aggression under response choice conditions. *Aggressive Behavior*, *29*(2), 95-106.

Zeichner, A., & Pihl, R. (1979). Effects of alcohol and behavior contingencies on human aggression. *Journal of Abnormal Psychology*, *88*(2), 153-160.

Zhou, J., Wang, X., Li, L., Cao, X., Xu, L., & Sun, Y. (2006). Plasma serotonin levels in young violent offenders: Aggressive responding and personality correlates. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *30*(8), 1435-1441.

Zimmer, L. & Morgan, J. P. (1997). *Marijuana myths and marijuana facts*. New York: The Lindesmith Center.

Vita

Joshua Parker Phillips

Date of Birth: November 12, 1980

Place of Birth: Lexington, Kentucky

Education:
University of Kentucky, Lexington, Kentucky
BS in Psychology, 2004
MS in Clinical Psychology, 2007

Honors:
University of Kentucky
Graduation Cum Laude, 2004
University of Kentucky, Department of Psychology
Graduation with Honors in Psychology, 2004
Kentucky Psychological Association
Runner-up, Paul Hager Graduate Student Poster Award, 2009
University of Kentucky, Department of Psychology
Excellence in Clinical Performance Award, 2009

Professional Positions:
University of Kentucky, Lexington, KY
Undergraduate Research assistant, 2002 – 2004

University of Kentucky, Jesse G. Harris Psychological Services Center
Therapist, August 2005 – 2010.

University of Kentucky, Lexington, KY
Teaching assistant, September 2005 – 2010.

University of Kentucky, Testing and Counseling Center, Lexington, KY
Therapist, August 2006 – May 2007;
Process observer in group therapy, Summer 2006.

Eastern State Hospital, Lexington, KY
Psychology Practicum Student, June 2008 – June 2009

Department of Neurology, University of Kentucky Hospital, Lexington, KY
Psychology Practicum Student, August 2007 - August 2008

University of Kentucky, Lexington, KY
APAGS Student Representative, March 2006- March 2007

James H. Quillion Veterans Affairs Medical Center, Johnson City, TN
Psychology Intern, July 2010 – July 2011

Professional Publications:

Giancola, P., Levinson, C., Corman, M., Godlaski, A., Morris, D., Phillips, J., & Holt J. (2009). Men and women, alcohol and aggression. *Experimental and Clinical Psychopharmacology*, 17 (3), 154–164.

Phillips, J. & Giancola, P. (2008). Experimentally-induced anxiety attenuates alcohol-related aggression in men. *Experimental and Clinical Psychopharmacology.* 16 (1), 43-56.

Phillips, J. & Giancola, P. (2008). A psychological and neurobiological perspective on the empathy and aggression relation. In B.P. Reimann (Ed.), *Personality and Social Psychology Research.* Hauppauge, NY: Nova Science Publishers, Inc.

Phillips, J. & Giancola, P. (2007). The effects of contextual empathy on aggressive behavior in men and women. In: A. M. Columbus (Ed.), *Advances in Psychology Research, Vol 49,* 197-208. Nova Science. New York.