Utah State University
DigitalCommons@USU

All Graduate Theses and Dissertations | Graduate Studies, School of

5-1-2011

A Multigroup Analysis of Reintegrative Shaming Theory: An Application to Drunk Driving Offenses

Elizabeth J. Dansie
Utah State University

Recommended Citation
Dansie, Elizabeth J., "A Multigroup Analysis of Reintegrative Shaming Theory: An Application to Drunk Driving Offenses" (2011). *All Graduate Theses and Dissertations.* Paper 857.
http://digitalcommons.usu.edu/etd/857

This Dissertation is brought to you for free and open access by the Graduate Studies, School of at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact becky.thoms@usu.edu.

A MULTIGROUP ANALYSIS OF REINTEGRATIVE SHAMING THEORY: AN APPLICATION TO DRUNK DRIVING OFFENSES

by

Elizabeth J. Dansie

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

Approved:

Dr. Christopher J. Johnson Major Professor	Dr. Maria C. Norton Committee Member
Dr. Scott C. Bates Committee Member	Dr. Michael P. Twohig Committee Member
Dr. Kerry E. Jordan Committee Member	Dr. Byron R. Burnham Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2010

Copyright © Elizabeth J. Dansie 2010

All Rights Reserved

ABSTRACT

A Multigroup Analysis of Reintegrative Shaming Theory:

An Application to Drunk Driving Offenses

by

Elizabeth J. Dansie, Doctor of Philosophy

Utah State University, 2010

Major Professor: Dr. Christopher J. Johnson
Department: Psychology

A restorative justice alternative to crime prevention termed reintegrative shaming theory by Braithwaite has seen increased attention as an alternative to retributive justice, although empirical investigations of its efficacy are limited. The purpose of the present study was to test confirmatory measurement and structural models of reintegrative shaming theory in order to assess the underlying theoretical model and the application of this theory in response to drunk driving offenses. Nine latent constructs were included in these models: reintegration, stigmatization, perceived fairness, self-esteem, shame-guilt, embarrassment-exposure, unresolved shame, offender responsibility, and family support.

Multigroup structural equation modeling was used to assess for measurement invariance of indicators used to measure these nine latent constructs between 724 drunk driving offenders randomly assigned to traditional court processing versus offenders assigned to reintegrative shaming conferencing following arrest. Partial metric and

partial scalar invariance were found. Thus, analyses proceeded by conducting tests for significant differences in the latent means between groups. Offenders assigned to conferencing reported significantly higher mean values on the constructs reintegration, perceived fairness, self-esteem, shame-guilt, and family support, supporting Braithwaite's theory.

Finally, a structural model was hypothesized based upon Braithwaite's theory to assess the relationships between the latent constructs. Three additional structural paths were included to achieve an acceptable model fit. This structural model was found to be partially invariant between groups. As predicted, a higher level of reintegration was associated with greater perceived fairness, while a higher level of stigmatization was related to decreased self-esteem and lower perceived fairness. In turn, greater self-esteem and perceived fairness were significantly related to higher reported experiences of shame-guilt and lower ratings of embarrassment-exposure. Greater perceived fairness also corresponded to lower reported unresolved shame. Finally, greater shame-guilt was significantly related to greater offender responsibility and family support, while unresolved shame was significantly related to less offender responsibility acceptance. The findings from the current study support Braithwaite's hypotheses regarding the importance and benefits of disapproving of the criminal act and not the person, while allowing offenders to accept responsibility for their actions and attempt to remediate the wrong that they committed.

(182 pages)

ACKNOWLEDGMENTS

I would first like to thank my advisor, Dr. Christopher Johnson, for his support, dedication, and time commitment to me and the completion of my degree. I would also like to thank Dr. Jamison Fargo, whose love for research and learning greatly contributed to the love I have for science today. Finally, I would like to give special thanks to my family and friends for all of the encouragement and guidance they have provided me throughout my doctoral program.

Elizabeth J. Dansie

CONTENTS

Page

ABSTRACT iii

ACKNOWLEDGMENTS v

LIST OF TABLES viii

LIST OF FIGURE ix

CHAPTER

I. PROBLEM STATEMENT 1

II. LITERATURE REVIEW 6

Restorative Justice 7
Nature of Shame 25
Empirical Investigations of Reintegrative Shaming Theory 38

III. METHOD 63

Data 63
Procedure 65
Measures 66
Statistical Analyses 72
Hypotheses 87

IV. RESULTS 92

Descriptive Statistics 92
Tests for Factorial Validity 93
Structural Model Testing 104

V. DISCUSSION 114

Summary of Findings 114
Limitations and Future Research Directions 131
Conclusions 138

REFERENCES 140
Page

APPENDICES 156

Appendix A: Covariance Matrix for Offenders Assigned to Conferences ... 157
Appendix B: Covariance Matrix for Offenders Assigned to Court 162

CURRICULUM VITAE 167

LIST OF TABLES

Table		Page
1	Indicator Items Comprising Latent Constructs and Descriptive Statistics	68
2	Indicator Loadings for Configural Baseline Model	95
3	Assessing Factorial Invariance Between Conference and Court Offenders	99
4	Fit Indices for Invariance Tests	100
5	Tests for Latent Mean Differences Between Groups	102
6	Correlations Among Latent Factors	105
7	Fit Indices for Structural Model	106
8	Direct Effects for Structural Model	108
9	Standardized Total and Indirect Effects for Structural Model	112

LIST OF FIGURES

Figure		Page
1	Hypothesized confirmatory measurement model for all offenders	89
2	Hypothesized structural model for all offenders	90
3	Final constrained structural model for the conference group	109
4	Final constrained structural model for the court group	110

CHAPTER I

PROBLEM STATEMENT

Throughout the United States, the rates of crime and repeat criminal offending are incredibly high, with the traditional punitive court system having a modest influence as a deterrent of crime (Langan & Levin, 2002). Specifically 25-67% of offenders arrested in 1994 for crimes ranging from minor misdemeanors to serious felonies were re-arrested after release (Langan & Levin, 2002). This large proportion of individuals committing repeat offenses after being prosecuted through the traditional justice system makes clear the need for a possible alternative form of justice to attempt to stop crime. In response to this problem, a restorative justice alternative to crime prevention termed Reintegrative Shaming Theory (RST; Braithwaite, 1989) has been implemented in various countries across the world, such as New Zealand, Australia, and the United States.

Restorative justice is an alternative form of justice that utilizes mediated conferencing, as opposed to punishment, to deter criminal behavior (Braithwaite, 1999; Llewellyn & Howse, 1998; Van Ness & Strong, 1997). Thus, the offender, victim, and anyone impacted by a crime come together in a physically and emotionally safe environment to discuss the crime, its impact on all parties, and arrive at a successful resolution to the criminal offense. Further, the offender is given the opportunity to accept responsibility and offer appropriate reparation for their behavior (Marshall, 1999). This form of justice is very different from the traditional retributive court process, where

punishment is utilized to deter future criminal behavior (Umbreit, Coates, & Vos, 2002; Zehr, 1985).

While many varied forms of restorative justice are practiced and have been found to be effective alternatives to retributive justice practices (Latimer, Downden, & Muise, 2005; Nugent, Williams, & Umbreit, 2004), proponents of a particular form of restorative justice emphasize the importance of the emotional dynamics occurring in a restorative conference. Through Braithwaite's RST (1989; Ahmed, Harris, Braithwaite, & Braithwaite, 2001), successful reparation of a crime occurs via a sequence of shame, acknowledgement, and transformation. Within this *ideal* framework, the offender commits a wrongdoing, and although the offender may experience shame both before and after the conference, ideally the offender experiences shame during the conference. The offender then can choose to acknowledge the experience of this emotion within the caring environment of the conference setting surrounded by loved ones and attempt to make restitution for their crime. The acceptance of responsibility and attempts of apology then defuses the offender's shame. Ideally, once a restitution agreement is made between the offender and victim, the relationship is transformed from one of disrespect towards a relationship of equality.

Various core concepts are central to the understanding and success of RST, including *shaming* as a regulatory practice and the experience of *shame* as an emotional experience. Importantly, the type of shame an individual experiences (i.e., more adaptive *shame-guilt,* or less adaptive *unresolved shame* or *embarrassment-exposure*), can also influence the success of a reintegrative shaming conference (Harris, 2006).

Along with shaming and shame, the terms *reintegration* and *stigmatization* are essential components of the theory. Reintegration as defined by Braithwaite (1989) involves respectfulness and approval of the offender as a person, while conveying disapproval for the criminal act. In contrast, stigmatization involves disapproving of the offender as a person and labeling the individual as a deviant. Following, shaming can be reintegrative or stigmatizing in nature and Braithwaite (1989) predicts that reintegrative forms of shaming are likely to lead to fewer future offenses in comparison to shaming conducted in a stigmatizing fashion.

Although RST has been posited to be an effective alternative to retributive justice and has hence been implemented in various cities throughout the world, little empirical testing of RST has been conducted. Specifically, after conducting a literature review of the empirical investigations of RST, the present author identified only 18 total studies relevant to RST, while only 11 directly tested the effectiveness of the application of RST. Further, while these studies were conducted in the context of multiple offense-types (e.g., bullying, white-collar crime, minor criminal offenses), none of the studies attempted to assess the underlying theoretical model as adapted from Braithwaite's (1989) hypotheses. Thus, the goals of the proposed study were to: (a) test the theoretical model underlying RST, and (b) assess the effectiveness of this theoretical model when implemented in response to actual criminal offending.

This review of the literature further guided the present study as certain central tenets of the theory have received little or no empirical attention, and thus were included in the present model. In particular, the impact of stigmatizing versus reintegrative

shaming practices on the acceptance of responsibility by the offender, and perceived family support following an offense have not been investigated. This focus on responsibility acceptance by the offender is essential to assess, as Braithwaite contends that allowing the offender to accept responsibility on their own (without pressure from others) will lead to the most successful outcome for both victim and offender. In terms of family support or interdependent relationships, none of these previous investigations assessed family support as an outcome variable. This finding was surpring because Braithwaite hypothesized that crime rates are lower in societies with individuals in more interdependent relationships. Thus, feelings of strong family support following an offense may likely lead to less future criminal behavior.

Finally, the impact that individual differences in offenders on reintegrative versus stigmatizing shaming has received little empirical attention. Only one study (Ahmed & Braithwaite, 2004) looked at the impact of self-esteem as an individual difference variable influencing the impact of RST on offenders. Self-esteem is an important construct to assess in this realm, because self-esteem is closely linked to how individuals react to shaming, the type of shame they experience, and thus their likelihood of committing future offenses (Leary, 2005; Rosenberg, 1979). Thus, it appeared that greater attention should be paid to the mediating impact of self-esteem.

For the current study, data were used from the *Reintegrative Shaming Experiments (RISE) in Australia, 1995 – 1999* (ICPSR; Sherman, Braithwaite, Strang, & Barnes, 2000), which was sponsored by the United States Department of Justice: National Institute of Justice. These data were from the only randomized-controlled trial of the

efficacy of RST in reducing criminal recidivism that has been conducted (Tyler, Sherman, Strang, Barnes, & Woods, 2007), where 724 drunk driving offenders were randomly assigned to RST conferencing or traditional court processing after being arrested. Although the researchers collected data on a variety of constructs hypothesized by Braithwaite (1989) to reduce crime (i.e., perceived fairness, self-esteem, family support, and offender responsibility), the relationships between these constructs were left untested, and a reduction in recidivism rates was the primary dependent variable investigated.

Thus, a secondary data analysis using multigroup structural equation modeling (SEM) techniques was conducted with the Tyler and colleagues (2007) data to address the proposed goals of the study. Specifically, multigroup SEM was used to compare offenders randomly assigned to RST conferencing or court processing on nine latent constructs hypothesized by Braithwaite to reduce criminal offending (i.e., reintegration, stigmatization, fairness, self-esteem, shame-guilt, unresolved shame, embarrassment-exposure, family support, and offender responsibility). This statistical technique allowed for the testing of the underlying theoretical model of RST, the comparison of conference and court offenders on their mean scores on the aforementioned latent variables (i.e., the application of RST in response to crime), and allowed the researcher to test if the causal structure among the latent variables proposed by Braithwaite was empirically supported.

CHAPTER II

LITERATURE REVIEW

According to the 1997 crime census conducted by the United States Department of Justice: Bureau of Justice Statistics (Langan & Levin, 2002), an estimated 1,408,337 violent crimes were reported that year, while 9,843,481 property crimes were reported. Moreover, many individuals apprehended for law-breaking are re-arrested: of the 272,111 offenders discharged from prison in 1994, it is estimated that 67.5% were rearrested for a felony or serious misdemeanor, 46.9% were reconvicted, and 25.4% were re-sentenced for a new crime (Langan & Levin, 2002). It is clear from these annual statistics that many individuals are negatively impacted by crime and re-offense rates are high. Thus, investigations are needed to determine which factors reduce crime and are empirically supported, and how effective these factors are when implemented in real-world settings. Problematically, when empirical testing on various forms of crime deterrence has been conducted, the testing of the actual theoretical models underpinning these approaches has widely been ignored. Thus, when attempts at reducing crime are unsuccessful (e.g., there was no significant decrease in rates of recidivism), it is impossible to determine why the failure occurred.

A restorative justice alternative to crime prevention termed RST (Braithwaite, 1989) has been implemented in an attempt to reduce criminal offending. Like many models of justice and crime prevention, little empirical testing of the theoretical model underlying this form of justice has been conducted. Thus, the goals of the present study were to: (a) test the theoretical model underlying RST, and (b) assess the effectiveness of

this theoretical model in response to actual criminal offending using multigroup SEM. Before a discussion of the present study is provided, an overview of different forms of crime deterrence will be provided, followed by a more detailed discussion of the tenets of RST and how RST is hypothesized to be an especially effective mechanism by which crime can be curtailed.

Restorative Justice

Restorative Versus Retributive Justice

One important factor for deterring possible first and re-occurring offenses is the type of justice being served for a crime. Justice (considered here to be an administration of the law or infliction of punishment (Oxford English Dictionary, Simpson & Weiner, 1989]) can take many forms. However, the question remains: "What form should this justice take?" The most common type of justice currently in use in the United States is *retributive* justice (Llewellyn & Howse, 1998). The focus of retributive justice is on the process of establishing the guilt of an offender, and then imposing some form of punishment for that offense (Zehr, 1985). Ultimately, this process is used to achieve the goal of establishing or re-establishing the social equality in the relationship between the offender and the victim of the crime, where each individual in a relationship has a right to be treated with equal respect and dignity. The focus of retributive justice is on establishing the *social* equality between victim and offender. Thus, this form of justice recognizes that a crime can affect the direct victims of a crime as well as impact other people indirectly, such as family or community members.

The use of punishment as the main instrument to restore social equality is flawed for various reasons (Braithwaite & Pettit, 1990). Firstly, the type of punishment that is dealt does not account for contextual differences (i.e., the same punishment is generally dealt for a specific type of offense) inherent in each individual offense. Thus, there is no consideration for the type of punishment that may best suite the offense and be the best option to restore social equality. Secondly, because the goal of retributive justice is on establishing guilt and meting out punishment, the focus is on the outcome of what has happened, rather than the process of what can be done to restore the relationships between those impacted by the offense (Zehr, 1990). Thus, in most cases, the needs of the victim are highly ignored, and crime is instead treated as an offense against the state, rather than an offense against individuals. Thirdly, punishment is nonvoluntary and is imposed on an offender, making retributive justice an isolating experience that actually increases the social inequality between victim and offender. No responsibility is placed upon the offender to right the wrong that was done, but rather the offender is required to passively endure the punishment (Llewellyn & Howse, 1998).

While retributive justice is most commonly practiced in the United States, an alternative form of justice that has existed for centuries, known as *restorative* justice, is gaining popularity (Braithwaite, 1999; Llewellyn & Howse, 1998; Van Ness & Strong, 1997). As summarized by Llewellyn and Howse (2002), precolonial African societies, contemporary Japanese societies, and aboriginal cultures in Canada and Australia have utilized principles of restorative justice in resolving the sequelae of a crime. Through the lens of restorative justice, the occurrence of crime is not just a violation of the law, but it

is an occurrence that can harm individuals or communities and destroy relationships (Van Ness & Strong, 1997; Zehr, 1990). Similar to retributive justice, the goal of restorative justice is to restore the equality in social relationships, but this end is achieved through very different means.

This alternative form of justice aims to overcome many of the weaknesses inherent in retributive justice. Restorative justice, like retributive justice, emphasizes that more individuals are impacted by crime than the victim and offender (Llewellyn & Howse, 2002; Marshall, 1999; Zehr, 1990). Importantly, unlike retributive justice, restorative justice recognizes that imposing a certain punishment for an offense actually increases the separation between parties. Rather than applying a stringent set of rules or punishments to each offense, justice is relational and context-dependent: consideration is given to all involved parties and all aspects of an offense and a dialogue between all parties is necessary in order to restore social equality. Subsequently, each approach to restoring a relationship may be different. By taking into account the particular aspects of each relationship and of the individual parties involved, restorative justice focuses on the process by which a relationship can be restored, rather than just if guilt was established and punishment was dealt (Marshall, 1999). Further, rather simply aiming to return a social relationship back to the status quo, restorative justice recognizes that an imbalance in the relationship may have existed prior to an offense. Thus, the focus is also to *establish* social equality where none previously existed. Within the restorative justice paradigm, due to the context-dependent nature of the justice process, a relationship can

be restored with a variety of means beyond that of mere restitution like payment or jail time (Llewellyn & Howse, 2002).

In addition to attempting to restore each relationship through varied means, practitioners of restorative justice emphasize that it also differs from retributive justice in that punishment is not nonvoluntarily imposed on an offender. Instead, within the restorative justice paradigm, the offender is an active, willing participant in the restorative process. Within restorative justice, responsibility for an act cannot be ignored. With the offender being an active participant in the process, the offender is no longer isolated from the relationship, which commonly occurs when punishment is dealt, and the focus on the relationship is maintained (Llewellyn & Howse, 2002; Marshall, 1999). When taking these principles of restorative justice into account, it appears that Marshall's (1996) definition most aptly defines restorative justice. According to Marshall, "Restorative justice is a process whereby all the parties with a stake in a particular offence come together to resolve collectively how to deal with the aftermath of the offence and its implications on the future" (p. 37, cf. Braithwaite, 1999, p. 5).

Process of Restorative Justice Conferencing

Restorative justice is an umbrella term encompassing three related categories of justice, including circles, conferences, and victim-offender mediation (Bradshaw & Roseborough, 2005; Latimer et al., 2005; Maxwell & Morris, 1993). Although each of these categories of restorative justice is slightly different, they all share the common goal of repairing the harm caused by a crime through re-establishing the social equality between parties. Through all three forms of restorative justice, all parties affected by an

offense (i.e., the wrongdoer, the victims, and the community) have an opportunity to come together in a conference-type setting to mutually decide how to provide appropriate reparation for a crime. Further, restorative justice can be implemented through any of these three forms during any stage of the criminal justice system: restorative justice can be used precharge, postcharge, presentence, postsentence, or if the offender is on parole (Latimer et al., 2005).

Although restorative justice conferences vary depending upon the context of each individual conference, Moore and McDonald (2000) provide a concise description of the process of conferencing. According to Moore and McDonald there are five stages: (a) the participants and the facilitator arrive and one individual that has previously agreed to begin the dialogue starts discussing the problem; (b) once the first person has given their perspective, the facilitator encourages those most directly affected to give their account of the events that transpired, followed by those less directly impacted; (c) once everyone has been given an opportunity to speak, the offender is given an opportunity to offer some form of symbolic reparation; (d) the facilitator encourages those most directly impacted to discuss what they believe are appropriate actions for reparation, and this begins a negotiation between all parties until an agreement is reached, and a plan of action constructed; and (e) the official agreement is signed. Ideally, an outcome is reached by all parties involved, with the intention of restoring the relationship between the parties to one of respect and equality. In contrast to retributive justice where context is not considered when punishment for a crime is administered, the determined outcome of restorative justice is context specific (i.e., thus varies widely between parties) and is

reached through the process. Commonly, the wrongdoer is given an active role in recommending what constitutes restoration. Ultimately, the wrongdoer must do whatever the victim requires to feel that an acceptable resolution has been reached (Llewellyn & Howse, 2002), such as community service, monetary payment, or even a simple apology. Again, the goal is to establish or re-establish the social equality between parties. Punishment typically does not meet this aim and is not commonly practiced as an acceptable resolution within the restorative justice paradigm.

For this restorative process to lead to a satisfactory conclusion, theoretically, all restorative conferences must include three main elements. First, the process must be voluntary for all those involved. This voluntariness is ensured by requiring the offender to acknowledge the occurrence of the event at the onset of the restorative process. Second, the process must ensure that the offender is truthful, that all parties hear this truth, and that the offender accepts responsibility for their actions while being willing to openly discuss their behavior. Third, the process must involve a face-to-face "encounter" and "engagement" between all parties. Bringing victim, offender, and community members together serves multiple purposes, including the greater likelihood that the facts of the situation will be verbalized, stereotypes that victim/offender have of each other can possibly be dispelled because parties are given the opportunity to engage with one another, and emotions or feelings can be expressed and received (Llewellyn & Howse, 1998). Although a facilitator is always present, the participants involved in the dialogue must drive the entire process.

Effectiveness of Restorative Justice: Narrative and Quantitative Reviews

As part of establishing or re-establishing the social equality between parties, restorative justice emphasizes reparation for the victim while concomitantly treating the offender with respect and fairness (Van Ness & Strong, 1997). Consequently, proponents of restorative justice have hypothesized that restorative justice should lead to satisfactory outcomes for both victim and offender, such as greater victim and offender satisfaction with the process and outcome, lower recidivism rates, and greater restitution compliance (Latimer et al., 2005). Further, these positive outcomes are hypothesized to be equally representative of juvenile and adult offenders. As summarized by Bonta, Wallace-Capretta, Rooney, and McAnoy (2002), a very large body of literature is currently being developed investigating restorative justice on a multitude of outcomes. Several authors have attempted to synthesize the empirical findings through narrative reviews (Braithwaite, 1999; Latimer & Kleinknecht, 2000; Marshall, 1999). Although useful, these narrative reviews fail to determine overall effect sizes for restorative justice processes, thus making it difficult to quantify the effect of restorative justice on the outcomes under study.

To address the need for a quantitative measure of the effectiveness of restorative justice, multiple empirical studies have been conducted to assess the impact restorative justice techniques have in the area of crime prevention and crime control. Specifically, when conducting a literature review on the empirical studies of restorative justice, the author of the present study identified 84 articles published in peer-reviewed journals. The methodology used in these studies varied widely (e.g., correlational investigations,

quasi-experiments, and even a few more rigorous experimental designs), as well as the outcome under investigation (e.g., reduced recidivism, victim satisfaction, restitution compliance). The findings from three well-conducted meta-analyses on the effectiveness of restorative justice in comparison to other traditional forms of punishment will be discussed (Bradshaw & Roseborough, 2005; Latimer et al., 2005; Nugen et al., 2004). Although each meta-analysis employed different inclusion criteria (e.g., differing definitions of "recidivism"), that lead to the inclusion of different empirical studies in some cases, each meta-analysis conducted arrived at a similar result: restorative justice practices were deemed more effective on multiple outcome measures in comparison to other forms of justice.

Specifically, when investigating the overall effectiveness of restorative justice on recidivism rates in juvenile offenders in 23 studies, Bradshaw and Roseborough (2005) found that the average Cohen's *d* effect size was 0.26, indicating that participation in restorative justice processes contributed a 26% reduction in recidivism. The possible influence of seven moderating variables on this effect size estimate was also investigated (i.e., quality, type of comparison group, offense-type, definition of re-offense, source of study, sample, and length of follow-up). Only the type of control group (e.g., those that were referred to a restorative conference but refused versus an alternative program such as diversion programs) was found to be a statistically significant moderator of the effect size. As discussed by these researchers, this effect size is more than double the reported effect size of 0.10 from a meta-analysis of traditional youth justice programs conducted by Lipsey in 1995. Similarly, the overall effect of participation in restorative justice on

recidivism in juvenile delinquents was also investigated through meta-analytic techniques by Nugent and colleagues (2004). With 15 individual studies meeting their inclusion criteria, these authors found that the odds of reoffending by restorative justice participants in comparison to those juveniles who participated in a nonrestorative process were .70, concluding that restorative justice should continue to be practiced.

Similar conclusions were reached in the most recent meta-analysis on restorative justice conducted by Latimer and colleagues (2005), investigating the effects of restorative justice on juveniles and adult offenders in 22 unique studies. Using a phi coefficient to assess the magnitude of the relationship between participation in restorative justice and the outcome under study, these authors evaluated the outcomes of victim and offender satisfaction, recidivism, and restitution compliance. Results indicated that more positive outcomes were obtained for those groups of individuals that participated in restorative justice conferences in comparison to the traditional justice system on all four outcomes under investigation (i.e., victim satisfaction, offender satisfaction, recidivism, and restitution compliance). When considering the overall conclusions generated by all three meta-analyses, the consistency among findings points to the possible positive effects of restorative justice practices on reducing recidivism rates and other important outcomes (e.g., greater victim satisfaction).

Reintegrative Shaming Theory

Restorative justice is a broader term that encompasses a wide range of different restorative practices distinct from retributive styles of justice, such as victim-offender mediation and peacemaking circles (Maxwell & Morris, 1993). Although restorative

processes have been found to be effective alternatives to retributive justice practices in curtailing crime (Bradshaw & Roseborough, 2005; Latimer et al., 2005; Nugent et al., 2004), proponents of a particular form of restorative justice emphasize that it is the experience and expression of emotions within the restorative conference that makes restorative justice more effective in comparison to retributive justice (Braithwaite, 1989). In particular, the experience and proper expression of the negative emotions shame, guilt, and embarrassment by all parties in the restorative dialogue serves as the focus upon which successful reparation occurs. Through an integrated theory, coined RST by Braithwaite (Ahmed et al., 2001; Braithwaite, 1989), successful reparation of a crime occurs via a sequence of shame, acknowledgement, and transformation within a conference setting. When developing this theory, Braithwaite integrated the then dominant criminological theoretical traditions of labeling theory (e.g., Tannenbaum, 1938), subcultural theory (e.g., Cohen, 1955; Miller, 1958), control theory (e.g., Durkheim, 1951; Reckless, 1967), opportunity theory (e.g., Merton, 1957), and learning theory (e.g., Sutherland & Cressey, 1978), and has since refined the theory in accordance with empirical findings (Ahmed et al., 2001). In order to more fully describe RST, the components of the theory will be defined and discussed in relation to crime deterrence, supporting the argument that reintegrative shaming may be a viable alternative to retributive justice.

Components of RST

Shaming and shame. The first key terms to be emphasized are *shaming* and *shame*. These core concepts share common features, but are distinct in that shaming is

considered a regulatory practice, while shame is considered an emotion experienced by any of the parties involved in the conference. According to Braithwaite (1989), shaming means "all social processes of expressing disapproval which have the intention or effect of invoking remorse in the person being shamed" (p. 100). Although all shaming is a form of social disapproval, shaming can actually operate at two levels. The first form of shaming involves an expression of disapproval following a transgression that communicates that the offender has lost esteem in the eyes of the individual communicating the disapproval, and thus signals possible rejection. This shaming can be subtle or more overt, such as a frown, laugh, shaking of the head, a direct verbal confrontation, admonition by a judge, or can be done indirectly through gossip. The mode in which shaming is conducted is culturally specific. Further, shaming may be communicated through formal punishment or be accompanied by punishment.

The second form of shaming occurs through societal practices that build consciences in children. Through the socialization of children in a family or community network, the disapproval that shaming communicates leads to a development of a conscience that internally deters criminal behavior (Ahmed & Braithwaite, 2004, 2005; Ahmed et al., 2001; Braithwaite, 1989, 1999). Thus, through socialization within the family and community context, children first learn to be responsive to external controls communicated through the parents' overt disapproval. Over time, internal controls of behavior become more effective as the child develops their own standards of "right" and "wrong" behavior, dependent upon the culture in which the socialization occurred. This development of a conscience as a form of social control is hypothesized to be a superior

crime deterrent in comparison to overt punishment because committing an act that will be disapproved of by valued others in the community is a symbolic reflection of a "flaw" in the character of the individual committing the act (Braithwaite, 1989).

As hypothesized through RST, shaming in any form is posited to lead to the experience of the negative emotion shame, wherein the threat of experiencing this painful emotion is hypothesized to deter criminal behavior. Copious amounts of research have been conducted investigating the true nature of shame in a variety of different academic arenas, although no consensus has been reached regarding one definition of this emotion. For example, work has been conducted in the clinical realm (Greenberg & Paivio, 1997; Lewis, 1971; Lindsay-Hartz, 1984), psychology of emotion (Leary, 2007; Tangney & Dearing, 2002; Tomkins, 1963, 1987), criminology (Braithwaite, 1989; Gilligan, 2003) and evolutionary psychology (Gilbert, 1997; Scheff, 1990).

In an attempt to summarize this vast array of literature, Ahmed and colleagues (2001) developed three conceptualizations of shame, which are the: (a) social threat conception, (b) personal failure conception, and (c) ethical conception. Briefly, researchers theorizing from the social threat conception of shame assert that shame is elicited when we believe that our behavior may elicit social rejection from others in our social group. Shame through the personal failure conception is defined as being elicited when we believe that we have failed to live up to some personal standard, thus our entire self is considered a failure. Finally, shame as an ethical threat incorporates the first two conceptions by centering on the assumption that shame is elicited when we believe we have committed a wrongdoing. This behavior is considered a wrongdoing because the

social group upon which we belong has a moral code that states that the behavior is wrong. Thus, we may be rejected because of that wrongdoing. This type of shame differs from the social threat conception because it focuses on the violation of a shared moral code that elicits shame, beyond the mere threat of social rejection (Ahmed et al., 2001; Harris & Maruna, 2006).

Braithwaite (1989) emphasizes the central role that important individuals in our social group (e.g., family and/or community members) play in leading to the success of a reintegrative conference. Thus, within the context of the present study, shame is hypothesized to be elicited from shaming from *valued others* in our social group. This shame results because of something about ourselves that elicited real or perceived social rejection or disapproval from those important others. Furthermore, from an evolutionary perspective of emotions, individuals have evolved the need to be accepted in the social group for survival and the experience of shame serves as a warning that individuals may be socially excluded (Leary, 2005; Scheff, 1990). In turn, this warning signal motivates us to behave in compliance with group norms to avoid further shaming, avoid the experience of this painful emotion, and deter criminal behavior. This view regarding shame most closely corresponds to a combination of Ahmed and colleagues' (2001; Harris & Maruna, 2006) social threat conception and ethical conception of shame. Like the ethical threat conception, it is recognized that the standards upon which our behavior is judged is based partly upon a shared moral code within the social group, but shame can be elicited for reasons other than a violation of a moral standard (e.g., shame can be elicited due to a feeling that we are unattractive).

Reintegration and stigmatization. Along with shaming and shame, the terms *reintegration* and *stigmatization* are essential components of the theory. Particularly, shaming is followed predominantly with either acts of reintegration or acts of stigmatization by the shamer; both reintegrative and stigmatizing shaming may be displayed by the same individual. As specified by Braithwaite (1989), when paired with shaming, reintegration involves respectfulness and approval of the offender as a person, while disapproving the criminal act. Efforts are made to terminate this disapproval through rituals or ceremonies of forgiveness to decertify the criminal as a deviant and integrate the criminal back into the community by communicating that the social bonds between the shamer and shamed have not been severed. Braithwaite (1989; Ahmed et al., 2001) hypothesized that this type of shaming will result in lower rates of crime because the offender is not ostracized from those who provide social support and who may express further disapproval of criminal acts. Thus, the offender does not have to turn to individuals who approve of criminal behavior for social support, groups which Braithwaite refers to as criminal subcultures (1989).

In contrast to reintegration, stigmatization involves disapproval of the offender as a person and labeling the individual as a deviant; this disapproval is characterized by disrespect, and ceremonies are not conducted to decertify the individual as a criminal. Subsequently, the offender is outcast from their social group. Braithwaite (1989) specified that this type of shaming may result in increased criminal behavior because the individual is rejected by their own social group who disapproves of crime. Stigmatized

offenders will then rely upon and be further influenced by individuals in criminal subcultures who do not consider crime an elicitor of shame.

It should be noted that in the original conceptualization of reintegration and stigmatization within RST, Braithwaite (1989; Makkai & Braithwaite, 1994) hypothesized that reintegration and stigmatization are independent of shaming, and are on opposite poles of the same continuum. Thus, the type of shaming exhibited in a restorative conference could be plotted somewhere along this continuum. Upon testing the dimensionality of these constructs using confirmatory factor analysis, Braithwaite (Ahmed et al., 2001) revised this assumption: although the shaming construct did display very small correlations with reintegration and stigmatization (as hypothesized), the reintegration by stigmatization bivariate correlations were small and nonsignificant and therefore deemed independent, rather than being polar opposites. This evidence provides support for the assumption that a restorative dialogue or conference can be both reintegrative and stigmatizing, as mentioned above.

In contending that reintegrative shaming will be more effective at reducing criminal behavior, Braithwaite and Mugford (1994) sought to identify the characteristics of reintegration ceremonies that make them successful. This fieldwork was conducted in Auckland, New Zealand, and Australia with the data being gathered from conferences of 23 juvenile offenders. From their observations, Braithwaite and Mugford identified 14 conditions of successful reintegration ceremonies and many of these characteristics reflect Braithwaite's assumptions regarding reintegrative shaming. These 14 conditions will be summarized as follows: (a) the criminal act is defined as wrong and irresponsible,

and is uncoupled from the person, (b) denunciation of the crime is done by and in the name of the victim, (c) the victims, offenders, and family members, or the nonauthoritative actors, are in the control of the way in which the conference proceeds, (d) the perpetrator expresses his/her support of the values upheld in the law and the interests of the victim, (e) the physical and emotional distance between conference participants is limited while empathy between individuals is enhanced, (f) the separation of the victim and any fear or shame experienced is terminated by rituals of forgiveness, (g) the ceremony is flexible and context-specific, and (h) reintegration agreements must be enacted as promised.

Clearly, many criminal trials do not exhibit these characteristics, and could instead be considered "degradation ceremonies" (Garfinkel, 1956). These ceremonies follow a sequence of disapproval-degradation-exclusion, in contrast to reintegrative ceremonies that follow a pattern of disapproval-nondegradation-inclusion (Braithwaite & Mugford, 1994). Presented in contrast to reintegrative ceremonies, Braithwaite and Mugford also discuss eight characteristics of degradation ceremonies that are present in many courtroom proceedings. For example, within these degradation ceremonies, the crime and the offender are *coupled*, while the offender is distanced from family supporters as well as the victim. The offender, in fact, stands out among all those present at the trial, which makes the expression and experience of empathy difficult.

Interdependency and communitarianism. While shaming can occur in the context of the family environment (which leads to conscience building) or in the context of a restorative conference, the societal conditions in which a person lives can also be

more or less conducive to reintegrative forms of shaming. Two interrelated concepts are essential to consider: *interdependency* and *communitarianism*. Individuals in more interdependent communities rely on and are relied upon by others to achieve valued outcomes. More interdependent relationships make individuals more susceptible to shaming and to the experience of shame because each person is highly reliant on being accepted in the social group. Thus, any action that prompts possible rejection or disapproval from those in the social group is likely to elicit the negative emotion shame in the offender. Individual characteristics that are most highly associated with interdependency include being employed, commitment (e.g., to friends, family members, a job), being under 15 or over 25 years of age, and being female (Braithwaite, 1989).

Relatedly, collectivism, (or communitarianism according to Braithwaite) is a condition of societies wherein individuals are enmeshed in interdependent relationships. This type of society is characterized by a high value placed on the needs of others and the social group, and individuals develop personal attachments to others in their community that take precedence over individual needs. As emphasized by Braithwaite (1989), interdependency and communitarianism are highly relevant to RST. Communitarian societies are likely to foster shaming because people are highly involved in each others' lives. Further, individuals in communitarian societies are more likely to live within a network of supportive relationships that use *reintegrative* shaming, because it is more difficult in interdependent societies (which are commonly smaller in size) to be able to simply categorize someone as a criminal and treat them as an outcast (Braithwaite, 1989).

When stating why reintegrative shaming works as a crime deterrent, Braithwaite succinctly expressed the critical importance of supportive relationships:

> Both the specific and general deterrent effects of shame will be greater for persons who *remain* [italics added] strongly attached in relationships of interdependency and affection because such persons will accrue greater interpersonal costs from shame. This is one reason why reintegrative shaming makes for more effective social control than stigmatization. (1989, p. 81)

Thus, strong family and/or community support can serve as a buffer from being outcast if an individual commits a crime, which Braithwaite hypothesized will lead to less future crime because offenders will not need to turn to criminal subcultures for support.

Summary of RST. Within Braithwaite's (1989; Ahmed et al., 2001) theoretical conceptualization of reintegrative shaming, once an offender commits an offense, some form of shaming or disapproval is communicated to the offender regarding their actions. This shaming can be relayed more overtly or covertly, and typically has its greatest impact if communicated from a valued other, such as a close family member, friend, community member, or any individual upon which the criminal is interdependent. Any form of disapproval connotes shaming, and this shaming can be more reintegrative or stigmatizing in nature, hypothetically leading to very different outcomes for the likelihood that the offender will commit a future crime. Shaming that is communicated reintegratively disapproves of the act while accepting the person (i.e., leading to more positive outcomes for victim and offender), while in contrast, shaming that is stigmatizing condemns the entire person based upon their criminal actions (i.e., leading to more negative outcomes). This shaming then leads the offender to experience the emotion shame and deters the criminal from committing future offenses in order to avoid

feeling this negative emotion. Typically, reintegrative shaming is more effective and more likely to be practiced in communitarian societies characterized by interdependent relationships between members, as individuals in these communities more likely rely upon and are themselves relied upon by others in the community for physical, social, and emotional support than communities that are independent.

Nature of Shame

Although Braithwaite (1989) hypothesized the potential positive impact of the practice of reintegrative shaming because of the offender's *experience* of the shame emotion, not all researchers agree with this positive nature of shame. Instead, some researchers (Dearing, Stuewig, & Tangney, 2005; Gilligan, 2003; Tangney & Dearing, 2002) support the notion that it is the experience of the emotion *guilt*, not shame, that is the emotion that leads to positive outcomes such as lowered rates of recidivism, and state that shame can be a very destructive emotional experience. Still other researchers (Ahmed et al., 2001; Harris, 2003) argue from empirical investigations of the factor structure of shame that shame and guilt are not actually distinct emotions, but instead there are three forms of shame that mediate the relationship between the type of shaming that is communicated and the likelihood of future adaptive or maladaptive resolutions to a crime. A discussion of these topics will be provided in order to expand upon the conception of shame upon which the present study is based.

Shaming and Shame as Destructive Experiences

Throughout the psychological and criminal justice literature, there is no consensus regarding definitions of shame and guilt, nor if these constructs are truly distinct. A number of theorists and researchers have attempted to distinguish between these emotions on a variety of dimensions, including a focus on the self versus the other, being angry at oneself versus angry at one's behavior, a focus on internality versus externality, the behavior occurring privately or publicly, and immoral behavior versus nonmoral behavior (see Tangney & Dearing, 2002, for a review). The most widely endorsed distinction to be made of shame and guilt is based upon work by Lewis (1971) that was popularized in empirical investigations of shame and guilt by Tangney (Tangney & Dearing). From Lewis' conceptualization,

> The experience of shame is directly about the *self*, which is the focus of evaluation. In guilt, the self is not the central object of negative evaluation, but rather the *thing* done or undone is the focus. In guilt, the self is negatively evaluated in connection with something but is not itself the focus of the experience. (p. 30; emphasis in original)

Shame is considered a focus on the entire self by the self, or a real or imagined other, (Lewis's phrases, "How could *I* do that?"), while a person experiencing guilt focuses on their behavior or action ("How could I do *that*?").

Consistent with Lewis's conceptualization of shame and guilt, many researchers in the fields of criminology, sociology, as well as experimental psychology have adopted the stance that shaming and shame are strongly related to negative and possibly destructive outcomes (Dearing et al., 2005; Gilligan, 2003; Maxwell & Morris, 2002). For example, in the field of psychology a variety of studies have been conducted on the

association between shame and many dependent variables, such as depression (Luyten, Fontaine, & Corveleyn, 2002), social anxiety (Grabhorn, Stenner, Stangier, & Kaufhold, 2006), perfectionism (Stoeber, Harris, & Moon, 2006), and alcohol and drug use (Dearing et al., 2005). The results of these studies suggest that proneness to respond with the emotion shame is significantly associated with each of these harmful outcomes. Further, in a review of the conclusions drawn regarding the influence of shame on outcomes such as empathy and self-esteem, Tangney and Dearing (2002) have drawn similar conclusions and stated that shame is a "dark" immoral emotion, and that it is instead the experience of guilt that leads to empathic responses towards individuals upon which we have harmed.

Similarly, in the fields of criminology and sociology, caution is urged regarding the experience and practice of shaming. For example, prison psychologist Gilligan stated that shame is "the basic psychological motive, or cause, of violent behavior" (2003, p. 1154). Further, sociologist Scheff (1988) argues that the experience of shame can lead to anger and violence through recursive shame-shame and shame-rage spirals, where the experience of shame leads to more shame, and when unacknowledged, can be directed outwardly at others in the form of anger. The potential for negative outcomes resultant from shaming and shame are emphasized from an applied context by Maxwell and Morris (2002), as well as Van Stokkom (2002). These researchers practice restorative justice and emphasize that any type of shaming can be construed as stigmatizing. Further, they contend that shaming as a crime deterrent may actually result in a greater

risk of blocked communication between victim and offender, rather than a reparation of the damage done by the offense.

From the viewpoint of these researchers, it would be detrimental and potentially harmful to use shame inductions as a remedial action to crime. The question thus arises regarding why Braithwaite and colleagues (1989; Ahmed et al., 2001) would suggest that shaming and shame could serve as an effective crime deterrent and should be practiced in restorative settings. Part of this controversy stems from a mismatch between how shame is conceptualized by researchers. Although multiple definitions of shame abound, Ahmed and colleagues (2001) emphasized that the diversity in which shame is conceptualized can be summarized by three points: "(a) how we feel others think of us, (b) how we feel about ourselves, and (c) the views about what is ethically shameful that are shared by us with others" (2001, p. 79). Inherent in this synopsis is the notion that shame not only involves an emotional reaction to a violation of what we and other individuals consider ethical, or a failure to live up to our own personal standards, but also centers around the importance of our social relationships with others. Thus, the experience of shame is linked to an increase in adherence to valued social norms, when failure to comply with those shared values (e.g., treating others with respect, not hurting or stealing from others) not only signals that we are not living up to our own personal standards of what is ethical, but that our social group will disapprove of our violation. This view is in stark contrast to how shame is operationalized in Tangney and colleagues' (2002) model, where shame is treated as a proneness to react to negative evaluation with a global devaluation of the self.

Not only are there multiple conceptualizations of shame in the literature, but there also seems to be a controvery about what is regarded as "shaming" in Braithwaite's theory. It appears that when individuals such as Gilligan (2003) speak of shaming, they are referring to stigmatizing forms of shaming, which Braithwaite would agree are potentially destructive. For example, Gilligan emphasized that shaming occurs when individuals "reject," "abandon," and "ignore" others, treat others with "indignancy," and label them as "weak," "unskilled," and "incompetent" (2003, p. 1168). Moreover, although Scheff (1988) recognized that shame can be linked to violence through shame-rage spirals, he cogently stated that shame has two roles: (a) shame can be constructive when it is *acknowledged*, thus resulting in a social monitoring of the self that maintains our moral behavior and relationships with others; or (b) shame can be destructive or "pathological" when shame is *unacknowledged* or by-passed, giving rise to "a potentially limitless spiral" (1988, p. 400). Even Braithwaite acknowledged that direct forms of shaming have the potential to be stigmatizing if the context is already highly shaming, and that indirect forms of shaming in restorative conferences that allow the wrongdoer to confront their crime themselves, apologize, and attempt to right the wrong done are the most reintegrative (Ahmed et al., 2001).

Three Types of Shame

Relevant to the present study, like Lewis and Tangney, Braithwaite (1989) originally hypothesized shame and guilt to be distinct constructs. Shame is aroused after stigmatizing shaming where the individual feels bad for committing a crime, while guilt is aroused after reintegrative shaming where the individual feels morally responsible for

their bad act. This original assumption has since been revised after a principal component analysis conducted by Harris (2003). With a sample of 720 drunk driving offenders assessed using an interview format, Harris found that shame and guilt were not distinct, but instead composed one component that he labeled "shame-guilt." Through his analyses, two other emotion-related components were also reported that he termed "unresolved shame" and "embarrassment-exposure." Although these three constructs were found to be distinct, Harris did find moderate intercorrelations amongst them. Specifically, statistically significant correlations included the correlation between shame-guilt and embarrassment-exposure (0.60 for conference and 0.52 for court), followed by embarrassment-exposure and unresolved shame (0.34 for conference and 0.15 for court).

In contrast to Braithwaite's (1989) original assumptions regarding shame and guilt, Harris (2003) instead found that the shame-guilt component was comprised of items such as feeling ashamed of oneself over one's behavior, knowing and feeling bad because of doing something wrong, worry over losing respect or being rejected by family members and friends, and worry because of harming another individual. Shame-guilt is associated with acknowledged and defused shame, and has been found to mediate reintegrative approaches and more positive outcomes (i.e., less criminal activity, such as bullying; Ahmed et al., 2001; Harris, 2006). This shame-guilt component was distinct from unresolved shame, as items that loaded on the latter component measured ongoing feelings from the act, such as being continually bothered by thoughts about the act and obsessive thoughts about whether the bad act was indeed wrong. Unresolved shame is commonly experienced following stigmatizing shaming practices, which is associated

with shame displacement. This form of shame is associated with a resistance to accept responsibility for one's actions, while directing negative emotionality to another individual in the form of anger and blame. Thus, unresolved shame has been found to be an important mediator between stigmatizing shaming and more negative outcomes, while hindering the reparative process of the conference (Harris, 2006). Finally, items that loaded on the embarrassment-exposure component consisted of feelings of self-consciousness and concerns about others' evaluations. Although not considered as painful an experience as unresolved shame, embarrassment-exposure is hypothesized to be experienced more frequently following stigmatizing shame as well. This assertion is based upon the conclusion that embarrassment is most commonly experienced in social situations where an individual is the center of attention, making them feel uncomfortable (Harris, 2006).

Central to the current study is that through regression analyses, Harris (2003) found that the shame-guilt component significantly predicted greater empathy responses, while being significantly negatively related to anger and hostility. In contrast, unresolved shame was significantly predictive of greater anger and hostility. These findings are in accord with the aforementioned statements on acknowledged and unacknowledged (by-passed) shame proposed by Scheff (1988). To reiterate, Scheff stated that it is shame that is unacknowledged (unresolved shame) that leads to outward-directed anger and violence, while shame that is acknowledged (shame-guilt) is adaptive and leads to enhanced social monitoring. The central importance of acknowledged shame is supported by the work of Tomkins (1963) and Lynd (1958). Aptly stated by Lynd, "The

very fact that shame is an isolating experience also means that if one can find ways of sharing and communicating it, this communication can bring about particular closeness with others…" (p. 66). Thus, rather than hiding the experience of shame or projecting anger and blame onto others, shame that is successfully acknowledged within a conference setting can allow a bridge of understanding to be built between victim and offender, thus providing a type of "symbolic reparation" for a wrong that was committed (Retzinger & Scheff, 1996). Empirical support has been found for this assumption regarding the acknowledgment of shame as well (Ahmed & Braithwaite, 2005; Losoncz & Tyson, 2007; Makkai & Braithwaite, 1994; Murphy & Harris, 2007). Braithwaite (Ahmed et al., 2001) revised his hypotheses of shame and guilt to reflect these findings. For the purposes of the current study and to align with Braithwaite, the aforementioned conceptualization of shame and guilt composing one construct was employed.

Link between Shame and Self-Esteem

Braithwaite's (Ahmed et al., 2001) conceptualization of shame adopted for the current study suggests that shame not only involves negative feelings aroused because of a violation of one's own or important others' standards, but is inherently linked to our social relationships. The real or perceived opinions others in our social group have of us are inherently linked to the arousal of shame, as shame is elicited when we believe others disapprove of us or our actions. This "social threat" conception of shame is intimately linked to an individual's level of self-esteem, and that self-esteem will impact not only how an individual reacts to a shaming experience, but also plays a vital role in how an

individual either acknowledges or fails to acknowledge their shame (Gailliot & Baumeister, 2006; Sommer & Baumeister, 2002).

The inherently social nature of the shame experience echoes early works by Cooley (1922), who contended that social monitoring of the self is continuous: we are always "living in the minds of others without knowing it" (p. 208). Cooley's "looking glass self" maintains that we see ourselves through the eyes of others, and this social monitoring always has an evaluative component. Scheff (1988) reiterated this view through his deference-emotion system, where he asserted that shame is aroused by real or negative evaluations of the self and that "shame is *the* social emotion, arising as it does out of the monitoring of one's own actions by viewing one's self from the standpoint of others" (p. 398; italics in original). Similarly, from an evolutionary view, Gilbert and McGuire (1998) contend that the experience of shame serves as a warning signal that our actions may reduce our social "attractiveness" to the group.

Coupled with the idea of the looking glass self and that shame involves real or perceived rejection by the social group, Scheff (1990) extended this argument and stated that these constant, low-level states of shame and pride partially make up how we feel about ourselves. These self-feelings, or self-esteem, could thus be considered "a summary measure, representing the balance between pride and shame states in a person's life, taking into account not only duration but also intensity" (p. 284). Thus, our experiences of pride and shame highly impact our self feelings. This socially-oriented conceptualization of self-esteem is also evident throughout the works of Rosenberg (1979), who created one of the most widely used measures of trait self-esteem.

Rosenberg (1979) defined self-esteem as "the totality of the individual's thoughts and feelings having reference to himself as an object" (p. 7), and stated that not only does our self-attitude give rise to pride and shame specifically, but also contended that we cannot know ourselves *except* through the eyes of others. Much like Scheff's idea of self-esteem being a summary measure of pride and shame, Leary and Down's (1995) viewed self-esteem as a "sociometer" that functions to maintain interpersonal relationships. Self-esteem as a sociometer "(a) monitors the social environment for cues indicating disapproval, rejection, or exclusion; and (b) alerts the individual via negative affective reactions when such cues are detected" (Leary & Downs, 1995, p. 130). This measure of our "relational value" involves not only our immediate responses to an event (i.e., state self-esteem), but also reflects our internalized feelings of ourselves (i.e., trait self-esteem). Supporting the sociometer view of self-esteem, empirical work has demonstrated that a sense of belongingness explained a significant amount of variance in participants' ratings of self-esteem (Gailliot & Baumeister, 2006). Our inclusionary status can impact our self-views.

When reviewing the nature of self-esteem, it is clear that the social view of self-esteem can be easily integrated into Braithwaite's model of shame and shaming practices (Ahmed et al., 2001; Braithwaite, 1989). According to Braithwaite (1989), the type of shaming that is imposed on an offender (reintegrative or stigmatizing) will differentially impact how the offender views themselves and their likelihood to commit crime. For example, if after being arrested for a drunk driving offense the offender is shamed by being treated with intolerance, disrespect, and is labeled as a deviant "criminal" by those

in their social network, the offender's self-views and self-esteem are likely to be negatively impacted. This is because individuals see themselves through the eyes of others, and via the sociometer, the individual will experience shame because their relational value has been lowered due to how others have reacted to their offense.

According to Leary (Leary & Downs, 1995) and Scheff (1990), these stigmatizing behaviors have the possibility of impacting the state or trait level of self-esteem of the individual, depending on how often and pervasive these shaming behaviors are practiced. Moreover, according to Leary (Leary & Downs, 1995), individuals with low levels of trait self-esteem are likely to be hypersensitive to disapproval and interpret many reactions by significant others, whether reintegrative or stigmatizing, as negative and connoting possible rejection. Thus, an individual's level of self-esteem will influence how they perceive the shaming behaviors of others, and those with low self-esteem are likely to interpret any shaming as stigmatizing. Research by Sommer and Baumeister (2002) support this hypothesis. Specifically, social rejection differentially impacted individuals with low versus high self esteem. In a series of three studies, Sommer and Baumeister found that individuals with varying levels of self-esteem responded differently to implicit forms of rejection: individuals with low self-esteem were more likely to hopelessly give up, withdraw, and be unlikely to persist in a difficult task following rejection in comparison to those with high self-esteem.

Influence of Self-Esteem on Shame Emotions

Continuing the synthesis of this relational view of self-esteem with the ideas of Braithwaite (1989), not only do the shaming behaviors practiced impact the individual's

state and trait level of self-esteem, but high or low levels of self-esteem can potentially impact the type of shame the offender experiences. In an empirical investigation of Braithwaite's claims (1989; Ahmed et al., 2001), Harris (2003) found that the type of shaming employed, overtly or covertly, directly impacted whether the individual reacted with shame-guilt (shame acknowledgement) or unresolved shame (shame displacement). To integrate this logic with the argument above, a drunk driving offender who has low self-esteem and is shamed (by reintegrative or stigmatizing methods) is likely to experience unresolved shame, due to the offender's rejection sensitivity. When experiencing this form of shame, the individual is likely to resist accepting responsibility for their actions and direct their negative emotionality outwards, which in turn makes the reparative process of the conference more difficult to achieve. As reported by Harris (2003), this experience of unresolved shame is predictive of responding with greater anger and hostility. In contrast, an individual with high self-esteem is likely to respond with unresolved shame when shamed only through stigmatizing disapproval. Leary (2005) has contended that those with high self-esteem will be less sensitive to rejection cues than those with low self-esteem. These individuals can adaptively cope in the face of potential rejection, as supported by the empirical work by Sommer and Baumeister (2002). Further, as Rosenberg stated, "A person with high self-esteem is fundamentally satisfied with the type of person he is, yet he may acknowledge his faults while hoping to overcome them" (1979, p. 31).

Not only has self-esteem been found to be inherently linked to shame, but empirical investigations have also found an individual's level of self-esteem to be

correlated with different types of negative outcomes. These findings are in line with Harris's (2003) conclusions mentioned above. For example, researchers have demonstrated that low self-esteem is correlated with fewer economic prospects, poor physical and mental health, and even criminal offending (Cox, 1996; Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005; Leary, Schreindorfer, & Haupt, 2004; McGee & Williams, 2000; Oser, 2006; Trzesnieski et al., 2006). In a review of the role of low self-esteem and behavioral problems, Leary and colleagues (2003) cited several, empirically based studies of low self-esteem being implicated in a range of dysphoric emotional states and behavioral problems. For example, people with low self-esteem were more likely to experience social anxiety, jealousy, anger, depression, feelings of isolation, loneliness, and insecurity, while being more prone to engage in substance abuse, risky sexual behavior, and aggression. Thus, it appears that an individual's level of self-esteem can impact (a) how they respond to shaming and, in turn, acknowledge their shame, and (b) their likelihood of participating in criminal activity.

In contrast to the view that low-self esteem can lead to a host of negative outcomes, some researchers instead have demonstrated that exceptionally high levels of self-esteem (i.e., narcissism) are related to an array of negative outcomes. Some of these outcomes include increased aggression, the implementation of self-protective strategies and a hindrance of self-regulation abilities, and prejudice against out-group members (Baumeister, Heatherton, & Tice, 1993; Baumeister, Smart, & Boden, 1996). This assumption leads to the conclusion that there is actually a "dark side" of high self-esteem (Baumesiter et al., 1996). Finally, further controversy within the self-esteem realm has

arisen with the hypothesis that self-esteem only plays a negligible role (positive *or* negative) in predicting future adjustment (Baumeister, Campbell, Krueger, & Vohs, 2003; Blackhart, Nelson, Knowles, & Baumeister, 2009; Bushman & Baumeister, 1998). In the most recent review to date on self-esteem, Baumeister and colleagues (2003) found little support for the hypotheses that self-esteem is associated with academic performance, drug use and smoking, violence, or antisocial behavior. In fact, the most desirable correlates of high self-esteem were an increased level of subjective happiness and increased initiative. Clearly, although multitudes of studies have been conducted to determine the role that high and low self-esteem plays in predicting life outcomes, no consensus currently exists regarding the impact of self-esteem on human functioning in a variety of realms, meriting further empirical attention to be paid to the impact of self-esteem on criminal behavior.

Empirical Investigations of Reintegrative Shaming Theory

Braithwaite's (1989; Ahmed et al., 2001) theory of reintegrative shaming can be considered a specific form of restorative justice where the goal is not simply to inflict punishment on an offender for a criminal act. Instead, reintegrative shaming practices are used to repair and restore the relationship between the offender and victim that was harmed by the offense. Although retributive justice (i.e., punishment) is the most common response to criminal offending in the United States, recent meta-analyses (Bradshaw & Roseborough, 2005; Latimer et al., 2005; Nugent et al., 2004) conducted on restorative justice have found that restorative justice in the form of victim-offender

mediation and peacemaking circles can be more effective alternatives to retributive justice in reducing recidivism. Although these findings are promising, two problems within this literature are apparent: (a) reintegrative shaming conferences were not included in any meta-analyses of restorative justice, and (b) when RST was investigated, the underlying theory specifying the relationships between many of the core RST constructs was not empirically assessed. Thus, in the following section, a critical review of the empirical investigations of the empirical RST literature will be presented.

In conducting a literature review on RST, the present author identified 18 investigations of RST, but only 11 of these 18 articles were actual empirical tests of the theory. The other seven investigations indirectly tested the theory or interpreted and/or re-interpreted findings through the lens of RST, without actually conducting a direct test of the efficacy of the theory (Baumer, Wright, Kristinsdottir, and Gunnlaugsson, 2002; Houts, 1996; Lu, 1999; McAlinden, 2005; Miethe, Lu, & Reese, 2000; Vagg, 1998; Zhang, 1995). Thus, at first glance it appears that the theory has been subjected to a greater number of empirical investigations (in a variety of different contexts or cultures) than it actually has, which could lead other researchers or practitioners of RST to draw unwarranted conclusions regarding the efficacy of RST.

Seven Indirect Tests of RST

As examples of indirect testing of RST, four early attempts at generalizing the theory focused on the efficacy of certain tenets of RST in different cultures (Baumer et al., 2002; Lu, 1999; Vagg, 1998; Zhang, 1995). To elaborate, Vagg used self-report data on juvenile offenders in Hong Kong to assess Braithwaite's (1989) hypothesis that

reintegrative shaming practices are more likely to occur in interdependent societies, as exemplified by Hong Kong. Vagg found that youth who commit crimes in Hong Kong are quickly labeled as deviants by those individuals outside of the family and are sanctioned heavily, even for first time offenses. These findings lead Vagg to question Braithwaite's assumptions that reintegrative, in contrast to stigmatizing shaming, is practiced in interdependent communities. The problem inherent in this conclusion is that this study began with the assumption that Hong Kong is an interdependent and/or communitarian society, while providing no test for the validity of this claim. Rather, previously collected data and police records were assessed in relation to the assumption about the interdependency of the culture.

The same basic technique was used by Baumer and colleagues (2002) in studying the Icelandic population, where these authors wished to assess Braithwaite's hypothesis that communitarianism (i.e., collectivism) is associated with lower crimes rates. Similar to Vagg (1998), the authors did not attempt to measure the actual level of communitarianism in Iceland, nor did they actually measure if reintegrative techniques were used with offenders. Instead, they just assumed that Iceland is a communitarian society, thus they must use reintegrative practices. The authors then compared recidivism rates in Iceland to a variety of other nations and determined that because Iceland's recidivism rates are similar to those in other noncommunitarian societies, reintegration does not actually lead to less recidivism. The methods and conclusions drawn by Baumer and colleagues (2002) suffer from the same limitations as Vagg (1998), thus great caution is urged if attempting to assess the efficacy of RST from these two studies.

Similar to Vagg (1998), Zhang (1995) attempted to assess Braithwaite's claims that greater interdependency leads to higher rates of shaming practices by comparing Asian and African American families in Los Angeles, California. Zhang found that the Asian families were not significantly more interdependent than the African American families, nor did they differ on their rates of shaming. However, because the basic premise that the two groups should differ on their level of *interdependency* was not supported by the data, the test of the differences in the *rates of shaming* practiced is not meaningful. Finally, Lu (1999) assessed a form of reintegrative shaming termed *bang jiao* in Shanghai, China and concluded that criminal recidivism rates were lower in the two *bang jiao* practicing neighborhoods that were studied. Although this author offered support for RST in the form of *bang jiao*, actual data on recidivism rates were not provided, nor were recidivism rates for neighborhoods that do not practice *bang jiao* provided or even mentioned. Thus, the internal validity of Lu's claims is weakened as a result of not providing evidence of the recidivism rates of a comparison group.

The remaining three indirect attempts to test or apply RST were conducted with different classes of offenses (Houts, 1996; McAlinden, 2005; Miethe et al., 2000). For example, Miethe and colleagues (2000) compared recidivism rates for those individuals that attended drug court with a matched sample of nondrug court individuals from a public data source of offenders in Nevada. These authors conducted their study under the assumption that drug court has many of the same properties as RST; thus, those that are prosecuted through drug court should have lower rates of recidivism. Contrary to this hypothesis, those individuals prosecuted through drug court had higher rates of

recidivism. The authors attempted to explain this finding by stating that their "field observations" actually demonstrated that stigmatizing forms of shaming were practiced in drug court (rather than reintegrative shaming as they assumed). Thus, they concluded that their study actually does support some of the assumptions of RST (i.e., the stigmatizing shaming that actually was practiced in drug court lead to higher rates of crime). The validity of their conclusions is questionable for a variety of reasons: offenders were not randomly assigned to treatment or control groups, the authors do not present any data besides anecdotal evidence of their field observations, and the authors do not report data on field observations from the nondrug court hearings. Thus, it is difficult to even determine what type of shaming practices actually occurred in the two conditions, which is necessary in order to make any further conclusions based on comparisons between the two groups.

Also working within the context of drug-related offenses, Houts (1996) attempted to explain the efficacy of certain facets of a 12-step alcoholism program through the lens of RST, while not actually testing the theory. Specifically, Houts interviewed 59 members of 12-step programs to determine the extent to which the factors of interdependency and communitarianism were present in these 12-step programs. Houts found evidence for the existence of interdependent relationships, communitarianism, and reintegrative shaming practices within the 12-step paradigm, and thus concluded that these "nontreatment factors" could be plausible partial explanations for why 12-step programs are successful. Thus, this study did not propose to test the theory, but instead attempted to use RST as a theory to explain success in 12-step programs. Similarly,

McAlinden (2005) did not conduct a test of RST, but instead argued that the facet of community involvement inherent in RST, as well as the notions of offender responsibility and reintegrative shaming practices, make RST a particularly viable candidate for effectively curtailing recidivism with sexual offenders.

Although these seven "investigations" of RST did not provide direct tests of the theory, these studies are important for two reasons. Firstly, these studies demonstrate that the consideration of reintegrative shaming as a form of restorative justice is becoming more widely accepted as a plausible alternative to traditional court processes to curtail crime, even in the United States. Secondly, and more importantly, the use of RST to interpret findings in a variety of contexts suggests that researchers and practitioners already believe in the efficacy of certain tenets of the theory, although they have not been rigorously and empirically investigated. Thus, further direct testing of the theory should be conducted in order to determine if the theory can and should be implemented in such a wide range of settings.

Eleven Direct Tests of RST

In contrast to the weak conclusions that can be drawn from the investigations discussed above, firmer conclusions regarding the effectiveness of RST in relation to a reduction in criminal offending can be drawn from the 11 articles identified as direct empirical tests of the theory (Ahmed & Braithwaite, 2004, 2005; Botchkovar & Tittle, 2005; Hay, 2001; Losoncz & Tyson, 2007; Makkai & Braithwaite, 1994; Murphy & Harris, 2007; Tittle, Bratton, & Gertz, 2003; Ttofi & Farrington, 2008; Tyler et al., 2007; Zhang & Zhang, 2004). Thus, in this section, the conclusions drawn regarding RST from

each of these studies will be presented. However, caution is urged when assessing the conclusions from these investigations because of (a) the limited number of studies conducted on RST, and (b) methodological weaknesses inherent in each of the studies that limit the claims of causality for each study (Shadish, Cook, & Campbell, 2002).

When attempting to integrate and summarize the findings from the 11 empirical investigations of RST, it becomes apparent that the diversity of these studies with regard to the offenses that are studied, the methodologies used to study RST, the components of RST that are tested, and the populations in which RST is investigated is so large that no simple summarization would suffice. Thus, the first step in attempting to draw broad conclusions regarding the efficacy of RST as exhibited by these studies is to roughly categorize the investigations as providing "strong" or "partial" support for the tenets of RST. Strong support was defined as any study that found empirical support for all or all but one of the tenets that were tested. Partial support was defined as any study where the majority of the tested tenets were supported, yet two or more were not supported. Based on this rather coarse categorization scheme, it appears that the support found for RST is promising, but inconclusive. Specifically, six of the studies provided strong support for the components of RST studied (i.e., Ahmed & Braithwaite, 2004, 2005; Losoncz & Tyson, 2007; Makkai & Braithwaite, 1994; Murphy & Harris, 2007; Ttofi & Farrington, 2008) and five studies were categorized as providing partial support (Botchkovar & Tittle, 2005; Hay, 2001; Tittle et al., 2003; Tyler et al., 2007; Zhang & Zhang, 2004). In order to draw more fine-grained conclusions regarding the efficacy of RST as evidenced

by these 11 studies, the investigations will be grouped according to offense type and discussed in greater detail below.

White collar crime. In the earliest empirical investigation of RST (that was conducted with white collar criminals) Makkai and Braithwaite (1994) investigated if there was a differential increase in compliance with regulatory standards in nursing homes if a high or low level of "disapproval" (i.e., shaming) was paired with high or low levels of reintegration. These authors found strong support for the reintegration construct of RST. Those nursing homes that were inspected by individuals that demonstrated high levels of disapproval and high levels of reintegration reported the greatest change in increased compliance with regulatory standards. Further, this effect was only evidenced in nursing homes where there was interdependency between the director of the nursing home and the inspector, thus also supporting Braithwaite's hypothesis regarding the importance of interdependent relationships in leading to reduced offending (Makkai & Braithwaite, 1994).

In the only other investigation of the efficacy of RST with white collar crime, Murphy and Harris (2007) similarly found strong support for RST. Murphy and Harris specifically investigated if being reintegratively, in contrast to being stigmatically shamed, was associated with less reported evasion of taxes two years later. These authors also tested if shame-related emotions mediated this relationship between shaming and criminal re-offending. In support of Braithwaite's theory, higher perceived levels of reintegrative forms of shaming (which also indicates low levels of stigmatizing shaming) were significantly associated with lower reported recidivism in the form of tax evasion.

Further, the relationship between forms of shaming and recidivism was partially mediated by shame acknowledgement or displacement. Thus, the core assumptions of the impact of reintegrative versus stigmatizing shaming practices were strongly supported by their data.

Minor criminal offenses. In addition to the two studies on RST with white-collar crime, limited research on the efficacy of RST has also been conducted in the area of minor criminal behavior such as drug use and theft. The support for RST in this category of criminal offenses is equivocal, but warrants further investigation of RST in this context (Botchkovar & Tittle, 2005; Tittle et al., 2003; Tyler et al., 2007). For example, Tittle (Botchkovar & Tittle, 2005; Tittle et al., 2003) conducted two studies with various colleagues to assess three major hypotheses of RST: (a) reintegrative shaming is negatively associated with future misbehavior, (b) stigmatizing shaming is positively associated with future misbehavior, and (c) the magnitude of the above associations would be greater if participants were more socially integrated. In both studies, participants (located in North Carolina and Russia) were asked to project their future likelihood of committing four types of crimes (i.e., violence, large theft, minor theft, and drug use) using interviewing techniques. These authors only found support for the second hypothesis listed above in both studies. Specifically, they found that reported stigmatization lead to projections of future crime, reported reintegration was either not statistically associated with projections of future crime or was associated with increased (rather than decreased) projections of misbehavior, and that the relationships investigated were not moderated by the social integration construct. The authors did note that these

conclusions should be viewed with caution because of a variety of weaknesses inherent in the study design, including "incomplete measurement, absence of a full array of contingent variables, misreporting, and other problems inherent to surveys" that could have lead to them not finding more favorable results (Tittle et al., 2003, p. 612).

In a study also in the realm of minor criminal offenses, specifically drunk driving offenses, Tyler and colleagues (2007) sought to test the hypothesis that restorative conferences would lead to greater feelings of being reintegrated and a greater sense of being treated fairly by law enforcement. These feelings of reintegration would then lead to greater future compliance with the law one and two years later. Unlike all other investigations previously discussed, the Tyler and colleagues study analyzed data from the only randomized controlled trial of RST to date. The broader study in which these data were collected is referred to as the Reintegrative Shaming Experiments (RISE; Sherman et al., 2000), where offenders committing multiple offenses (e.g., drunk driving, minor theft) were randomly assigned to take part in either a restorative conference using RST or a traditional adjudicative procedure. The Tyler and colleagues study is the only peer-reviewed publication investigating a reduction in recidivism due to RST from this overall randomized-controlled trial. For this investigation, they analyzed data from 900 drunk driving offenders that were arrested and assigned to conference or court.

When investigating the impact of RST on compliance in this analysis, Tyler and colleagues (2007) failed to find a significant difference between the RST conferences and traditional court processing on police-recorded or self-reported recidivism rates. On the other hand, those offenders randomly assigned to a RST conference did report greater

efforts to not drive while drunk and viewed the law as more legitimate and fair, in comparison to the control group. Further, they found that those that viewed the law as more legitimate were less likely to recidivate and that the items assessing reintegration were significant predictors of this "legitimacy" construct when using multiple regression analyses. Thus, the authors concluded that the constructs of reintegration and perceived fairness had a significant impact on lowering rates of future offending. This conclusion was consistent with Braithwaite's theory, because those offenders that reported being treated reintegratively and fairly were less likely to re-offend. Although there were null findings for the treatment effect on recidivism, the authors concluded that this could have been due to the weakness of the treatment across all offenders assigned to treatment, rather than a failure of the theory (Tyler et al., 2007).

Adolescent delinquency. Along with the efficacy of RST being evaluated in the context of adult offenders (i.e., as was done with the white collar and minor criminal offenses), RST has also been investigated in the context of adolescent delinquency (Hay, 2001; Losoncz & Tyson, 2007; Zhang & Zhang, 2004). In one of the earliest tests of RST, Hay investigated the importance of shaming paired with high and low levels of reintegration (i.e., where Hay operationalized low levels of reintegration as stigmatization) on delinquency. He also assessed if higher levels of interdependency between parents and adolescents lead to more reintegrative shaming practices. Hay found strong support for the hypothesized negative relationship between shaming and delinquency, but his findings were not entirely supportive of RST: adolescent delinquency was found to be lower when shaming was used, but delinquency was lower

at *both* high and low levels of reintegration. Thus, the effect of shaming was significant, but the interaction between shaming and reintegration/stigmatization was not found. This study also supported Braithwaite's claims regarding the importance of interdependency, as parents that were more interdependent with their children were more likely to use reintegrative shaming techniques.

Losoncz and Tyson (2007) conducted a similar study with the same population in order to investigate the influence of reintegration, stigmatization, and child-parent interdependency on projected rates of delinquency. The results of this investigation provided very strong support for all three tenets of the theory that were tested: (a) reintegration was significantly negatively related to delinquency, (b) stigmatization was significantly positively related to delinquency, and (c) child-parent interdependency was indirectly related to delinquency through reintegration and stigmatization. Thus, much like Hay (2001), those parents that were close with their children were more likely to use reintegrative shaming and less likely to use stigmatizing shaming, which lead to lower rates of projected delinquency.

Finally, Zhang and Zhang (2004) also sought to test the efficacy of shaming, reintegration, and the interaction between these two constructs (i.e., reintegrative shaming) on predatory delinquency in youth, while controlling for the influence of the child's attachment to his/her parents and peers. Specifically, these authors investigated both peer and parent shaming and reintegration on delinquency in two waves of data collection. Although Zhang and Zhang (2004) found partial support for the main effects of shaming and reintegration on offending, the key test of this study was of the

interaction between these two constructs (i.e., forming the reintegrative shaming construct as hypothesized by Braithwaite) in association with offending behavior. The interaction for both parents and peers was not statistically significantly related to offending behavior. Thus, this study did not support one of the key arguments of Braithwaite's theory.

At this point, it is essential to emphasize that Makkai and Braithwaite (1994), Hay (2001), and Zhang and Zhang (2004) all investigated the efficacy of RST by assessing the significance of an interaction between shaming and high and low levels of reintegration. In the revision of the theory in 2001, Braithwaite and colleagues (Ahmed et al., 2001) predicted that there may be important

> shaming, reintegration, and stigmatization main effects but no interaction effects in contexts heavily laden with shame and no main effects but interaction effects for these variables in contexts where limited shame is normally experienced. (Ahmed et al., 2001, p. 43)

Thus, it may be unwise to assume that the lack of support found for these tests of interactions between shaming and reintegration are indicative of a failure of RST to predict offending behavior, if the contexts in which the interactions were tested were already inherently shaming.

Bullying. Finally, RST has been investigated in the context of school bullying behavior in three studies (Ahmed & Braithwaite, 2004, 2005; Ttofi & Farrington, 2008). Ahmed and colleagues (Ahmed & Braithwaite, 2004, 2005) conducted two similar studies to assess the relationship between self-initiated bullying behavior and shaming practices. In their first investigation (Ahmed & Braithwaite, 2004), the authors used path analysis to assess the importance of various RST-related predictors of bullying (e.g.,

family disharmony, stigmatizing and nonstigmatizing shaming, self-esteem, shame acknowledgment, and shame displacement), along with other potential predictors (e.g., liking for school, impulsivity). They found that shame acknowledgment (i.e., similar to shame-guilt) was significantly negatively predictive of bullying, while shame displacement (i.e., similar to unresolved shame) was significantly positively predictive of bullying, in accordance with RST. Also in agreement with RST, higher levels of family disharmony and stigmatizing shaming were positively predictive of bullying, while the opposite pattern was found for nonstigmatizing shaming practices. Finally, self-esteem played a mediating role through shame-acknowledgement: individuals with high self-esteem were less likely to acknowledge their shame, leading to increased bullying behavior.

In a similar investigation of school bullying in Bangladesh, Ahmed and Braithwaite (2005) investigated the importance of reintegrative and stigmatizing shaming, forgiveness, and shame management. Similar to Ahmed and Braithwaite (2004), shame acknowledgment was predictive of less bullying, while shame displacement was predictive of increased bullying. Moreover, the authors found that reintegrative shaming and parental forgiveness of a wrongdoing were significantly predictive of less bullying, although stigmatizing shaming was not significantly associated with bullying behavior. Thus, both of these investigations were highly supportive of the majority of RST hypotheses that were assessed.

In a third study of RST in the context of school bullying, Ttofi and Farrington (2008) investigated the impact of parental bonding (i.e., interdependence), reintegrative

and stigmatizing shaming practices, and shame acknowledgment/displacement on bullying. In line with RST, the authors found that stigmatizing shaming was significantly negatively predictive of shame acknowledgment. The opposite pattern was observed for reintegrative shaming, although that effect was not statistically significant. Further, the effect of the shaming variables on bullying behavior was mediated by the manner in which the child managed their shame (i.e., acknowledged vs. displaced shame). In terms of the importance of interdependency on reducing offending behavior, increased levels of mother bonding, but not father bonding, was significantly predictive of less bullying. Thus, much like Ahmed and Braithwaite (2004, 2005), the results provided strong support for most of the tested key theoretical assumptions of RST.

Summary of empirical support for RST. Upon reviewing the 11 direct tests of components of RST, it is tempting to draw overall conclusions regarding the efficacy of this type of restorative justice. For example, from the studies previously conducted, it appears that reintegrative shaming practices generally resulted in less offending, while stigmatizing shaming practices lead to further offending. Likewise, interdependent personal relationships reduced the likelihood of criminal behavior, and the type of shame that was experienced (shame-guilt or unresolved shame) was differentially related to offending behavior. Finally, it appeared that RST was more successful in reducing offending behavior when the offense was classified as a white collar crime or juvenile delinquency/bullying, in comparison to minor criminal offenses such as stealing, drunk driving, and violent acts.

Although at first glance these claims seem warranted, two issues arise that limit the ability to draw these conclusions. First, an extremely limited number of studies have been conducted on RST to allow for broad generalizations to be made. Moreover, an even smaller number of studies have been conducted within each offense category. Thus, in order to determine if RST is perhaps more effective for a certain class of offense or offender type, further testing of RST must be conducted in a variety of contexts.

Second, the large degree of heterogeneity of study characteristics within the empirical RST literature makes it even more difficult to make overall generalizations of the efficacy of RST on reducing criminal behavior. Shadish and colleagues (2002) provide an extensive discussion of external validity claims (i.e., the ability to draw valid inferences about causal relationships when there is variation in study characteristics). They further specify that generalizations are made on four study characteristics (i.e., units, outcomes, treatments, and settings). Unless effects are robust, extensive heterogeneity between studies on certain study characteristics can make it difficult to generalize study findings to the overall relationships between constructs that the empirical studies represent. This problem is present within the empirical literature on RST; not only are various types of units, outcomes and settings present within each study, but this problem is compounded by the extremely small number of studies conducted to assess the efficacy of RST. To elaborate, RST has been studied with white collar criminals, school children, and the general adult population (all "units") in a variety of different cities and countries. Similarly, RST has been investigated within the context of compliance with nursing home standards, tax evasion, minor criminal offenses, juvenile

delinquency, and bullying (all "outcomes"), and in a variety of locations such as nursing homes, schools, homes, and public places (all "settings"). Lastly, even the implementation of the "treatments" differed between studies. Thus, it is difficult and premature to firmly draw overall conclusions regarding the efficacy of RST with the limited number of studies that have been conducted, although the available evidence is supportive of the theory. This finding leads to the conclusion that further systematic research on RST is necessary.

Study design weaknesses. Along with these external validity issues, factors that influence the ability of researchers to make valid inferences regarding the causal impact of RST on offending behavior (i.e., internal validity) also need to be addressed. Within the empirical RST literature, two notable factors exist that could plausibly negatively impact the internal validity of these studies: (a) the overabundance of correlational designs being used, and (b) within these correlational studies, the differential manner in which respondents report their offending behavior.

As was previously noted, the Tyler and colleagues' (2007) study using the RISE data is currently the only experimental investigation of RST. In contrast, the other 10 empirical investigations make use of correlational methodology. Specifically, multiple regression, path analysis, and SEM were used to assess the impact of different facets of RST in reducing offending behavior. Although methods such as path analysis and SEM can be used to approximate causality between variables, the correlational data used in the aforementioned 10 studies does not provide an appropriate test of reintegrative shaming theory for two reasons. Firstly, participants did not actually attend a restorative

conference, thus there is no evidence that they have actually been exposed to reintegrative shaming. Secondly, no investigation was conducted to compare the differential impact of reintegrative shaming or traditional court processing on re-offense rates or other similar outcomes. Thus, although constructs such as shaming and shame acknowledgment are assessed in relation to offending behavior, it is difficult to determine the manner in which these constructs are related. In contrast, in the Tyler and colleagues' (2007) study, participants were randomly assigned to receive reintegrative shaming or traditional court processing, thus strengthening their internal validity claims. Although the quality and worth of research findings are not dictated by the methodology chosen, it is clear that causal claims are strengthened by experimental research, thus making it clear that future RST studies should make use of experimental or even quasiexperimental designs to assess the efficacy of these constructs (Shadish et al., 2002).

Not only is the ability to draw causal conclusions regarding the efficacy of RST limited because of the dominance of correlational designs, but this problem is potentially exacerbated by the differential manner in which offending behavior is assessed. Three different techniques have been used within the 11 empirical investigations to assess delinquent or criminal offending within the RST literature: (a) actual records of offense, (b) self-reported frequency of past offending, and (c) self-reported projections of future offending. Making use of actual records of offending behavior when captured accurately and completely is superior to the other two forms of measuring the dependent variable when assessing RST, as will be explained below. Unfortunately, the use of this technique is not always possible. Only two empirical investigations of RST assessed recidivism

rates via actual recorded behavior (i.e., Makkai & Braithwaite, 1994; Tyler et al., 2007). On the other hand, five RST studies relied solely on a self-report measure of past offending behavior (i.e., Ahmed & Braithwaite, 2004, 2005; Murphy & Harris, 2007; Ttofi & Farrington, 2008; Zhang & Zhang, 2004). In each of these studies, participants were asked to indicate how often they committed the key offending behavior being assessed. Finally, four studies (i.e., Botchkovar & Tittle, 2005; Hay, 2001; Losoncz & Tyson, 2007; Tittle et al., 2003) required participants to project the likelihood that they would commit various offenses in the future.

The use of the latter two methods poses various potential limitations in terms of the accuracy of self-reporting of criminal behavior that could impact the internal validity of study findings. Various researchers have criticized the validity of findings when self-report measures are used in the context of criminal offending (e.g., see Gendreau, Irvine, & Knight, 1973; Hare, 1985; Schretlen & Arkowitz, 1990). More specifically, when turning attention to self-reported past behavior, the first issue that arises is a causal-order problem (Hay, 2001). If the independent variable is measured concurrently with the dependent variable or current attitudes about offending behavior, a possible confound arises. Specifically, offenders' current attitudes can influence their reporting of past behavior. A second validity issue specific to reporting of past behavior is summarized by Osberg and Shrauger (1990). These researchers state that when looking at self-reporting of behavior, people in general have difficulty with accurately knowing the real contributors to their behavior, thus the ability to report this information accurately is compromised (see also, Nisbett & Wilson, 1977). Similar validity issues arise when

researchers ask respondents to predict the likelihood of future misconduct, as is also frequently done in the RST literature. Osberg and Shrauger (1990) also addressed this issue and contended that people have a reduced capacity to accurately predict their future behavior because of inherent biases in inference processes about our own behavior (see also, Gilbert, 2006).

A final criticism to both self-reporting of past behavior and predicting future behavior concerns socially desirable responding (SDR; Paulhus, 1984). According to Paulhus, when responding to self-report measures, people have a general tendency to respond in a manner that presents them in the most favorable light. When applying SDR to the self-reporting of criminal behavior, we can easily see that when respondents are asked to report their past criminal history or the likelihood that they will commit various criminal offenses in the future, individuals may not report this information accurately and truthfully. This misreporting is because they (either consciously or subconsciously) want to depict their behavior more or less favorably. This potential problem not only applies to the outcome variable in the empirical studies of RST, but could also apply to methodology used that entails asking parents to self-report the type of shaming they use with their children. Potentially, parents that want to make themselves appear most favorable are less likely to report using stigmatizing versus reintegrative shaming. Thus, because SDR and the other self-report criticisms mentioned above can threaten the validity of study findings, it is clear that researchers should attempt to incorporate actual recorded rates of criminal offending, rather than predominately relying upon self-report measures of both predictor and outcome variables when studying RST.

Central RST tenets not empirically investigated. Although there is much diversity in many study characteristics within the RST literature (e.g., type of offense being investigated, type of method used to assess study outcomes), not all of the central RST tenets have been studied. Specifically, the study of certain constructs such as shaming, shame, shame acknowledgement, and interdependency have dominated the literature, with virtually no research being conducted on other key RST constructs.

First, the impact of stigmatizing versus reintegrative shaming practices on the acceptance of responsibility by the offender and perceived family support following an offense have not been investigated. In his original formulation of the theory, as well as in his later revision, Braithwaite (1989; Ahmed et al., 2001) specifies the importance of these constructs. Braithwaite stated how overt disapproval done reintegratively can be an effective deterrent of crime as it allows the offender to accept responsibility on their own (without pressure from others), leading to the most successful outcome for both victim and offender. Braithwaite contends that "the genius of well-conducted restorative justice processes is that they confront wrongdoing indirectly, implicitly inviting the wrongdoer themselves to be the one who directly confronts it, apologizes, and seeks to right the wrong" (Ahmed et al., 2001, p. 33).

Second, along with offender responsibility, the attention paid to the interdependency/family support construct has been limited. Although 9 of the 11 empirical studies that have been conducted included some form of family support or interdependency variable, none of these studies investigated a family support/ interdependency construct as an outcome measure. As hypothesized by Braithwaite

(1989) crime rates are lower in societies with individuals in more interdependent relationships, and that feelings of interdependency can be influenced by the type of shaming that occurs. In interdependent societies, individuals rely upon each other for physical and emotional support to a greater extent. In an attempt to remain connected to those in their social group, these interdependent individuals react more positively to reintegrative shaming practices. The central importance of this constructs makes it imperative to assess how connected and interdependent an offender feels following a crime, because this perceived family support system is likely to impact their likelihood of committing further offenses. Further, the operationalization of the interdependency construct has varied greatly from study to study (i.e., interdependency was differentially measured by items assessing parental attachment, relying on others for support, or mother bonding). Thus, like the offender responsibility construct, future research should systematically investigate the role of interdependency as an outcome variable within RST.

Third, the impact that individual differences in offenders have reintegrative versus stigmatizing shaming has received little empirical attention. Using SEM techniques, one study (Ahmed & Braithwaite, 2004) included shame and guilt proneness, self-esteem, empathy, impulsiveness, and internal locus of control as individual difference variables in the model. As emphasized above, self-esteem is closely linked to how individuals react to shaming, the type of shame they experience, and thus their likelihood of committing future offenses. Although the Ahmed and Braithwaite (2004) study included a self-esteem construct in their model, they did not investigate how reintegrative versus

stigmatizing shaming practices differentially impact the level of self-esteem of the individual. In fact, their final path model included only a direct path from self-esteem to shame acknowledgment and they did not investigate how self-esteem was impacted by the form of shaming used. But, according to the researchers mentioned above (Leary et al., 2003; Rosenberg, 1979; Scheff, 1990), self-esteem plays a vital role in how we react to criticism and disapproval, the type of shame we experience from that disapproval, and further behavioral problems resultant from that level of self-esteem. Thus, it appears that future research should also direct attention to the mediating impact of self-esteem.

RST Literature Conclusions: Current Study

Upon reviewing the empirical RST literature, several conclusions became apparent that dictated certain methodological decisions to be made and factors to consider in terms of the current study. First, it is clear that although 18 studies were identified within the empirical RST literature, there is a scarcity of actual empirical tests of core RST tenets. Thus, for the present study, the importance of and relationships between RST tenets were investigated using rigorous SEM techniques, rather than "indirect tests" of RST, as discussed above.

Second, the model testing for the current study was conducted using data collected from a randomized-controlled trial of the efficacy of RST. Specifically, a secondary data analysis was conducted using the RISE data (Sherman et al., 2000), as was done in the study by Tyler and colleagues (2007) where the authors investigated the impact of RST on drunk driving recidivism rates. Thus, the data utilitzed for the present study may potentially overcome some of the weaknesses inherent in the 10

correlational studies within the RST literature as a true experiment was conducted. Moreover, SEM techniques will be used with these data to estimate the causal relationships between variables based upon theoretically derived hypotheses.

Third, a secondary data analysis of the Sherman and colleagues data has the potential to overcome many of the weaknesses discussed regarding problems with self-report measures. Specifically, unlike other RST investigations, offenders were not asked to self-report on their past behavior, or to predict their future behavior. Although the current study side-stepped these possible methodological limitations, much of the data collected for the RISE study is self-report in nature, thus it may be afflicted by problems with SDR. Although SDR can potentially be problematic with all self-report measures, it was presumed to be less of a weakness in the current study because offenders were asked to respond to items less susceptible to desirable responding (e.g., the type of shaming they perceived to have experienced, questions regarding perceived family support, self-esteem, and fair treatment), rather than questions implicitly laden with tendencies to respond favorably (e.g., reporting on the likelihood of committing future crimes).

Finally, the current study built upon the breadth of RST constructs studied within the empirical literature by assessing relationships between constructs rarely or never studied. Specifically, offender responsibility, interdependency, and self-esteem were investigated in relation to other important RST constructs (i.e., reintegration, stigmatization, shame-guilt, unresolved shame, embarrassment-exposure, and perceived fairness). Although the researchers for the RISE project collected data on these three constructs, they did not analyze the importance of these constructs or their relationships

to more frequently studied constructs. Thus, SEM was used to assess the relationships between previously studied and unstudied constructs. Using SEM to analyze the structural relationships between constructs allowed the researcher to address the first overall purpose of the current study: to test the theoretical model underlying RST to determine whether Braithwaite's hypotheses regarding the effect of RST on crime-deterrent constructs (e.g., perceptions of reintegration, stigmatization, fairness, and offender responsibility) are empirically supported. Similarly, this statistical technique also enabled the researcher to address the second aim of assessing the effectiveness of these RST constructs when implemented in response to actual criminal behavior. Specifically, drunk driving offenders randomly assigned to RST conferencing or court processing were compared using multigroup SEM.

CHAPTER III

METHOD

Data

The present study is a secondary analysis of data from the *Reintegrative Shaming Experiments (RISE) in Australia, 1995 – 1999* (Sherman et al., 2000), sponsored by the United States Department of Justice: National Institute of Justice. Data were retrieved from the National Archive of Criminal Justice Data at the University of Michigan on December 15, 2008. The purpose of the original study was to compare the effectiveness of standard court processing with restorative justice conferencing using RST on police-recorded and self-reported rates of recidivism 1 and 2 years following the first offense. In the original study, the impact of RST was examined in offenders who committed four different types of offenses: (a) drunk driving at any age with a blood alcohol level over 0.08, (b) juvenile property offending with personal victims, (c) juvenile shoplifting, and (d) violent crimes committed by youth younger than 30.

Participants were recruited by Australian Federal Police in the Australian Capital Territory region. Specifically, when an arresting police officer made an arrest that they believed was eligible for inclusion in the study, they called a 24-hour number that was in permanent custody of a member of the research team. The officer was then asked 10 screening questions to determine if the case was indeed eligible. If the case was deemed eligible, an envelope containing the treatment assignment (i.e., conference or court) was opened and the officer was immediately notified in order to take the appropriate action. Although participants were randomly assigned to conference or court processing,

attending a conference was voluntary while attending court was mandatory. Out of the 900 individuals apprehended and recruited for the study, only seven refused to attend conferencing (Ahmed et al., 2001).

All individuals apprehended were assigned to court processing or restorative justice conferencing using the randomized controlled trial method. The criteria used for randomization included:

> (1) the crime was one of the target offenses, (2) the offender(s) had made a full admission of responsibility for the offense, (3) a sergeant had approved the case being sent to RISE, (4) there was no reason to believe the offender would object to a conference if the case was assigned to a conference, (5) the offender did not have any outstanding warrants or bonds requiring him/her to go to court, (6) the offender lived in the ACT region, and (7) the police officer referring the case agreed to accept the RISE recommendation for all co-offenders in the case regardless of whether it was assigned to court or conference. (Sherman et al., 2000, p. 5)

For the present study, analyses included only individuals arrested for drunk driving as the sample size for the other three offender groups was deemed unacceptable for SEM analyses (e.g., $n < 100$). Nine hundred drunk driving offenders were referred to the original RISE study, but participant loss did occur due to a variety of circumstances. First, 17 cases assigned to court and 23 assigned to conference were abandoned by the police (due to lost records by the police or officers being unable to contact the offender). Second, two cases per treatment condition were also lost because they were already participants in one of the RISE experiments. Third, one individual randomly assigned to court and 21 individuals randomly assigned to conference actually received the incorrect treatment (which most often occurred because the individual was assigned to conference and did not attend, thus they were sentenced to a court hearing). Finally, 104 individuals

assigned to court and 52 individuals assigned to conference were unable to be interviewed (although observational data was collected on these individuals), and were excluded from the present analyses. Despite these limitations, Barnes (1999) found that the two groups did not significantly differ on any major demographic characteristics.

There were additional eligibility criteria for the offenders arrested for drunk driving offenses, including: "(a) the offender's blood alcohol content (BAC) was over .08 at the time of the incident, (b) the offense did not involve an accident, (c) the offender was not a police officer, and (d) the offender was eligible for VATAC (Voluntary Agreement to Attend Court; Sherman et al., 2000, p. 4). Thus, no drunk driving cases included in the RISE dataset involved a direct victim of the crime, as the researchers sought to maintain homogeneity between offenses. Within the drunk driving sample, 374 offenders assigned to reintegrative shaming processing and 350 individuals processed through the standard Australian court system were used in the present analyses.

Within this entire sample, the majority of the offenders were male (75.6%) and nonjuvenile (95.2%), with a mean age of 30 years. Additionally, the majority of the sample had completed at least 12 years of formal schooling (55.6%), was currently employed full time (64.3%), had never been married (57.4%), and was born in Australia (80.2%).

Procedure

Three sources of data were collected for all study participants by trained RISE research staff: (a) observations of conference or court proceedings, (b) self-report items

via semistructured interviews with participants, and (c) traditional paper-and-pencil self-report where participants answered questions privately. Data collected were Likert-type items ranging from 1 to 8 for the observational data and 1 to 4 or 1 to 5 for the interview and self-report data. The reliability of the conference and court observations was assessed through percentage of agreement between trained observers (Harris & Burton, 1998). In a study to assess the reliability of these observations, Harris and Burton concluded that between-observer agreement was acceptable. Interviews were completed with participants 2 to 6 weeks (on average, 5-6 weeks) following the implementation of the treatment (i.e., court or conferencing). Interviews lasted, on average, 87 minutes in comparison to court cases, which only lasted an average of 7 minutes in duration.

Measures

All observational items were compiled into the "Global Observational Ratings Instrument," which was completed at the conclusion of both the court hearings and the conferences. This measure was created to assess the observer's overall impressions of the occurrences at the hearing or conference (e.g., offender behavior, type of shaming used). The items assessed through interviewing and self-report formed the "Act Justice Survey," which was subjected to pilot testing with offenders not included in the RISE study. The Act Justice Survey included five categories of items to assess core reintegrative shaming concepts based on Braithwaite's delineation of the theory: (a) participants' perception of how they were treated throughout the process, (b) participants' perceptions of their reaction to the offense they committed, (c) participants' perceptions

about support, being shamed, and how they felt during the case, (d) demographic questions, and (e) questions pertaining to their self-esteem and emotions. Due to the sensitive nature of the self-esteem and emotion items, the participants responded to this particular subset of questions by answering privately without a trained interviewer administering the questions. For the purposes of the present study, a subset of these items was chosen to measure nine constructs.

An iterative process was used to choose the items from the Act Justice Survey and Global Observational Ratings Instrument. A review of the empirical RST literature was first conducted to determine which latent constructs were important to include in the measurement and structural models for RST. Over 100 items were then initially chosen to represent 14 latent constructs to describe Braithwaite's theory. Once these items were chosen, descriptive statistics were calculated to investigate the distribution and pattern of missingness for each item, which resulted in a reduction in the number of usable items. Exploratory factor analyses and reliability analyses using *SPSS version 17 for Windows* (SPSS Inc., 2008) were then conducted to investigate the factor structure of the hypothesized 14 latent constructs. These analyses resulted in a further reduction in the number of items to be used, as well as a reduction in the number of latent constructs to be included in the model. A total of 40 items were originally chosen to measure nine latent constructs. As will be described in greater detail in the results section, four final items were removed in order to achieve acceptable fit for the baseline measurement model, yielding a final total of 36 items with nine latent constructs. These items are presented below in Table 1, followed by the nine latent constructs measured by these items.

Table 1

Indicator Items Comprising Latent Constructs and Descriptive Statistics

Latent construct	Indicator item	Conference (N = 374) *M*	Conference (N = 374) *SD*	Court (N = 350) *M*	Court (N = 350) *SD*
Reintegration					
RS1	At the end of the conference/court case did people indicate that you were forgiven?	1.05	1.12	0.00	1.85
RS2	Did others at the conference/court case say that you had learned your lesson and now deserve a second chance?[a]	1.86	1.73	-0.02	2.33
RS3	Did you learn from the conference/court that there are people who care about you?[a]	2.50	1.32	0.47	1.89
RS4	During the conference/court case did people suggest that they loved you regardless of what you did?[a]	1.81	2.36	0.07	3.37
RS5	During the conference/court case did people talk about aspects of yourself which they like?[a]	1.48	1.35	-0.57	2.31
Stigmatization					
SS1	Were you treated in the conference/court as though you were likely to commit another offence?[a]	-0.43	1.42	-0.18	1.38
SS2	Did people during he conference/court case make negative judgments about what kind of person you are?[a]	-0.78	1.47	-0.63	1.53
SS3	During the conference/court case were you treated as though you were a criminal?[a]	0.05	1.04	0.37	1.07
SS4	During the conference/court case were you treated as though you were a bad person?[a]	0.07	0.89	0.09	1.02
Self-Esteem					
SEE1	All in all, I am inclined to feel that I am a failure.[b]	2.98	1.07	2.57	0.93
SEE2	I wish I could have more respect for myself.[b]	1.40	0.72	1.32	0.74
SEE3	I certainly feel useless at times.[b]	1.21	0.58	1.06	0.56
SEE4	I give up easily.[b]	3.66	1.52	3.59	1.64
SEE5	At times I think I am no good at all.[b]	1.73	0.75	1.59	0.77

(*table continues*)

Latent construct	Indicator item	Conference (N = 374)		Court (N = 350)	
		M	*SD*	*M*	*SD*
Perceived fairness					
FAR1	You felt you had the opportunity to express your views in the conference/court.[b]	3.36	1.39	1.90	1.16
FAR2	All sides got a fair chance to bring out the facts in the conference/court.[b]	3.96	1.72	2.20	1.37
FAR3	You felt you had enough control over the way things were run in the conference/court.[b]	1.89	0.96	1.11	0.84
FAR4	The conference/court took account of what you said in deciding what should be done.[b]	2.14	1.03	1.51	1.00
FAR5	You feel that you were treated with respect in the conference/court.[b]	2.54	1..45	2.21	1.19
Shame-guilt					
SG1	During the conference/court case I felt that the offence I committed was wrong.[b]	2.59	1.08	1.90	0.92
SG2	In the conference/court case I felt angry with myself for what I had done.[b]	1.62	1.09	1.31	1.01
SG3	During the conference/court case I felt ashamed of what I did.[b]	1.80	1.01	1.15	0.86
SG4	During the conference/court case I felt ashamed of myself.[b]	1.28	1.28	0.81	1.00
SG5	I felt bad in the conference/court because my actions had hurt others.[b]	0.93	1.23	-0.13	1.48
Unresolved shame					
U1	Since the conference/court have you found yourself continually bothered by thoughts that you were unfairly or critically judged by people at the conference/court case?[b]	-1.46	1.57	-0.59	1.58
U2	Do you feel that some of the things brought up in the conference/court case are unresolved in your mind?[b]	-1.11	1.46	-0.78	1.59
Embarrasment-exposure					
E1	During the conference/court case I felt awkward and aware of myself.[b]	1.40	1.08	1.11	0.97
E2	In the conference/court case I felt embarrassed because I was the center of attention.[b]	1.03	1.40	1.06	1.39
E3	During the conference/court case I felt so exposed I wished I could just disappear.[b]	-1.04	1.70	0.11	1.36
E4	In the conference/court I felt uneasy because I was surrounded by people who were supposed to be more important than me.[b]	0.01	1.07	0.31	1.26

(*table continues*)

Latent construct	Indicator item	Conference (N = 374)		Court (N = 350)	
		M	*SD*	*M*	*SD*
Offender responsibility					
RES1	To what extent die the offender accept they had done wrong?[c]	3.67	1.32	4.96	2.54
RES2	How sorry/remorseful was the offender for their actions?[c]	5.42	3.51	2.89	4.25
RES3	How much responsibility did the offender take for their actions?[c]	7.38	2.65	6.13	2.98
Family support					
FAR1	As a result of the conference/court case I feel more proud of being a member of my family.[b]	4.82	3.08	2.45	1.43
FAR2	The conference/court case has brought my family closer together.[b]	2.92	1.81	1.78	1.24
FAR3	The conference/court has increased the respect we hae for one another in my family.[b]	3.83	2.07	2.04	1.25

Note. RS = reintegration, SS = stigmatization, SEE = self-esteem, SG = shame-guilt, U = unresolved shame, E = embarrassment-exposure, FAR = fairness, FS = family support, and RES = offender responsibility. Items with an "a" superscript are on a 1-4 scale; those with a "b" superscript are on a 1-5 scale, and those with a "c" superscript are on a 1-8 scale.

Reintegration

The reintegration construct was measured using five items that captured the support and forgiveness shown to the offender following an offense (e.g., "How much support was the offender given during the court case?"; $\alpha = 0.81$).

Stigmatization

The stigmatization construct was assessed using four items that measured the disrespect and rejection shown to the offender following an offense (e.g., "During the conference/court case were you treated as though you were a bad person?"; $\alpha = 0.76$).

Self-Esteem

Offender self-esteem was modeled using five items from the Rosenberg (1979) Self-Esteem Scale (e.g., "I give up easily"; $\alpha = 0.75$).

Perceived Fairness

The perceived fairness construct was measured using five items that assessed the offender's feelings of being treated fairly during court or conference (e.g., "The police were fair during conference/court"; $\alpha = 0.84$).

Shame-Guilt

The shame-guilt construct was modeled using five items that measured the degree to which the offender felt ashamed or sorry following an offense (e.g., "During the conference/court case, I felt ashamed of what I did"; $\alpha = 0.86$).

Embarrassment-Exposure

The embarrassment-exposure construct was formed using four items regarding the feelings of awkwardness and exposure experienced by the offender following an offense (e.g., "During the conference/court case I felt awkward and aware of myself"; $\alpha = 0.67$).

Unresolved Shame

The unresolved shame construct was composed of two items to measure the experience of negative, unresolved emotionality following an offense (e.g., "Things brought up in the conference/court are unresolved in your mind"; $\alpha = 0.67$).

Offender Responsibility

Offender responsibility was measured by three items that assessed the degree of responsibility taken by the offender following an offense (e.g., "To what extent did the offender accept they had done wrong?"; $\alpha = 0.81$).

Family Support

The family support construct was modeled using three items created to measure the offender's perception of their family's acceptance of them following a criminal act (e.g., "I feel proud to be a member of my family"; $\alpha = 0.84$).

Statistical Analyses

Structural equation modeling (SEM) techniques were used to examine the proposed hypotheses. All analyses were conducted using LISREL 8.8 (Jöreskog & Sörbom, 2006). In accordance with the recommendations of Anderson and Gerbing

(1988), a two-step modeling approach was used to create a confirmatory measurement model in the first step, and a structural model in the second step. Within each of these steps, a multigroup structured means modeling (SMM) approach was conducted to examine invariance (i.e., whether the same model is invariant across groups; Thompson & Green, 2006; see also Byrne, 1991; Meredith, 1993; Widaman & Reise, 1997) for offenders randomly assigned to either RST conferencing or traditional court processing. This technique allowed for testing differences in the latent means between groups. The multigroup SMM assesses whether offenders assigned to conferencing or court reported differential mean scores on each of the nine constructs listed above.

Use of multigroup SEM allowed the researcher to assess both a confirmatory measurement model and a structural model for tests of invariance. The confirmatory and structural models were specified in accordance with Braithwaite's (1989) predictions within the framework of RST. Thus, one possible underlying theoretical model of RST was assessed. In the following sections, the two-step modeling approach will be discussed, followed by a description of the steps followed to examine invariance. It is important to note that invariance testing was conducted at both steps of the two-step modeling approach. Thus, invariance between conference or court offenders was assessed for the confirmatory measurement model as well as the structural model.

Two-Step Modeling Approach

Step one: Confirmatory measurement model. A confirmatory measurement model was tested that specified the relationship between the observed indicator variables and the latent constructs. All latent constructs within the confirmatory model were

allowed to freely covary. As was explained above, items were chosen from the Act Justice Survey and Global Observational Ratings Instrument to measure the nine latent constructs in the model, before any statistical analyses were conducted.

When specifying a confirmatory measurement model it is recommended to specify a model where each latent construct is unidimensional in order to accurately interpret the substantive meaning of a construct (Anderson & Gerbing, 1982). Thus, no items within the confirmatory measurement model in the current study were allowed to load on more than one latent construct, nor were any measurement errors specified to be correlated. This type of unidimensional measurement is also referred to as congeneric measurement (Joreskog, 1971b), and allows for clearer interpretation of the constructs being estimated. All latent constructs were composed of at least two or more indicators to increase content validity and reliability of each latent construct (Anderson & Gerbing, 1988). Further, Bollen (1989) noted that in order for a model with two or more latent factors to be identified, each latent factor must be composed of at least two indicators.

Along with specifying the pattern of loadings for the measurement model, the identification of the model must be also established. For a model to be identified, two conditions of the model must be met: (a) the number of free parameters in the model must be less than or equal to the number of observations in the covariance matrix being estimated, and (b) every latent variable must have a scale. As instructed by Kline (2005), one indicator per latent construct was fixed to 1 (i.e., referred to as a unit loading identification constraint). The indicator that is fixed to 1 becomes the reference variable and assigns a scale to each latent construct based on the common variance of the

indicator that was fixed to 1. Kline emphasizes that it is preferable to fix one indicator per factor in multigroup modeling as opposed to fixing the factor variance. This method simplifies interpretation in multigroup analysis because interpretation stems from unstandardized, rather than standardized coefficients. Unstandardized coefficients are necessary because comparisons made across groups on model parameters requires that the measured variable scores be on the same scale for both groups and standardizing scores leads to a differential rescaling of scores for each group (Widamin & Reise, 1997). Fixing the variance of a factor to 1 results in a standardization of the estimates and would produce inappropriate estimates in the present application.

Once the measurement model was properly specified to test the a priori hypotheses, the parameters of the measurement model were estimated. Following the recommendations of Finney and DiStefano (2006), model parameters (i.e., parameter values, standard errors, and fit indices) were estimated using the Maximum Likelihood (ML) normal theory estimator, as the data met the following conditions: (a) the data were continuous or categorical with four or more categories, and (b) the data were multivariate-normally distributed. All analyses were conducted using the variance-covariance matrix, rather than a correlation matrix, because standard ML estimation assumes that the variables are unstandardized, and analysis of standardized variables can lead to inaccurate standard error estimates (Kline, 2005).

Goodness of model fit was assessed according to the recommendations of Hu and Bentler (1999; see also Marsh, Hau, & Wen, 2004). Models were deemed acceptable if one of the following two criteria was met: (a) Non-normed Fit Index (NNFI) and the

Comparative Fit Index (CFI) were greater than or equal to .96 *and* the standardized root mean squared residual (SRMR) was less than or equal to .09, or (b) SRMR was less than or equal to .09 *and* the root mean squared error of approximation (RMSEA) was less than or equal to .06. The likelihood ratio χ^2 was not used as a means of determining whether the model displayed substantial global misfit because of its sensitivity to sample size (Bentler & Bonett, 1980), although this statistic was used when comparing invariance between models, as will be discussed in detail below.

In addition to assessing the acceptability of model fit indices, models were also examined by evaluating the strength and significance of all indicator loadings (Byrne, 1991; Kline, 2005). Model respecification and reestimation were performed in order to produce an acceptably fitting model on occasions when the model did not produce acceptable fit according to the global fit indices or nonsignificant indicator loadings were present. Model respecification was guided by theory (as recommended by Young, 1971) and the modification indices produced by LISREL (Sorbom, 1989). Modification indices specify the minimum amount of decrease in the overall model χ^2 if a specific parameter were freed. Large model modification values indicate that the fit would improve substantially if a certain indicator was allowed to freely load on a certain latent construct. Problematically, relying on modification indices to respecify a model can lead to problems with capitalization on chance because of idiosyncratic characteristics of the particular dataset being analyzed (MacCallum, Roznowski, & Necowitz, 1992). Thus, in the current study, no model modifications were made unless substantively justified by theory.

Examination of indicator variables. Anderson and Gerbing (1988) maintain that there are four options that can be used if an indicator is deemed unacceptable. If it appears that an indicator is contributing to unacceptable model fit, the researcher can specify the indicator to load on a different factor, delete the indicator, allow the indicator to load on multiple factors, or specify correlated measurement errors between indicators. In order to preserve the unidimensionality of the model, these authors recommend that only the first two options should be used. Thus, for the current study, any indicators that had nonsignificant loadings were removed from the model. This removal was done in a successive fashion, where only one indicator was removed at a time and the model reexamined after the removal of each item. Significant indicators remaining in the model were then "interpreted as unstandardized regression coefficients that estimate the direct effects of the factors on the indicators" (p. 176, Kline, 2005). All indicators that were fixed to 1 to set the scale of the factor were not tested for statistical significance because they are constants in the model.

Step two: Structural model. Once acceptable data-model fit was established for the measurement model, a structural model was estimated that specified the relationships between the latent constructs (Anderson & Gerbing, 1988). Specifically, the final model estimated from the confirmatory factor analysis in step one, including any parameters freed during the first step, was used to assess the structural relations between the latent constructs. In the structural model, direct structural paths were specified between latent constructs in order to test a subset of the theoretical assumptions of RST as theorized by Braithwaite (1989).

In order to determine whether the structural model was supported by the data, the goodness of fit criteria were employed as outlined above. When unacceptable model fit was found (either for overall goodness of fit or due to non-significant structural paths) the model was respecified by adding or deleting structural paths and then reestimated (Kline, 2005). No structural paths were added or deleted unless done in accordance to substantive theory, in order to reduce the likelihood of capitalization on chance due to sample-specific characteristics in the data (MacCallum et al., 1992).

Measurement Invariance Testing

To reach conclusions regarding the fit of the measurement and structural models across two groups (e.g., conferencing vs. court), some model parameters must be constrained to equality across the two groups. Once chosen, parameters are constrained between the two groups, the model is estimated, and if the fit of the model is still acceptable, that constrained parameter is deemed "invariant" between groups. The parameters that are constrained to be invariant dictate the level of model equivalence tested by the researcher and the conclusions that can be drawn from the invariance analyses. Specifically, multiple components of a measurement or structural model can be held equivalent between groups, such as the indicator loadings, intercepts, variances/covariances, and error terms. The greater number of model parameters constrained between groups, the more restrictive the model, thus stronger conclusions regarding model equivalence can be drawn.

Listed in order of increasing restrictiveness (i.e., more constraints placed on the models), these levels of invariance are configural, metric ("weak"), partial metric, scalar

("strong"), partial scalar, and strict (Meredith, 1993; each of these forms of invariance will be described in greater detail below). Briefly, partial metric invariance is established when the two models being compared have generally invariant patterns of item loadings for each factor, with some loadings freed (constraints are released) across models. Similarly, partial scalar invariance is established when there is invariance across intercepts, with some intercepts freed. Scalar factorial invariance is demonstrated when between-group equivalence is established for the factor loadings and intercepts. To establish strict factorial invariance (which is rarely met in practice because it is the most restrictive form of invariance as all primary model components are held to equivalence), the factor loadings (i.e., slopes), intercepts (i.e., means), error variances and covariances, and factor variances and covariances are constrained to be invariant across the different groups (Meredith, 1993; Thompson & Green, 2006).

Thompson and Green (2006) note that with SMM models it is unnecessary for a model to maintain strict factorial invariance in order to reach valid conclusions regarding factorial invariance. Additionally, Byrne, Shavelson, and Muthen (1989) have stated that valid conclusions can be reached under conditions of partial metric and partial scalar invariance (i.e., not all of the factor loadings or intercepts are constrained to equality). Specifically, Bryne and colleagues maintain that the minimum requirement for proceeding under conditions of partial metric invariance is that at least one indicator loading per factor must be equivalent between groups. Similarly, to establish partial scalar invariance, at least one intercept per factor must also be invariant.

These increasingly strict forms of invariance testing were performed during step one and step two of the modeling process. Thus, these techniques were used to establish the most rigorous and acceptable confirmatory measurement model between groups. To reiterate, once a measurement model with acceptable fit was estimated, direct paths between the factors were specified in order to conduct step two of the modeling. In addition to estimating an acceptably fitting structural model, the invariance of the structural paths in the model between groups was assessed.

χ^2 difference testing. As described above, joint fit criteria (Hu & Bentler, 1999; Marsh et al., 2004) are used to determine the acceptability of data-model fit, or what is also referred to as the *practical* fit of the model. When testing for factorial invariance by constraining some parameters to equality between groups, the χ^2 likelihood ratio test is used to assess *relative* fit (Widaman & Reise, 1997), meaning the degradation in fit resulting from the imposed constraints. When conducting invariance testing, some parameters are constrained to equality between groups, depending upon the type of invariance being assessed. The fit of the constrained model is then compared to the fit of the unconstrained model. If there is a large discrepancy in fit between the unconstrained and constrained models, constraints placed on the parameters have degraded the fit. Thus, those constrained parameters are considered different, or variant between the two groups. In order to determine whether the discrepancy between the two models is important or "significant," the difference in the χ^2/df ratio between the two models is used. Thus, the relative fit of the constrained model in comparison to the unconstrained model is assessed.

If the difference in χ^2/df ratio between the constrained and unconstrained models is statistically significant, there is evidence for the less restrictive model providing a better fit to the data. If this is the case, the constraints placed upon the noninvariant parameters are then relaxed (Widaman & Reise, 1997), meaning some constraints are freed (i.e., removed) and those parameters are again allowed to vary between groups. A specification search is then conducted to determine which parameter(s) are not equivalent across the groups. This specification search is aided by the use of modification indices (Sorbom, 1989), as mentioned above. Following the recommendations of Byrne and colleagues (1989), noninvariant parameters for each specified model may be freed, one at a time, and the model reestimated. Constraints on the model parameters are released until the χ^2/df ratio between the constrained and unconstrained model is no longer statistically significantly. A nonsignificant $\Delta\chi^2$ indicates that any remaining invariant parameters are not introducing ill-fit. Importantly, the decision to release a constraint placed upon a parameter in the model is only made if there is substantive and theoretical justification for releasing the constraint (Byrne et al., 1989; Jöreskog, 1971a). Thus, constraints were not released just to improve the fit of the model. When conducting both steps of Anderson and Gerbing's (1988) two-step model building approach (i.e., developing an acceptable confirmatory measurement model and acceptable structural model), χ^2 difference testing was conducted to establish an acceptable model for both groups.

Configural invariance. To begin factorial invariance testing, the first step is to establish configural invariance between the groups being compared (Meredith, 1993; Thompson & Green, 2006; Widamin & Reise, 1997). Configural invariance is

established when the same *pattern* of fixed and free indicator loadings holds for both groups. Thus, the same number of factors or latent constructs must be present for both groups. In addition, these indicators must be explained by the latent factors with the same pattern of nonzero and zero indicator loadings (e.g., a path loading being estimated between a latent construct and an indicator versus a path not existing). This type of invariance is referred to as nonmetric invariance because no constraints are placed upon the models between groups. All parameters in a configural model are freely estimated and are allowed to vary between groups (Widaman & Reise, 1997). Although all parameters are free, when configural invariance is tested, the models for the two groups are "stacked" and analyzed separately, but simultaneously. Then, the fit of this stacked model is determined by evaluating the global fit indices described above. The χ^2 and *df* from this simultaneous estimation is the sum of the χ^2 and *df* values for the court and conference groups, because the χ^2 values are additive (Byrne, 1994). The establishment of configural invariance is critical. The configural model serves as a baseline model for subsequent invariance testing. If the data do not exhibit configural invariance, then more restrictive tests of invariance for the indicators and the intercepts are unwarranted (Bollen, 1989).

Metric invariance. Once configural invariance is established, the equivalence of indicator loadings between groups is assessed. This test establishes whether participants in the different groups responded to the items of the measure in the same manner. If metric or partial metric invariance is not established, differential item functioning may be present (Kline, 2005). Differential item functioning is problematic because it makes it

difficult to determine if the reported scores on items being measured are truly due to differences between the groups, or because the items are actually measuring the latent constructs in a different manner for each group.

Metric, or weak factorial invariance, is established when the factor loadings for the observed indicators to the latent constructs are equivalent between the groups. Metric invariance is established by constraining all of the indicator loadings to equivalence across groups and estimating the fit of the model. This metric model is then compared to the baseline configural model via a χ^2 difference test. If the $\Delta\chi^2$ test is significant, then the metric invariance assumption is not met and a search is conducted to determine which parameters should be freed (Thompson & Green, 2006; Widaman & Reise, 1997). If some of the constraints placed upon the indicators are relaxed, all subsequent analyses proceed under partial metric invariance (Byrne et al., 1989).

Scalar invariance. Once metric or partial metric invariance is established, scalar invariance is assessed. This condition implies that "subjects with the same value on the latent construct should have equal values on the observed variables" (p. 641; Hong, Malik, & Lee, 2003). Scalar invariance determines if the means of the observed variables are equivalent between the groups. To assess for scalar invariance, constraints are placed upon the intercepts (i.e., means for the measured variables). Importantly, all constraints placed upon the model to establish metric invariance remain in the model when assessing for scalar invariance (Thompson & Green, 2006; Widaman & Reise, 1997) and all indicators are allowed to freely vary between groups remain free. The *metric* invariant model serves as the baseline model of comparison with the intercept-constrained model,

and a χ^2 difference test is conducted to determine whether the fit of the intercept-constrained model is significantly worse than when the intercepts are allowed to freely vary. If the χ^2 difference test is statistically significant, a search is conducted to determine which intercepts should be allowed to freely vary between groups, and these constraints are then relaxed. Subsequent analyses then proceed under conditions of partial scalar invariance (Byrne et al., 1989). Meredith (1993) noted that the establishment of scalar (or partial scalar) invariance is essential if comparisons between latent mean scores are to be conducted.

Latent mean differences and factor intercept invariance. If the conditions of configural, metric, and scalar invariance (or partial invariance) are upheld, differences in the latent means between groups (i.e., offenders assigned to conference or court) can be assessed. This testing provides information regarding whether groups reported similar mean values on each of the nine factors in the model. When testing for full or partial scalar invariance, the differences in the latent means between groups are also estimated, although the values of the latent means for both groups cannot be estimated directly (Hancock, 1997). Instead, estimation of the differences in the latent means is accomplished by allowing the factor means, or K, to be freely estimated in one group while being fixed to zero in the second group. The group whose means are fixed to zero is then considered the reference group. The value for K that is obtained for the group that is freely estimated is thus the mean difference between groups on the latent construct (Hancock, 1997).

Statistically significant differences on the latent means can be assessed by examining standard *t*-score values, and any *t*-score greater than 2.00 is considered statistically significant at $\alpha = 0.05$. The magnitude of latent mean differences can be assessed by computing a Cohen's *d* effect size estimate for each mean difference. Cohen's *d* is calculated by dividing the mean difference for each factor by the pooled standard deviation across both groups (Cohen, 1988).

Along with the estimation of the differences in the latent means, factor invariance can be assessed if scalar (or partial scalar) invariance is established (Byrne et al., 1989; Meredith, 1993). With the use of the scalar (or partial scalar) invariant model that is obtained above, the equivalence of the factor means is assessed by first fixing the factor intercepts (factor means) to zero in both groups. All invariance testing is conducted in a hierarchical fashion, and because the models are nested, the fit of the models is compared via a χ^2 difference test. If the fit of the scalar (or partial scalar) model where the factor intercepts are fixed to zero (i.e., which presumes that the means are equal in both groups) is significantly worse than the scalar invariant model where the intercepts are allowed to be freely estimated, evidence is provided that there is a significant difference in some or all of the latent means between groups. This test serves as an omnibus test of the differences between all nine latent means between the two groups and is much like a traditional ANOVA used to compare mean differences in three or more groups (Thompson & Green, 2006).

Although this overall test of factor invariance is important, it does not provide information regarding *which* specific latent means are different between groups. To

determine which specific latent variables significantly differ between the two groups, the model where all factors are fixed to zero is modified and just one latent is constrained to equality at a time. All remaining latent variable means are allowed to be freely estimated. This model is then compared to the final scalar invariant model. A significant $\Delta\chi^2$ indicates that the latent constrained to equality is nonequivalent between the groups (Thompson & Green, 2006). This test was conducted for all nine latent factors individually, in order to determine if a significant decline in fit resulted when each latent factor was tested, thus providing evidence for the invariance of each latent factor.

Structural model invariance. Once an invariant (or partially invariant) confirmatory measurement model is estimated, directional structural paths between factors may be specified, in accordance with step two of Anderson and Gerbing's two-step approach (1988). These structural paths represent hypotheses of correlation or causation between factors. All equality constraints placed upon the confirmatory measurement model remain when testing the structural models. The acceptably fitting structural model with no equality constraints placed upon the structural paths between groups serves as the baseline model to assess the invariance of the structural parameters. The goodness of fit of this model may be assessed according to the joint fit criteria specified by Hu and Bentler (1999).

In addition to estimating an acceptable structural model according to the steps discussed in detail above, the invariance of the structural parameters between groups is examined. This test provides information regarding whether the structural paths are equivalent between groups. Invariance testing of the structural parameters proceeds by

first fixing all structural paths to equality between the two groups. A χ^2 difference test is then conducted between the unconstrained (i.e., baseline model) and fully-constrained structural models. This procedure is analogous to the measurement invariance testing delineated above. A statistically significant $\Delta\chi^2$ value indicates that the fit of the constrained model is significantly worse, indicating that some or all of the structural paths are noninvariant between groups. To determine which structural path constraints should be relaxed, equality constraints on the structural paths are freed, one by one, and the relative fit of the model in comparison to the unconstrained model is re-assessed using a χ^2 difference test after each constraint is released. Thus, the invariance between groups for each individual structural path in the model is examined, to determine whether the data empirically support the hypothesized relationships between factors.

Hypotheses

SEM with multigroup analysis was conducted to first create a confirmatory measurement model. This model was created to assess whether the nine hypothesized latent constructs were measured similarly (i.e., invariant) between offenders randomly assigned to RST conferencing or traditional court processing following arrest for a drunk driving offense. In the original RISE study, participants were randomly assigned to one of two treatment groups (RST conferencing or court) to determine if attending RST conferencing led to greater perceived fairness and lowered rates of recidivism. These researchers (Tyler et al., 2007) hypothesized that the two groups would report *different* values on the items in the two measurement instruments. Similarly, in the current study,

it was hypothesized that although the measurement instrument should measure all nine constructs similarly between the two groups (i.e., thus displaying metric invariance), the random assignment of offenders to two different treatment conditions should result in different values on some of the means for the indicators (i.e., thus displaying partial scalar invariance). Further, if there was indeed a treatment effect, offenders in the two groups should also display different mean values on the factors (i.e, the test of differences in latent means and factor intercept invariance). Importantly, multigroup SEM is a powerful analytic tool that can be used to identify which specific mean values for the intercepts are invariant and which differ significantly between groups. In regards to the structural model, it was hypothesized that the structural parameters would be invariant between groups, thus providing evidence that the theoretical model derived from Braithwaite's (1989) hypotheses was empirically supported for both groups.

The specific hypotheses for the measurement invariance tests are provided below, followed by the structural hypotheses. The hypothesized confirmatory measurement model is provided in Figure 1, while the hypothesized structural model is displayed in Figure 2.

Measurement Invariance Hypotheses

1. The data will display metric invariance between offenders assigned to conference or court; the strength and significance of all indicator loadings will be invariant between conference or court offenders.

2. The data will display partial scalar invariance between offenders assigned to conference or court; the intercepts (means) for some of the individual indicator items will

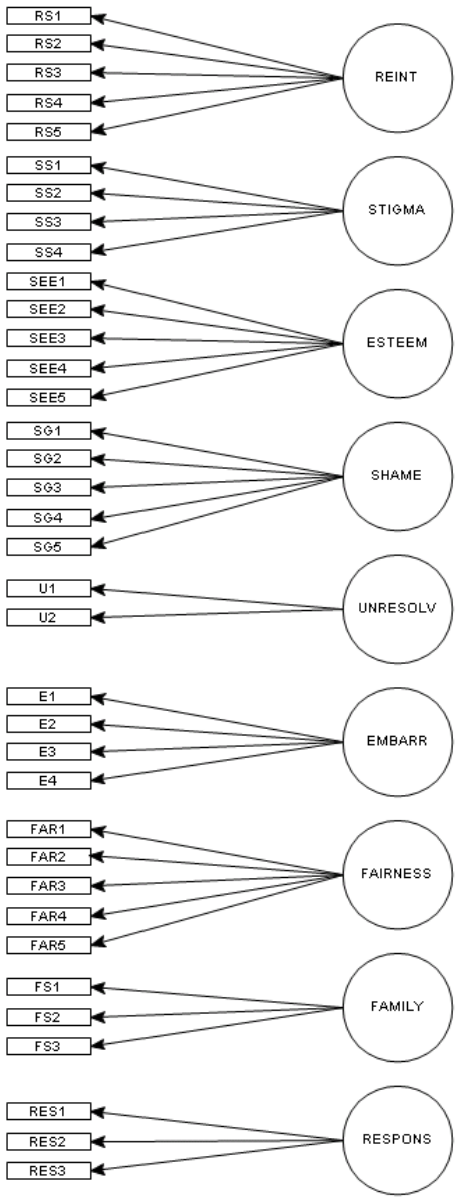

Figure 1. Hypothesized confirmatory measurement model for all offenders.
Note. Although all latent constructs were allowed to freely covary, for simplicity, all two-headed arrows representing correlations among latent constructs are omitted from the figure. RS = reintegration, SS = stigmatization, SEE = self-esteem, SG = shame-guilt, U = unresolved shame, E = embarrassment-exposure, FAR = fairness, FS = family support, and RES = offender responsibility.

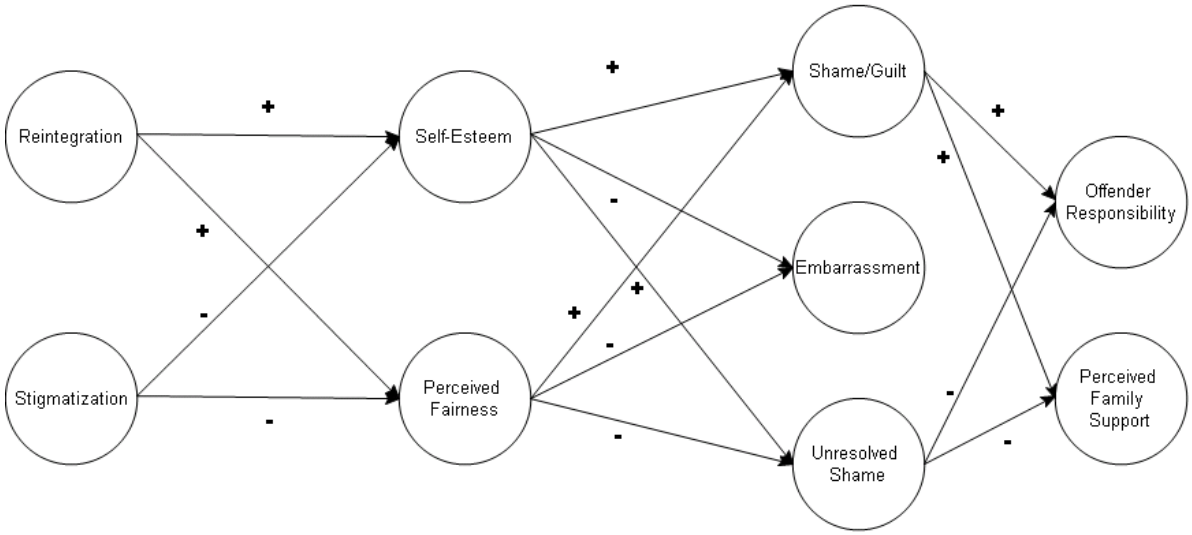

Figure 2. Hypothesized structural model for all offenders.
Note. "+" signs indicate a hypothesized positive association between factors, while a "-" sign indicates a negative association between factors.

be different for conference or court offenders. Due to the random assignment of offenders to conference or court, participants assigned to conference will report higher means on some indicators measuring reintegration, fairness, self-esteem, shame-guilt, family support, and offender responsibility. Offenders randomly assigned to court will report higher means on some of the indicators measuring stigmatization and unresolved shame.

3. The data will not display factor intercept invariance: the factor means will be different for offenders assigned to conference or court. Specifically, offenders assigned to conference will report higher latent mean scores on the latent constructs reintegration, fairness, self-esteem, shame-guilt, family support, and offender responsibility. Offenders

assigned to court will report higher latent mean scores for the constructs stigmatization, unresolved shame, and embarrassment-exposure.

Structural Hypotheses

1. The self-esteem construct will serve as a mediating variable between the types of shaming experienced and the emotion felt by the offender.

2. The perceived fairness of the conference/court will serve as a mediating variable between the type of shaming experienced and the emotion felt by the offender.

3. The type of shaming (i.e., reintegrative versus stigmatizing) experienced by the offender will indirectly influence the acceptance of responsibility by the offender and their perceived familial support following an offense, through the self-esteem, fairness, and emotion constructs.

4. The type of emotion experienced by the offender will differentially impact their responsibility and familial support. Shame-guilt will be the strongest predictor of increased offender responsibility and perceived familial support, while unresolved shame will be the strongest predictor of decreased offender responsibility and perceived familial support.

CHAPTER IV

RESULTS

Descriptive Statistics

Before proceeding with SEM analyses, preliminary analyses were conducted in order to assess how well the data met the assumptions of SEM. Descriptive statistics in the form of means and standard deviations for each item were computed and are presented in Table 1. Along with these descriptive statistics, skewness and kurtosis values were calculated to assess for the normality of each indicator item, because maximum likelihood estimates can be attenuated when the normality assumption is violated (Finney & DiStefano, 2006). No indicator items displayed skewness (i.e., > ±3.0) or kurtosis (i.e., > ±10.0) values exceeding criteria as specified by Kline (2005). Thus, all analyses proceeded with the use of maximum likelihood estimation.

In regards to missing data, a small percentage of data was missing from the 724 total cases used in the current analyses (no variable had data missing for more than 20 cases). The missing data displayed no discernable patterns and was deemed to be missing at random. Multiple imputation with maximum likelihood estimation was conducted using LISREL (Jöreskog & Sörbom, 2006) so all cases could be included in the SEM analyses (Kline, 2005).

Tests for Factorial Validity

Multiple-group SEM techniques were used to assess the proposed hypotheses. Models were estimated using maximum likelihood estimation procedures with the data variance-covariance matrix (see Appendices A and B). All invariance testing proceeded in a hierarchical fashion by testing nested models in the order of configural, metric, and scalar invariance.

Configural Invariance

Statistical analyses proceeded by developing an adequate confirmatory measurement model. The first step was to assess configural invariance between groups, where the invariance of the pattern of factor loadings for both groups was assessed (Figure 1). The test for configural invariance determines if acceptable fit for the measurement model is achieved for both conference and court offenders, where the same model is estimated for both conference or court offenders separately, but simultaneously. This model then served as the baseline model for the subsequent metric invariance testing. Forty items were included in the initial model with nine latent constructs, where all latent constructs were allowed to freely covary. One indicator per latent construct was constrained to 1 in order to scale the latent construct. The fit of this initial model was adequate according to the fit specified by Hu and Bentler (1999), $\chi^2 = 2599.52$, $df = 1408$, NNFI and CFI = 0.94, SRMR = 0.062, RMSEA = 0.048 ($90\%_{CI} = 0.045 – 0.051$).

In order to improve model fit and increase model parsimony for each latent construct, items were removed based on a consideration of item content and by

investigating each item's reliability and standardized factor loading. Items were removed successively by removing one item and then reassessing the fit of the model. Four items were removed, for a total of 36 retained items in the measurement model (see Table 1). The fit of this revised 36-item measurement model was adequate for the stacked model where the factor loadings are estimated separately and simultaneously for conference or court offenders, $\chi^2 = 1842.79$, $df = 1116$, NNFI = 0.95, CFI = 0.96, SRMR = 0.059, RMSEA = 0.042 ($90\%_{CI} = 0.039 – 0.046$). The standardized and unstandardized indicator loading values are presented for conference or court offenders in Table 2. All factor loadings were statistically significant at $p < .05$. The fit of the model thus supported configural invariance, providing evidence for the hypothesis that the pattern of fixed and free parameter loadings in the nine-construct confirmatory measurement model is invariant between conference or court offenders.

Metric Invariance

The second step for testing measurement invariance was conducted by examining the equivalence of the indicator loadings (weak invariance) between court and conference groups by constraining all indicator loadings between groups to equality. The difference in fit of the baseline configural model (Model 1) and the indicator-constrained model (Model 2) was compared via a χ^2difference test because Model 2 is nested within Model 1. A significant $\Delta\chi^2$ indicates that the fit of the constrained model is significantly worse, meaning that not all of the indicator loadings are equivalent between groups. For the present study, this was indeed the case, as the difference between the χ^2 statistic for Model 1 and Model 2 was statistically significant at $p < .05$ (see Table 3 for a summary

Table 2

Indicator Loadings for the Configural Baseline Model

	Standardixed (unstandardized)	
Indicator	Conference	Court
FAR1	0.62 (0.76)	0.81 (0.98)
FAR2	0.59 (0.70)	0.81 (0.97)
FAR3	0.71 (1.00)	0.71 (1.00)
FAR4	0.61 (0.86)	0.71 (1.01)
FAR5	0.53 (0.58)	0.74 (0.81)
SS1	0.54 (0.66)	0.60 (0.73)
SS2	0.45 (0.53)	0.66 (0.79)
SS3	0.78 (1.00)	0.78 (1.00)
SS4	0.73 (0.85)	0.74 (0.86)
RS1	0.73 (1.23)	0.56 (0.94)
RS2	0.72 (1.28)	0.54 (0.96)
RS3	0.61 (1.00)	0.61 (1.00)
RS4	0.78 (1.55)	0.52 (1.03)
RS5	0.76 (1.28)	0.55 (0.93)
SEE1	0.68 (0.72)	0.48 (0.51)
SEE2	0.55 (0.72)	0.56 (0.74)
SEE3	0.71 (0.93)	0.70 (0.91)
SEE4	0.43 (0.44)	0.48 (0.49)
SEE5	0.77 (1.00)	0.77 (1.00)
F1	0.73 (0.94)	0.73 (0.93)
F2	0.72 (0.87)	0.93 (1.13)
F3	0.82 (1.00)	0.82 (1.00)
RES1	0.95 (1.00)	0.95 (1.00)
RES2	0.66 (0.83)	0.57 (0.71)
RES3	0.69 (0.66)	0.84 (0.80)
SG1	0.49 (0.48)	0.71 (0.70)
SG2	0.73 (1.00)	0.73 (1.00)
SG3	0.85 (1.10)	0.94 (1.21)
SG4	0.92 (1.36)	0.90 (1.32)
SG5	0.65 (0.88)	0.50 (0.68)

table continues)

Indicator	Standardixed (unstandardized) Conference	Court
E1	0.82 (1.00)	0.82 (1.00)
E2	0.80 (1.05)	0.82 (1.08)
E3	0.36 (0.36)	0.56 (0.57)
E4	0.67 (0.82)	0.77 (0.94)
U1	0.71 (1.22)	0.95 (1.63)
U2	0.57 (1.00)	0.57 (1.00)

Note. FAR = Perceived Fairness, SS = Stigmatization, RS = Reintegration, SEE = Self-Esteem, FS = Family Support, RES = Offender Responsibility, SG = Shame-Guilt, E = Embarrassment-Exposure, and U = Unresolved Shame

of model fit for the measurement invariance testing). Thus, a decision was made to reject the full metric invariance model and a specification search was conducted to identify which indicators were not equivalent between groups, and should be freely estimated (i.e., allowed to differ between models).

Decisions with respect to which indicators should be freely estimated were based on modification indices (Sorbom, 1989) and substantive theory. As was done above when assessing configural invariance, equality constraints were relaxed successively by removing items one at a time and re-assessing model fit. In total, four items were allowed to freely covary between conference or court offenders (Model 3). Specifically, two items loading on the shame-guilt construct, one item loading on the family support construct, and one item loading on the offender responsibility construct were found to vary between the groups, while all other indicator loadings were found to be equivalent.

An assessment of the distributions of the four noninvariant parameters was conducted to attempt to determine why these four parameters were noninvariant between groups. The skewness, kurtosis, variance, and presence of floor or ceiling effects were

examined for these four indicators separately for offenders assigned to conference or court. When examining these data qualities, it appeared that differential skewness and kurtosis values on the two shame items could have accounted for the noninvariance that was detected with these two parameters. Specifically, for both shame items, the kurtosis values were discrepant (i.e., ~ 2.0) between the two groups, although none of the kurtosis or skewness values exceeded the recommended cut-offs for nonnormality. Further, for the shame item "My actions hurt others," the frequency distribution for the court group displayed somewhat of a floor effect. In contrast, the conference group responses did not exhibit this pattern (i.e., the frequency of offenders responding that their actions did not hurt others was much greater in the court group in comparison to the conference group). Although no problems were apparent for the family support item displaying noninvariance, the frequency of responses for the offender responsibility item, "Responsibility offender took," displayed a slight ceiling effect for both groups, but to a much greater degree for the offenders assigned to conferencing (i.e., conference individuals were more likely to take responsibility than court offenders). The finding of noninvariance on these four items indicates that these items may be measuring their respective constructs differentially between groups, and should thus be investigated more fully in future studies utilizing the Act Justice Survey and Global Observational Ratings Instrument.

Allowing four items to vary and constraining the remaining 31 items to equivalence between the groups resulted in a nonsignificant χ^2 test between Model 1 and the partially-constrained Model 3, as well as an adequate fitting model (see Table 3).

Between-group equivalence was not established for all indicator loadings, thus subsequent analyses proceeded under partial measurement invariance. These findings partially support measurement invariance Hypothesis 1.

Scalar Invariance

Having established partial metric invariance, tests of scalar invariance were conducted. Scalar invariance was assessed by constraining the intercepts (means) of the 31 items found to be metrically equivalent in the above analyses. This step utilizes the same techniques as the prior step, except constraints are now imposed on the intercepts (means of the indicator items). These models are also nested, thus a χ^2difference test was used to compare the relative fit of the partial metric model and the full scalar model (Model 4, see Table 3). Again, a significant χ^2 value indicates that not all of the intercepts between groups are equivalent, and if so, a specification search is conducted to determine which intercepts should be freely estimated. Importantly, in the current study, conference and court participants were randomly assigned to two different treatment groups. Thus, it was hypothesized that the groups would display only partial scalar invariance, as it was anticipated that they would have different mean values on some of the indicators.

When testing for full scalar factorial invariance in the current study, all intercepts were constrained to equality between conference or court offenders and Models 3 and 4 were compared using a χ^2difference test. As shown in Table 3, the difference in χ^2 values was statistically significant, thus a decision was made to reject the full scalar model, as

Table 3

Assessing Factorial Invariance Between Conference and Court Offenders

Model	χ^2 Difference	*Df* difference	Decision
Test of Full Metric Invariance (Model 1 vs. Model 2)	68.17	27	Reject
Test of Partial Metric Invariance (Model 1 vs. Model 3)	34.66	23	Accept
Test of Full Scalar Invariance (Model 3 vs. Model 4)	966.43	22	Reject
Test of Partial Scalar Invariance (Model 3 vs. Model 5)	17.99	10	Accept
Test of Factor Invariance (Model 5 vs. Model 6)	366.12	9	Reject

Note. "Reject" decisions were based on $\alpha = 0.05$ for the estimated difference in χ^2.

hypothesized. Constraining all intercepts also resulted in a significant decline in fit according to RMSEA, CFI, and NNFI values (see Table 4). Thus, a specification search was conducted to determine which intercepts should be allowed to freely vary between the conference or court offenders, with the use of substantive theory and modification indices to guide the process. Constraints imposed on the intercepts were then relaxed, one-by-one, and the fit of the model re-assessed after the removal of the constraint on each noninvariant intercept.

In total, 12 intercepts were allowed to freely vary between groups before a nonsignificant χ^2 difference was estimated. Importantly, as noted by Byrne and colleagues (1989), full metric or full scalar invariance is not a necessary prerequisite for conducting further tests of invariance (e.g., factor invariance), as long as one indicator or

Table 4

Fit Indices for Invariance Tests

Model	χ^2 (*df*)	RMSEA ($CI_{90\%}$)	NNFI	CFI
Configural Model (Baseline Model: Model 1)	1842.79 (1116)	0.042 (0.039 - 0.046)	0.95	0.96
Full Metric Invariance (Model 2)	1910.96 (1143)	0.043 (0.040 - 0.047)	0.95	0.95
Partial Metric Invariance (Model 3)	1877.45 (1139)	0.042 (0.039 - 0.046)	0.95	0.96
Partial Metric, Full Scalar Invariance (Model 4)	2843.88 (1161)	0.063 (0.06 -0.066)	0.90	0.91
Partial Metric, Partial Scalar Invariance (Model 5)	1895.44 (1149)	0.042 (0.039- 0.046)	0.95	0.96
Partial Metric, Partial Scalar, Full Factor Invariance (Model 6)	2261.56 (1158)	0.051 (0.048 - 0.055)	0.90	0.91

Note. RMSEA = root mean square error of approximation; CI = confidence interval; NNFI = nonnormed fit index; CFI = comparative fit index.

intercept per latent construct remains constrained to equality. In the current analyses, this condition was met. Thus, as shown in Table 3, a decision was made to accept this partial scalar model, Model 5, which indicated that conference or court offenders reported different mean values on 12 of the indicator items. Table 4 also displays the fit indices for Model 5 and it is clear that the fit of this model according to RMSEA, NNFI, and CFI improved substantially and was acceptable when the equality constraints were relaxed for the noninvariant intercepts. These findings provide partial support for measurement invariance Hypothesis 2, as the intercepts or means for only a portion of the indicators

were found to be invariant between groups. Further analyses thus proceeded under conditions of partial metric and partial scalar invariance.

Test of Latent Mean Differences

To test for significant differences in the latent means between offenders assigned to conference or court, the differences in latent means were estimated indirectly by fixing the latent mean scores to zero in the reference group, while allowing them to be freely estimated in the other group. The value that is estimated is thus the difference in latent means between groups (Hancock, 1997). In the current study, the court group was treated as the reference group. Thus the estimated values for the conference group served as the estimated difference in the latent means between conference and court offenders. The estimated values for the latent means for both groups, as well as the difference in these means, are presented in Table 5. The statistical significance of the mean difference on the latent means between the groups is evaluated using a *t* test. All *t* values were statistically significant (Table 5).

The positive or negative value of the mean difference indicates if the offenders assigned to court or conference reported higher values. As the court group was assigned as the reference group, positive values indicate that the conference group reported a higher mean value. All hypothesized mean difference values were supported, with the exception of the offender responsibility and unresolved shame constructs. Cohen's *d* effect size estimates were also calculated to assess the strength of these means differences (Cohen, 1988) and these estimates are provided in Table 5.

Table 5

Tests for Latent Mean Differences Between Groups

Latent	Conference (*M*)	Court (*M*)	Mean Diff.	*t* score	Cohen's *d*
Fair	1.85	1.16	0.69*	9.63	0.53
Stigma	0.05	0.36	-0.31*	-4.88	-0.21
Reint	2.49	0.49	2.01*	28.30	1.23
Esteem	1.73	1.59	0.14*	3.01	0.08
Family	3.84	2.04	1.80*	26.18	1.37
Respons	3.67	4.96	-1.29*	-9.78	-2.14
Shame	1.65	1.27	0.38*	5.55	0.32
Embarr	1.40	1.11	0.29*	3.32	0.28
Unresolv	-1.11	-0.78	-0.32*	-5.04	-0.16

Note. Fair = Fairness, Stigma = Stigmatization, Reint = Reintegration, Esteem = Self-Esteem, Family = Family Support, Respons = Offender Responsibility, Shame = Shame-Guilt, Embarr = Embarrassment, and Unresolv = Unresolved Shame. * indicates that the mean difference is statistically significant at $\alpha = 0.05$.

In addition to the estimation of the difference in the latent means, factorial intercept invariance between the two groups was assessed. To determine if the model displayed factorial intercept invariance, the factor intercepts (factor means) were fixed to zero in both groups (Model 6). Model 6 was nested within Model 5, thus these two models were compared via a χ^2 difference test to determine whether some or all of the latent means scores differed between groups (Thompson & Green, 2006). To determine which specific latent variables significantly varied between the two groups, for each

latent, Model 6 was modified where just one latent was constrained to equality, rather than all latent variables. All remaining latent variables were allowed to be freely estimated. This model was then compared to Model 5, and a significant decline in fit according to a χ^2 difference test with 1 degree of freedom indicated that that specific latent was different between the groups.

In the current study, the estimation of Model 6 where all factor intercepts were fixed to zero in both groups resulted in a significant decline in fit, as reflected in the increase in RMSEA and decrease in NNFI and CFI (see Table 4). Thus, constraining the factor intercepts to be equal resulted in an ill-fitting model, suggesting that the factor intercepts are non-equivalent between groups (i.e. much like an omnibus ANOVA to detect mean differences between groups). In order to determine which of the nine factor intercepts were non-invariant, each factor intercept was constrained to equivalence between the groups, one at a time, with all other intercepts being freely estimated. This modified Model 6 was then compared to Model 5 for each latent variable in an iterative fashion. Thus, this test was performed nine times (once for each latent construct), in order to precisely determine if each factor intercept was invariant between groups. For all nine latent constructs, the χ^2 difference test was statistically significant, thus providing further support for the hypothesis that the conference and court offenders would differ on all factor intercepts.

Structural Model Testing

According to Anderson and Gerbing's (1988) two-step modeling approach, once an adequate measurement model is developed, the structural parameters specifying the directionality between constructs can be tested. For the present study, the hypothesized structural model in Figure 2 served as an initial model to be tested for both conference and court offenders. It is important to note that when the structural model is tested, all constraints placed upon the measurement model in the previous steps remained in the structural model; all indicators and intercepts constrained to equality between the two groups remained constrained when testing the structural model. The magnitude of the relationships between the latent constructs was initially assessed by examining the correlation matrix between all latent variables, for conference or court attendees separately, as shown in Table 6.

When testing the hypothesized structural model in Figure 2 for both groups, the fit of the model was not sufficient when the structural parameters were unconstrained between groups (Table 7). Thus, parameters were added to the hypothesized model to achieve an acceptably fitting model. This re-specification was grounded in RST theory and the modification indices were used for guidance. The model was modified by including three additional structural paths: (a) shame-guilt to embarrassment-exposure, (b) reintegration to family support, and (c) stigmatization to unresolved shame. The addition of these three structural paths produced a model with acceptable goodness of fit

Table 6

Correlations Among Latent Factors

		Fair	Stigma	Reint	Esteem	Family	Respons	Shame	Embarr
Conference									
	Stigma	-0.47***							
	Reint	0.22***	-0.07						
	Esteem	0.15*	-0.13	0.12					
	Family	0.25***	-0.19**	0.46***	0.13*				
	Respons	0.10	-0.07*	0.10	0.07	0.04			
	Shame	0.25***	-0.01	0.44***	-0.15*	0.32***	0.28***		
	Embarr	0.02*	0.26***	0.32***	-0.34***	0.27*	0.16**	0.65***	
	Unresolv	-0.52***	0.62***	-0.14	-0.19*	-0.15*	-0.11	-0.02	0.19***
Court									
	Stigma	-0.69***							
	Reint	0.26***	-0.12						
	Esteem	0.10	-0.14*	0.01					
	Family	0.21**	-0.18**	0.41***	0.09				
	Respons	-0.02	-0.14*	-0.06	0.03	0.04			
	Shame	0.00	0.10	0.31***	0.00	0.22***	0.02		
	Embarr	-0.15**	0.24***	0.21**	-0.18**	0.15**	0.08	0.72***	
	Unresolv	-0.66***	0.71***	-0.14*	-0.07	-0.05	-0.18**	0.12	0.26***

Note. Conf = Conference group, Fair = Fairness, Stigma = Stigmatization, Reint = Reintegration, Esteem = Self-Esteem, Family = Family Support, Respons = Offender Responsibility, Shame = Shame-Guilt, Embarr = Embarrassment, and Unresolv = Unresolved Shame. *** = $p < .001$, ** $p < .01$, * = $p < .05$.

Table 7

Fit Indices for Structural Model

Model	χ^2 (*df*)	RMSEA ($CI_{90\%}$)	NNFI	CFI	SRMR
Hypothesized	2464.48 (1191)	0.054 (0.051 - 0.057)	0.92	0.92	0.097
Modified- 3 add. structural paths (Baseline model)	2047.53 (1185)	0.045 (0.042 - 0.048)	0.95	0.95	0.077
Fully-constrained structural model	2114.68 (1202)	0.046 (0.043 - 0.049)	0.94	0.95	0.085
Partially-constrained structural model	2065.77 (1197)	0.045 (0.042 - 0.048)	0.95	0.95	0.081

Note. RMSEA = root mean square error of approximation; CI = confidence interval; NNFI = non-normed fit index; CFI = comparative fit index; SRMR = standardized root mean residual.

Table 8). This modified and unconstrained model then served as a baseline model to determine if the structural paths were invariant between conference or court offenders.

In order to assess whether the structural model was invariant between conference and court offenders, all structural paths were fixed to equality between the two groups and a χ^2 difference test was conducted between the unconstrained and fully-constrained structural models. The difference in χ^2 values was statistically significant as shown in Table 7 ($\Delta\chi^2 = 67.15$, $df = 17$, $p < .001$). A specification search was conducted to isolate potential sources of ill-fit. Each equality constraint placed upon the structural paths was freed, one by one, and the fit of the model re-assessed after each constraint was released. In total, five structural parameters were allowed to freely vary between conference or court offenders (Table 8). Direct paths between the following latent constructs were allowed to be freely estimated between groups in the final

constrained structural model: (a) stigmatization to perceived fairness, (b) perceived fairness to unresolved shame, (c) perceived fairness to embarrassment-exposure, (d) shame-guilt to offender responsibility, and (e) shame-guilt to family support. The final partially-constrained structural models for the conference and court participants are displayed in Figures 3 and 4, respectively.

Allowing these five structural paths to be freely estimated for both groups resulted in a nonsignificant $\Delta\chi^2$ between the unconstrained structural model and the partially-constrained structural model ($\Delta\chi^2 = 18.24$, $df = 12$, $p > .05$). The fit of this partially-constrained structural model was acceptable, as shown in Table 7. Model fit can also be ascertained by investigating the squared multiple correlations for all outcome variables in the model, as this is an indication of the percentage of variance explained by these factors. For the final model, the squared multiple correlations were: fairness = 0.36, self-esteem = 0.02, shame-guilt = 0.43, embarrassment-exposure = 0.09, unresolved shame = 0.56, offender responsibility = 0.05, and family support = 0.13.

Direct Effects

Although this partially constrained model produced acceptable fit, not all structural parameters were statistically significant for both groups. To allow for easy comparison between the estimates of the structural paths for the constrained and unconstrained structural models, Table 8 shows the estimated standardized direct effects for both groups separately when the model was unconstrained between groups and the standardized direct effects for the final structural model when the structural paths were constrained to equality. For the equality-constrained standardized path loadings, separate values are provided for the conference and court Groups for the five parameters that were freed. As shown in Table 8, when modeled separately,

Table 8

Direct Effects for Structural Model

	Standardized path coefficients		
Path	Conference	Court	Constrained (Conference/Court)
Reint → Esteem	0.11	0.00	0.04
Reint → Fairness	0.20**	0.18*	0.19***
Reint → Family	0.44**	0.31***	0.35***
Stigma → Esteem	-0.17*	-0.13*	-0.14*
Stigma → Fairness	-0.44***	-0.69***	-0.45*** (-0.68)***
Stigma → Unresolv	0.39***	0.53***	0.44***
Esteem → Shame	0.05	0.13*	0.10*
Esteem → Unresolv	-0.06	0.04	-0.03
Esteem → Embarr	-0.31***	-0.18*	-0.26***
Fair → Shame	0.27***	0.10*	0.16***
Fair → Unresolv	-0.24***	-0.35***	-0.21*** (-0.42)***
Fair → Embarr	0.06	-0.15*	0.07 (-0.15)**
Shame → Response	0.28***	0.05	0.28*** (0.05)
Shame → Family	0.24***	0.12*	0.27*** (0.12)*
Shame → Embarr	0.73***	0.70***	0.73***
Unresolv → Response	-0.13**	-0.15**	-0.14**
Unresolv → Family	-0.21*	-0.03	-0.06

Note. Conf. – conference group. Fair – Fairness; Stigma = Stigmatization; Reint = Reintegration; Esteem = Self-Esteem; Family = Family Support; Respons = Offender Responsibilityp; Shame = Shame-Guilt; Embarr = Embarrassment; and Unresolv = Unresolved Shame. For the freed parameters, effect estimates for conference attendees are listed first, followed by court attendees in parentheses.
*** = $p < .001$; ** = $p < .01$; * = $p < .05$.

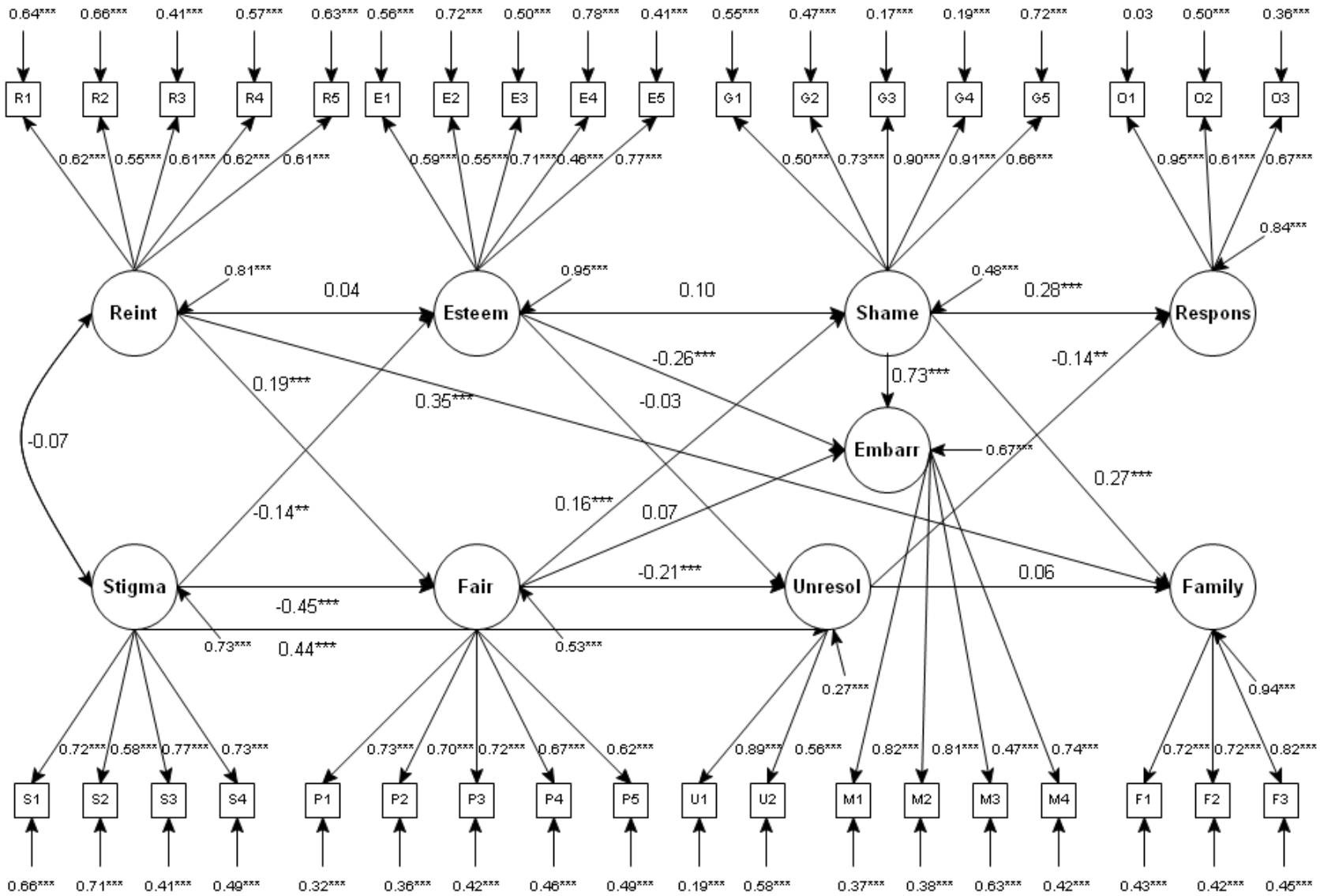

Figure 3. Final constrained structural model for the conference group. *** = $p < .001$, ** = $p < .01$, * = $p < .05$.

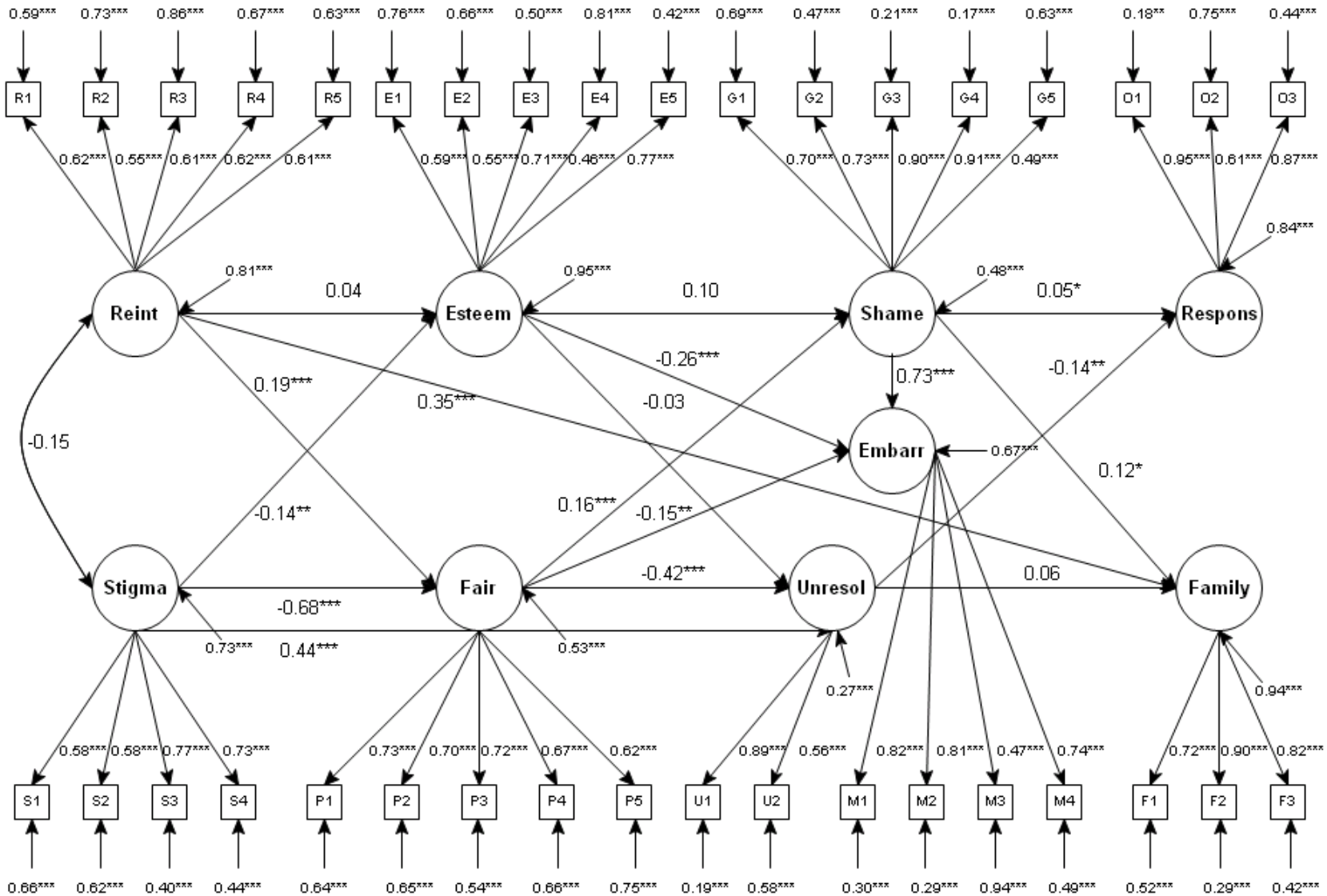

Figure 4. Final constrained structural model for the court group. *** = $p < .001$, ** = $p < .01$, * = $p < .05$.

11 of the structural paths were found to be statistically significant for both groups, while three were not statistically significant for either group. All path coefficients were in the hypothesized direction. Further, in two cases, the structural path was statistically significant for one group, while not significant for the other group (e.g., perceived fairness to embarrassment-exposure).

Indirect Effects

Table 9 shows the standardized total and indirect effects separately for participants assigned to conference or court, with values for the court attendees shown in italics. Although these values are for the equality-constrained structural model, the values for the total and indirect effects differ between conference or court participants because of the five structural paths that were allowed to freely covary between groups. Similar to the direct effect estimates discussed above, many of the key hypothesized relationships between constructs were statistically significant, supporting the strength and importance of the hypothesized relationships between latent constructs as specified by Braithwaite (1989). For example, the standardized total effects of reintegration and stigmatization on the outcome variables family support and offender responsibility were significant (with the exception of family support in the court group). An interpretation of this finding is that for the conference group, increased perceptions of being reintegratively shamed lead to increased family support and responsibility acceptance, while the opposite pattern resulted from increased stigmatizing shaming practices. The same interpretation applies for the court group, with the exception of the family support variable.

Table 9

Standardized Total and Indirect Effects for Structural Model

	Dependent variables						
Independent variable	Fair	Esteem	Shame	Unresolv	Embarr	Family	Respons
	Total effects						
Reint	0.20*	0.04	0.03*	-0.05*	0.00	0.31*	0.01*
	*0.18**	*0.05*	*0.01*	*-0.07**	*-0.04**	*0.42**	*0.01**
Stigma	-0.45*	-0.12*	-0.07*	0.64*	0.01	-0.04*	-0.09*
	*-0.68**	*-0.16**	*-0.02*	*0.67**	*0.14**	*-0.06*	*-0.11**
Fair			0.19*	-0.25*	0.07	0.06*	0.08*
			0.05	*-0.39**	*-0.16**	*0.04*	*0.07**
Esteem			-0.10*	-0.04	-0.30*	-0.02	-0.02
			*-0.09**	*-0.02*	*-0.24**	*-0.01*	*0.00*
Embarr			0.67*			0.16*	0.18*
			*0.77**			*0.10*	*0.04*
Shame						0.24*	0.27*
						0.13	*0.05*
Unresolv						-0.04	-0.10*
						-0.08	*-0.17**

(*table continues*)

	Dependent variables						
Independent variable	Fair	Esteem	Shame	Unresolv	Embarr	Family	Respons
			Indirect effects				
Reint			0.03*	-0.05*	0.00	0.01*	0.01*
			0.01	*-0.07**	*-0.04**	*0.01*	*0.01**
Stigma			-0.07*	0.12*	0.01	-0.04*	-0.09*
			-0.02	*0.27**	*0.14**	*-0.06*	*-0.11**
Fair			0.05			0.06*	0.08*
			*-0.12**			*0.04*	*0.07**
Esteem			-0.20*			-0.02	-0.02
			*-0.18**			*-0.01*	*0.00*
Embarr						0.16*	0.18*
						0.10	*0.04*

Note. Fair = Fairness, Stigma = Stigmatization, Reint = Reintegration, Esteem = Self-Esteem, Family = Family Support, Respons = Offender Responsibility, Shame = Shame-Guilt, Embarr = Embarrassment, and Unresolv = Unresolved Shame. Effects are listed separately for conference or court offenders, with court effects shown in italics.
* $p < .05$.

CHAPTER V

DISCUSSION

Summary of Findings

Throughout the United States, retributive justice practices (processing through the traditional court system) are most commonly utilized to deter criminal activity. Unfortunately, this typical response to crime has been found to be ineffective because rates of recidivism in the United States continue to remain high (Lagan & Levin, 2002). In response to this problem, the restorative justice alternative to crime prevention termed RST (Braithwaite, 1989) has been implemented in various countries throughout the world including the United States, but minimal empirical testing of this approach has been conducted. Thus, the goals of the present study were to test Braithwaite's theoretical model underlying RST and subsequently assess the effectiveness of this approach in response to offenders committing drunk driving offenses.

Multigroup SEM techniques were used to address these two goals. In the current study, nine latent constructs (i.e., reintegration, stigmatization, perceived fairness, self-esteem, shame-guilt, embarrassment-exposure, unresolved shame, offender responsibility, and family support) were measured by 36 directly observed variables (i.e., self-report items and other-report observational items). Using these 36 indicators, confirmatory measurement and structural models were tested between offenders assigned to RST conferencing or court processing following a drunk driving offense. Tests of

group invariance supported the predicted confirmatory measurement model, indicating that the items used to assess RST for the two groups of offenders functioned similarly between groups.

Direct structural paths, specifying expected covariance, between the nine latent constructs were then added to the measurement model to assess the fit of a structural model. The predicted structural model was also supported, providing evidence for the validity of the structural model developed in accordance with Braithwaite's hypotheses regarding the relationships between core RST constructs. Finally, to assess the efficacy of RST in response to crime, tests of latent mean differences between groups were conducted to determine if offenders randomly assigned to RST conferencing or court processing reported significantly different mean values on the nine aforementioned latent constructs. The findings support the hypothesis that attending RST conferencing in comparison to court processing is an effective response to crime, because more positive outcomes are evidenced following exposure to RST conferencing (e.g., greater perceived fairness or higher perceived family support). It appears that RST conferencing may be a viable alternative to the more frequently practiced retributive justice for drunk driving offenses. Thus, continued research and perhaps a wider implementation of this form of justice are warranted. A more detailed discussion of study findings and implications follows.

Measurement Invariance in Confirmatory Measurement Model

When assessing for measurement invariance using multigroup SEM techniques, a model was created that established the relationships between the manifest and latent variables. Once the fit of this model was acceptable for both groups (combined and when measured separately), equivalency constraints between groups were placed upon various parameters in the model (i.e., indicators, means). These equivalency constraints were used to determine if those parameters were invariant between groups. Comparisons using a χ^2 difference test between more constrained and less constrained models (nested models) were then made. Significant change in the χ^2 values between these nested models indicated that the relative fit of the model changed due to the constraint imposed on the model parameter, thus the model was deemed noninvariant between groups on that specific parameter. Depending upon the type of parameter that was being constrained to equality across groups, tests of configural (i.e., overall pattern of fixed and free indicator loadings), metric (i.e., indicators to factors), and scalar (i.e., intercepts for measures) invariance were conducted, as well as testing for significant differences in latent means.

In the current study, offenders arrested for drunk driving in Australia were randomly assigned to receive reintegrative shaming conferencing or traditional court processing after arrest. Measurement and structural invariance of a theoretically driven model was assessed between these two groups. When testing for factorial invariance between groups, the invariance of the confirmatory measurement model was assessed according to Anderson and Gerbing's (1988) two-step modeling approach. This analysis proceeded by first testing for configural invariance between groups. No further

invariance testing can proceed unless configural invariance is supported. The data in the current study supported the existence of configural factorial invariance between offenders assigned to conference or court: the fit of the confirmatory measurement model with nine latent variables measured by 36 items was acceptable for both groups. Thus the same pattern of fixed and free parameter loadings was empirically supported. This model then served as the baseline model for the subsequent invariance testing to address the proposed hypotheses.

Measurement invariance hypothesis 1: Metric invariance. Metric invariance between groups was tested using the configural measurement model. The hypothesis was that the strength of the indicator loadings to the latent constructs would be equivalent between offenders assigned to conference and court. This hypothesis was partially supported, because only four indicators were considered non-invariant between groups. Subsequently, the equality constraints placed upon these non-invariant indicators were relaxed. Model fit was acceptable when these four indicators were allowed to perform differently between the court and conference groups, while the remaining 32 of 36 (89%) indicators performed identically. Thus, the model displayed partial metric invariance, that according to Byrne and colleagues (1989), is sufficient in order to proceed to tests of scalar invariance.

The four items notwithstanding, measurement invariance to this point was supported. This is an important finding because in the next step, the hypothesis involves differences in means between the two groups (court vs. conference). Failure to find invariance in the current step would weaken any subsequent findings by allowing for an

alternate hypothesis: means might differ simply because the items were not performing similarly across the two groups, not because the treatment had an effect. Establishing measurement invariance via SEM rules out this alternate hypothesis and addresses a potential problem not addressed by performing a simple *t* test.

Measurement invariance hypothesis 2: Scalar invariance. Once partial metric invariance was established, the intercepts for the measures (i.e., means of the items), were constrained to equality between groups to test for scalar invariance. Although it was hypothesized that the model would display metric invariance, in the current study, it was hypothesized that the measurement model would only display partial scalar invariance. This hypothesis was posed because offenders were randomly assigned to two different treatment conditions (i.e., conferencing or court), thus it was hypothesized that if RST is effective, participants in the two different conditions would report different values on some of the items of the assessment instruments.

When assessing for scalar invariance, 12 of the intercepts (item means) were found to vary between offenders assigned to conference or court. Specifically, items representing seven of the nine latent constructs in the model (i.e., all latent constructs except offender responsibility and stigmatization) were noninvariant between groups. This partial noninvariance indicates that offenders randomly assigned to conferencing or court reported different mean values on 12 of the indicator variables in the model. Thus, measurement invariance hypothesis 2 was found to be partly supported. It must be noted that Byrne and colleagues (1989) maintain that at least one intercept per latent construct

must remain invariant between groups to assess for significant differences in the latent means and this condition was met in the current study.

Measurement invariance hypothesis 3: Factor intercept invariance. The measurement model in the current study displayed partial metric and partial scalar invariance. Thus, tests of factor intercept (latent means) invariance and significance testing of the differences in these latent means between groups were warranted. Again, as conference and court participants were subjected to two different forms of treatment, it was hypothesized that the factor means would be different between the groups. Specifically, the prediction was that offenders assigned to conference would report higher mean scores on the latent constructs reintegration, fairness, self-esteem, shame-guilt, family support, and offender responsibility. In contrast, offenders assigned to court would report higher latent mean scores for the constructs stigmatization, unresolved shame, and embarrassment exposure.

Much like hypotheses 1 and 2, hypothesis 3 was also partially supported by the data. Tests of factor intercept invariance revealed that when the factor means were held to equivalence between the two groups, the relative fit of the model decreased significantly. This finding provides evidence that the latent means were significantly different between the two groups. Significance tests of latent mean differences were then conducted to determine which means varied significantly between groups. In support of hypothesis 3, it was found that offenders assigned to RST conferencing reported greater perceived fairness, reintegration, self-esteem, family support, and shame-guilt, as well as lower stigmatization and unresolved shame. In opposition to the current hypotheses,

offenders assigned to conferencing also reported higher embarrassment-exposure and lower acceptance of responsibility. These findings lend general support to the claims made by practitioners and researchers of the restorative justice paradigm, as attendance of RST conferencing led to a variety of positive outcomes for offenders in comparison to attendance of a court hearing. Moreover, these positive outcomes have been found to be linked to a reduction in future criminal behavior (i.e., greater family support has been found to reduce the likelihood of future crime), thus making it even more important for the criminal justice community to consider RST as a viable alternative to traditional court processing.

Structural Model

Model estimation and re-specification. While the measurement model is a test of the relationships between the observed indicators and latent constructs, the structural model is a test of the relationships (i.e., covariances) *between* latent constructs**.** Once an acceptable measurement model was estimated, the parameters for the hypothesized structural model in Figure 2 were estimated. This model was first estimated with all structural parameters unconstrained between groups (i.e., relationships between constructs were allowed to differ in magnitude between groups). Once an acceptable model fit was achieved, the structural paths were constrained to equality between groups and the overall model fit was again assessed. In order to achieve acceptable fit, three additional paths were added to the model. The necessity of adding these paths indicates that the original model failed to capture significant covariance between the respective constructs. The additional paths were direct paths between: (a) shame-guilt and

embarrassment-exposure, (b) reintegration and family support, and (c) stigmatization and unresolved shame.

The addition of these three structural paths was supported by the theory of reintegrative shaming and previous research conducted with these constructs. Specifically, when Harris (Ahmed et al., 2001) tested the dimensionality of the three types of shame emotions, he found that the shame-guilt and embarrassment-exposure constructs were significantly correlated at 0.60 for those offenders attending conference, and 0.52 for court cases. It was not surprising that these constructs were highly correlated in the current study as well, thus a path that modeled this relationship was added.

In regards to the addition of the direct path between reintegration and family support, Braithwaite (1989) stated that reintegrative forms of shaming were practiced more frequently and more successfully in highly interdependent communities characterized by a high degree of social support. The hypothesized structural model in the current study proposed that the relationship between reintegration and family support would be mediated by self-esteem, perceived fairness, and the shame emotions experienced by the offender, but the data provided empirical evidence for the direct association between reintegrative shaming and subsequent feelings of family support. Thus, the addition of this structural path provides support for the importance of this direct relationship, as emphasized by Braithwaite.

Finally, in terms of the addition of the direct path between stigmatization and unresolved shame, although it was hypothesized that this relationship would also be

mediated by the self-esteem of the offender and the perceptions of being treated fairly, the data provided empirical support for the direct association between these constructs. In the earliest version of RST as well as in the later revisions, Braithwaite (1989; Ahmed et al., 2001) emphasized the detrimental impact of stigmatizing forms of shaming in terms of increases in subsequent criminal behavior, increased feelings of isolation from a support group disparaging of crime, and importantly, a failure to acknowledge the experience of shame. The addition of the direct path between stigmatization and unresolved shame reflects Braithwaite's, as well as Harris's (2003), assumptions regarding these constructs: both researchers emphasized that stigmatizing shaming is associated with unsuccessfully acknowledging the shame experience. Thus, the painful feelings resulting from the experience of shame are not discharged, leading to increased negative emotionality, including anger or decreased acceptance of responsibility for a crime.

Structural model interpretation. After estimating an acceptable structural model when all structural paths were freely estimated between groups, the structural parameters were fixed to equality between groups and the relative fit of the model was assessed. Fixing all structural paths to equality resulted in a statistically significant decline in fit. Overall, five structural paths were found to be non-invariant between the two groups and the equality constraints placed upon these paths were relaxed. This partially invariant structural model served as the final structural model for the present study and the estimates in Figures 3 and 4 were used to assess the strength of the relationships between the nine latent constructs. When interpreting the estimates in

Figures 3 and 4 from the model, five of the structural paths have different estimates (i.e., one for conference and one for court), because these structural paths were non-invariant between groups.

As can be seen in Table 8, the strength of all hypothesized direct structural paths, with the exception of three (i.e., reintegration to self-esteem, self-esteem to unresolved shame, and unresolved shame to family support), were supported by the data. As predicted, reintegrative and stigmatizing forms of shaming had a differential impact on self-esteem and perceived fairness. Specifically, a high level of reintegrative shaming was associated with greater perceptions of being treated fairly, while a higher level of stigmatizing shaming was related to decreased self-esteem and a lower level of perceived fairness (although the relationship between reintegration and self-esteem was not significant). In turn, higher reported levels of self-esteem and perceived fairness were significantly related to higher reported experiences of shame-guilt (i.e., adaptive shame) and lower levels of embarrassment-exposure (although the relationship between fairness and embarrassment-exposure was not significant for the conference group). Greater perceived fairness also corresponded to lower reported unresolved shame. Finally, greater shame-guilt was significantly related to greater offender responsibility and family support, while unresolved shame was significantly related to less offender responsibility acceptance.

Along with interpreting the direct effects, the total and indirect effects between constructs can also be interpreted (see Table 9). When looking at the total effect estimates, it is clear that reintegrative and stigmatizing forms of shaming have a

differential impact on the other constructs in the model, as predicted by Braithwaite (1989). For example, greater reintegrative shaming was found to be positively and significantly associated with greater perceived fairness, as well as greater shame-guilt, lower unresolved shame, greater perceived family support, and greater offender responsibility. The opposite pattern was found in relation to stigmatizing shaming practices. Similarly, higher levels of fairness were also related to increased offender responsibility and family support. These findings provide empirical support for Braithwaite's hypotheses regarding the importance of implementing reintegrative forms of shaming following a crime, as this adaptive form of shaming appears to be related to a variety of positive outcomes that are related to crime reduction.

Implications for the Theory of Reintegrative Shaming

Reintegrative and stigmatizing shaming. The present study provides strong empirical support for the differential impact of reintegrative versus stigmatizing shaming practices, through two avenues of findings: latent mean differences and the strength of the structural paths. To begin, and almost most importantly in terms of the application of these practices, offenders assigned to reintegrative shaming conferencing reported higher mean levels of being shamed reintegratively versus stigmatically, while the opposite pattern was found for offenders attending a traditional court hearing. This finding could be viewed as a manipulation check to assess how well reintegrative shaming was practiced in the treatment group. Along with reporting that they were treated more "reintegratively," conference attendees also reported statistically significantly higher

mean values on self-esteem, fairness, family support, and the experience of the shame-guilt emotion in comparison to court attendees.

Along with the latent mean comparisons, the strength of the direct and total effects provides further support for the hypothesized relationships between the type of shaming practiced and the experience of positive or negative outcomes. Specifically, Braithwaite's (1989) hypotheses that reintegrative shaming practices would be associated with higher self-esteem, greater perceived fairness, greater shame acknowledgement (i.e., the experience of shame-guilt), increased support from family or friends, and greater responsibility acceptance were supported, while the opposite pattern was hypothesized and supported in regards to stigmatizing shaming. These findings provide evidence to support the claims that reintegrative shaming practices should be employed following a crime, in order to lead to the most successful outcome for the offender and society.

In opposition to the stated hypotheses, individuals attending reintegrative shaming conferencing did not report higher mean values of offender responsibility or lower levels of embarrassment-exposure. Although the hypotheses regarding these mean differences were not empirically supported, the total effect of reintegrative shaming on offender responsibility was significant and positive (although small), while greater reintegrative shaming was associated with a reduction in unresolved shame (the opposite pattern of results were found for stigmatizing shaming). The reason why there was a contradiction in findings between the structural parameters and mean differences is unclear, but an attempt was made to discern why this outcome occurred. Two plausible explanations for these discrepant findings are as follows: First, the offender responsibility construct was

the only construct in the model that was measured solely by observational items, as no self-report items were collected to assess this construct in the original study. Thus, the RISE employees conducting the observations could have reported their findings unreliably or unsystematically and this could have impacted the results. Second, the offender responsibility and embarrassment-exposure constructs were measured with the fewest number of indicators in comparison to the other constructs in the model and these items were more highly skewed than the other items in the model (although the skewness values were not great enough to violate assumptions of normality). Thus, rather than concluding that reintegrative shaming practices are not associated with greater offender responsibility or less unresolved shame, the accuracy and reliability of the measurement of these constructs should be considered when assessing these findings. Moreover, because of these considerations and as is true of all structural models, further research should be conducted to more completely assess the relationships between these constructs.

Self-esteem and the shame emotions. The findings from the present study provide partial support for the hypothesized mediating role that self-esteem serves between reintegrative shaming and the resultant experience of the three measured shame emotions. Throughout the self-esteem literature, controversy exists regarding the positive or negative role that varying levels of self-esteem may play in leading to criminal behavior, drug use, happiness, or other life outcomes (e.g., Baumeister et al., 1993, 1996; Leary et al., 2003). The current study sought to shed light on this controversy by testing if the type of shaming an offender experiences impacts their level of self-esteem, and in

turn, if different levels of self-esteem may lead to the experience of positive or negative shame emotions.

In the current study, the direct relationship between reintegrative shaming and self-esteem was not statistically significant, nor was the direct relationship between self-esteem and unresolved shame. On the other hand, greater stigmatizing shaming was directly associated with lower reported self-esteem, and greater self-esteem was associated with greater reported experiences of shame-guilt, in support of the proposed hypotheses. In terms of the indirect relationships between the shaming constructs and the resultant three experienced shame emotions, the hypothesized relationships were generally supported for both groups of offenders (see Table 11); greater reintegrative shaming was indirectly and significantly associated with greater shame-guilt, less unresolved shame, and less reported embarrassment-exposure, while the opposite pattern was found in terms of stigmatizing shaming.

These findings are in direct accord with the propositions by Braithwaite, Harris, Leary, and Scheff, as previously discussed. Specifically, the type of shaming practiced was differentially related to the reported self-esteem or self-views of the offender, as discussed by Rosenberg (1979) and Scheff (1990). These findings can similarly be interpreted through Leary's (Leary & Downs, 1995) sociometer theory and Braithwaite's social threat conception of shame: being shamed (by qualitatively different forms of shaming methods) by others in our social group was associated with higher or lower self-esteem, thus demonstrating that losing esteem in the eyes of others can alter how we view ourselves.

Further, the results in the current study support Harris's (2003) findings regarding the impact of shaming on the experience of shame emotions. Specifically, higher reported levels of experiencing the adaptive shame emotion characterized by an acknowledgement of shame (i.e., shame-guilt) were associated with reintegrative shaming practices, while greater reported levels of shame displacement (i.e., unresolved shame) were reported in relation to greater stigmatizing shaming. In turn, greater shame-guilt was associated with greater responsibility acceptance and perceived family support, while the opposite pattern was found regarding the experience of increased levels of unresolved shame (with the exception that the relationship between unresolved shame and family support was not statistically significant). These findings are in direct correspondence with Harris's (2003) hypotheses and empirical findings regarding the relationships between the types of shame emotions and resultant positive or negative outcomes (e.g., responsibility acceptance, directing negative emotionality at the self versus another individual).

Although these findings are promising because of the direction and significance of the *direct* relationships between constructs in the structural model, the indirect relationships between self-esteem and the ultimate outcomes in the model (i.e., responsibility and family support) were found to be weak and nonsignificant. The coefficients of these indirect relationships were generally zero for both groups. Thus, interpretation of these results in terms of the controversy regarding the adaptive or maladaptive nature of high self-esteem does not appear warranted.

When attempting to determine why these relationships were weak, one plausible explanation became apparent. The items in the original RISE study chosen to measure self-esteem were taken from the 10-item Rosenberg (1979) Self-esteem Scale. Problematically, this scale measures self-esteem as a trait, rather than as a state experience. In the current study, it was hypothesized that differential levels of reintegrative or stigmatizing shaming would lead to changes in self-esteem (i.e., higher or lower self-esteem, respectively), but because the measure was created to assess self-esteem as a *trait* rather than a *state*, the measure may not be able to detect changes in self-esteem following the experience of different forms of shaming. Thus, the use of a trait measure may be inadequate for the current purpose.

Shaming and shame as beneficial, not detrimental, experiences. Running counter to the claims made by various criminological- and emotion-theorists (e.g., Gilligan, 2003; Maxwell & Morris, 2002; Tangney & Dearing, 2002), in the present study, forms of shaming and shame served beneficial functions for offenders following a drunk driving violation. Specifically, reintegrative forms of shaming and the experience of the adaptive shame-guilt emotion were associated with increased perceptions of fairness, responsibility acceptance, and family support. These findings are in direct contradiction to the generalizations that the emotion shame is a dark, troubling experience leading to violent behavior (Gilligan, 2003; Tangney & Dearing, 2002), and the presumption that *all* forms of shaming can be construed as stigmatizing and should be avoided (Maxwell & Morris, 2002).

Unlike reintegrative shaming and the experience of shame-guilt, stigmatizing shaming and unresolved-shame were indeed associated with more detrimental outcomes (i.e., lowered perceived fairness, responsibility acceptance, and family support), much like the outcomes predicted by these same researchers. These findings support the notion that the contradictions in the emotion and criminological literature regarding the adaptive or maladaptive nature of shaming and shame could be due to two reasons: (a) how these constructs are conceptualized in the literature, and (b) how these constructs are measured. As discussed previously, researchers such as Gilligan (2003) regard shaming as a form of rejection, abandonment, and disrespect, in line with Braithwaite's conceptualization of *stigmatizing* shaming. If a distinction is not made regarding the form that shaming takes, investigations of the impact that shaming may have on offender outcomes may continue to produce equivocal results.

Not only can disparate findings result from differential conceptualizations of shaming, but how shame is operationalized in empirical investigations can substantially impact the conclusions drawn regarding outcomes associated with the experience of shame. For example, most conclusions regarding the adaptive or maladaptive nature of shame employ the most widely used instrument, the Test of Self-Conscious Affect (TOSCA; Tangney, Wagner, & Gramzow, 1989), which operationalizes shame as highly maladaptive. Although this measure is the most widely used throughout the emotion literature (Robins, Noftle, & Tracy, 2007), it has been criticized by various researchers (Luyten et al., 2002; Sabini & Silver, 1997). For example, Luyten and colleagues (2002) concluded that although the TOSCA shame scale is positively correlated with many

maladaptive traits, drawing the overall conclusion that *shame* itself is maladaptive is unwarranted because the TOSCA "only measures maladaptive thoughts or behaviors associated with shame" (p. 1380). Specifically, shame is characterized by intensely negative self-evaluation, rumination over the offense that was committed, and intrapunitive responses. When operationalized in such a manner, it appears that shame as characterized by Tangney and colleagues is much like Braithwaite's (1989) and Harris's (2003, 2006) definition of *unresolved-shame*. Thus, much like the argument made previously regarding the importance of distinguishing between different types of shaming, when assessing the relationships between shame (as an experience) and various life outcomes, it is essential to differentiate between adaptive and maladaptive forms of shame as well.

Limitations and Future Research Directions

The findings from the present study are important on two levels. A dearth of research has been conducted to test the validity of the theory underlying reintegrative shaming practices in the criminal justice arena. Thus, the above conclusions can be used to guide further development or modification of a theory that is receiving increased attention in the criminal justice and emotion literature. Secondly, the findings from the present study can also be used to guide the application of RST. For example, empirical evidence from the present study supported the notion that reintegrative, in contrast to stigmatizing shaming practices conducted following a criminal offense, were associated with many positive outcomes for offenders. These positive outcomes include increased

perceived fairness, greater shame-guilt emotional responses, and greater offender responsibility. Thus, these findings can be used to support the continued or increased application of RST following a criminal offense. Although the present study is important along these lines, all empirical studies have limitations that should be addressed in order to guide future research in a similar area of study.

External validity. In the present study, criminal offenders committing a drunk driving offense were randomly assigned to undergo RST conferencing or the traditional adjudicative process following their arrest. The results provide support for the hypothesis that offenders receiving RST conferencing perceived their experience to be more positive than those who attended a traditional court hearing on a variety of outcomes. Ideally, RST conferencing would have the same beneficial effects when applied to all types of criminal offenses (e.g., theft, assault, white-collar crime), as hypothesized by Braithwaite (1989). Problematically, within the present study, only offenders that were arrested for a drunk driving offense were included, as the sample size was unacceptably low for the other offense types that were included in the original RISE investigation.

Shadish and colleagues (2002) provide a very in-depth discussion on a researchers' ability to generalize their research findings, or the external validity of a study. One type of study characteristic upon which generalizations are made is on the unit under investigation, or the sample being studied. One possible limitation to the current study is that, because the sample included only drunk driving offenders in Australia, generalizations regarding study outcomes can only be applied to other drunk driving cases and that RST may not be effective with other classes of offenses. Further,

there was no clear indication if the offender was a first-time or repeat drunk driving offender (although randomization to conferencing and court should eliminate any potential bias this may have in terms of main effects between groups). Although the external validity of the study may consequently be weakened, two points should be addressed that counter the argument against the applicability of RST to other offense types.

First, previous investigations of RST have been conducted with a variety of other offense types and, although they are correlational in nature, RST was found to be effective in cases where offenders committed a variety of offenses, such as white-collar crime (Makkai & Braithwaite, 1994; Murphy & Harris, 2007), bullying (Ahmed & Braithwaite, 2004; 2005; Ttofi & Farrington, 2008), and adolescent delinquency (Losoncz & Tyson, 2007). Therefore, despite employing a limited sample, the present study lends support to RST and makes clear that further studies of RST are warranted.

Second, it is often the case that there is no clear victim when an individual is arrested for a drunk driving offense and Braithwaite (1989) specified that an offender being in the presense of the victim of their crime during RST conferencing is one reason that RST is particularly effective. In the current study, offenders were only recruited to the RISE study if there was no victim resultant from the offense, in order to reduce variability between cases. A community representative instead attended the conference in place of a victim to discuss the potential negative impact that drunk driving can have on them or the community as a whole. In addition to the community representative, each offender was shown a video displaying car crashes occurring as a result of drunk driving,

emphasizing the harm that can be caused by this offense (Ahmed et al., 2001). Because no actual victim was present, it is possible that the results of the present study could be even stronger if the sample consisted of offenders whose crimes impacted a clearly identifiable victim. On the other hand, because this sample was specifically limited to victim-free offenses, caution is urged when making such a generalization without further study being conducted with drunk driving offenses involving a victim. Thus, the conclusions drawn from the present study should be used as a springboard for future research to be conducted with offenders of multiple offense types (and with a longer duration between offense and follow-up to ascertain the long-term impact of RST) in order to determine the limits to which RST should be generalized.

Along with generalizations to other units, Shadish and colleagues (2002) discuss that a second study characteristic upon which generalizations can be made is to other outcomes. The present investigation was conducted using an extant dataset and the types of outcomes that were assessed were chosen by the original investigators. Although generalizations may be made in regards to the nine latent constructs that were assessed, the external validity of the present study is limited, not only in terms of the nine constructs that were actually measured, but by the specific items that were used to measure these nine latent constructs. Similarly, the time-frame for which these constructs were assessed via the Act Justice Survey varied between 4-6 weeks, and it is unknown if the offender's ability to recall their experiences differed systematically based upon the duration of time that elapsed between offense and assessment. Thus, in future investigations of RST, it would be important to include additional measures of the same

constructs (e.g., state, rather than trait measures of self-esteem), as well as to measure other relevant constructs, while keeping the duration between offense and interview constant. Such important additional constructs could include the moderating impact that an offender's shame-proneness may have on their response to RST conferencing (Tangney & Dearing, 2002), or the level of collectivism of the community in which the offender resided (Braithwaite, 1989). These two constructs, as well as others, could play a vital role in determining the impact of RST on criminal offending.

Model re-specification. Similar to the criticism that the external validity of the present study may be limited to the units and outcomes actually studied, a second limitation is that the findings from the present study may have restricted generalizability because modification indices (Sorbom, 1989) were used when re-specifying the measurement model. Many authors have noted that sole reliance on modification indices to make model revisions can lead to problems with capitalizing on chance, or sample-specific idiosyncrasies in the data (e.g., Cheung & Rensvold, 1999; MacCallum, Wegener, Uchino, & Fabrigar, 1993; Young, 1971). This potential problem can lead researchers to build measurement and structural models that only describe the current sample, and do not generalize to other samples of data. In order to reduce the probability of capitalizing on chance, in the current study, only modifications that were consistent with theory were made. This reliance on theory, rather than numerical output to guide changes, precluded many modifications (e.g., correlated error terms, cross-loading indicators). Although these modifications would have led to an increase in model fit, they were not consistent with Braithwaite's specification of the theory.

Causality and equivalent models. When constructing a path model and using SEM estimation to assess the relationships between various constructs, theory is relied upon to determine the directionality of the relationships between constructs. This dependence upon theory is necessary because SEM is a statistical causal-modeling approach that often utilizes correlational data (Kline, 2005). The structural paths in the model may suggest causal or temporal relationships between latent constructs, but due to the correlational nature of the data, the graphical sequence of the constructs can be re-arranged without impacting the fit of the model. For example, rather than specifying that the forms of shaming lead to changes in self-esteem or perceived fairness, the fit of the model would remain the same if the model was specified to assess if changes in self-esteem or perceived fairness lead to changes in shaming. Although the fit would remain the same, many of these possible models do not make sense because theory would rule out many illogical patterns of relationships between constructs (e.g., shame-guilt causing self-esteem that causes forms of shaming).

Although claims of causality are somewhat limited due to SEM's reliance on modeling covariation among variables, steps can be taken to strengthen the argument for causal relationships between variables (Kline, 2005; Mulaik, 2000). In the current study, substantive theory was heavily relied upon in order to create and modify the structural model. For example, Harris (2003, 2006) predicted that forms of shaming lead to the experience of different forms of shame emotions that in turn lead to more adaptive or maladaptive outcomes for offenders. Claims of causality are also strengthened if the estimated structural model is supported, not only by theory, but by empirical evidence

from additional studies in that particular research domain. Thus, a form of "pattern matching" is conducted to determine if the results of the present study are in line with empirical findings from other investigations (Shadish et al., 2002). This was indeed the case in the present study. For example, Ahmed and Braithwaite (2004) found that shame acknowledgement was negatively associated with bullying behavior, while shame displacement was positively related to bullying. These findings are in direct accordance with the findings from the present study; shame acknowledgement (shame-guilt) was found to be predictive of more adaptive outcomes, while the opposite finding was supported in association with shame displacement (unresolved shame).

Finally, causal claims can be strengthened if the chosen directionality between constructs in a structural model mimics the actual temporal ordering of the variables as they were measured in time. In the current study, offenders were randomly assigned to receive RST conferencing or court processing. Following this administration of the treatment, interviews were conducted to assess the offender's self-esteem, perceived fairness, and all other constructs in the structural model. The structural paths modeled were drawn to correspond with this temporal sequence; reintegrative and stigmatizing shaming are the exogenous variables in the model, with all other constructs being ultimately predicted by these two constructs. Further, covariation in self-esteem was specified to be dependent upon the shaming constructs, rather than serving as a moderator for the impact of shaming on the other variables in the model. This specification was made because self-esteem was measured following the administration of the treatment.

Along the same lines, the problem of equivalent models must also be addressed. As noted above, SEM utilizes the covariation between variables to estimate causal models, and because the variance-covariance matrix is analyzed, the configuration among paths may be altered but the overall fit of the model would remain the same (Hershberger, 2006; Kline, 2005; MacCallum et al., 1993). With relatively simple models containing few constructs, the number of equivalent models is somewhat small, but more complicated structural models can yield hundreds or thousands of equivalent models. For example, in the present study there were 17 structural paths between nine latent constructs measured by 36 directly observed variables, yielding hundreds of thousands of possible equivalent models. It is recommended that if there are theoretically justified reasons for specifying competing equivalent models, these equivalent models should be estimated (Kline, 2005). In the present study, equivalent models were not tested because of the absence of theoretically driven alternatives, but future research in this area should be conducted with the inclusion of additional theoretical constructs to more clearly elucidate the causal relationships between constructs in RST.

Conclusions

Thousands of crimes are committed annually in the United States by new and repeat criminal offenders. Although a retributive form of justice is currently in place to deter criminal offending, the effectiveness of this court-based system is limited, thus meriting further research into possible alternative forms of justice. The current study sought to test the efficacy and underlying theory of one such form of justice termed RST,

using powerful statistical techniques to assess covariation between constructs. This restorative form of justice focuses on shaming the act and not the person, in order to convey the message that although crime will not be tolerated, committing an offense does not stigmatize an individual as a life-time offender with no chance to amend for their wrong-doing. The findings from the present study build upon the empirical RST literature by first assessing the reliability of the instrument used to assess RST. After determining the acceptability of the assessment instruments used for the measurement of RST constructs, the present study found support for the theoretical underpinnings of RST, as well as for the continuation of the application of RST in response to criminal offending. Future research on the positive outcomes of RST are thus warranted, because individual offenders benefiting from exposure to RST translates into an overall benefit to society.

REFERENCES

Ahmed, E., & Braithwaite, V. (2004). "What, me ashamed?" Shame management and school bullying. *Journal of Research in Crime and Delinquency, 41,* 269-294.

Ahmed, E., & Braithwaite, J. (2005). Shaming, shame, forgiveness and bullying. *Australian and New Zealand Journal of Criminology*, *38*, 298-323.

Ahmed, E., Harris, N., Braithwaite, J., & Braithwaite, V. (2001). *Shame management through reintegration*. Melbourne, Australia: Cambridge University Press.

Anderson, J. C., & Gerbing, D. W. (1982). Some methods for respecifying measurement models to obtain unidimensional construct measurement. *Journal of Marketing Research, 19,* 453-460.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411-423.

Barnes, G. (1999). *Procedural justice in two contexts: Testing the fairness of diversionary conferencing for intoxicated drivers.* Unpublished doctoral dissertation, Department of Criminal Justice and Criminology, University of Maryland, College Park.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. E. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4,* 1-44.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1993). When ego threats lead to self-regulation failure: Negative consequences of high self-esteem. *Journal of Personality and Social Psychology, 64,* 141-156.

Baumeister, R.F., Smart, L., & Boden, J.M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem*. Psychological Review, 103*, 5-33.

Baumer, E. P., Wright, R., Kristinsdottir, K., & Gunnalaugsson, H. (2002). Crime, shame, and recidivism: The case of Iceland. *British Journal of Criminology, 41,* 40-59.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Blackhart, G. C., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: A meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review, 13,* 269-309.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley.

Bonta, J., Wallace-Capretta, S., Rooney, J., & McAnoy, K. (2002). An outcome evaluation of a restorative justice alternative to incarceration. *Contemporary Justice Review, 5,* 310-338.

Botchkovar, E. V., & Tittle, C. R. (2005). Crime, shame, and reintegration in Russia. *Theoretical Criminology, 9,* 401-442.

Bradshaw, W., & Roseborough, D. (2005). Restorative justice dialogue: The impact of mediation and conferencing on juvenile recidivism. *Federal Probation, 69*(2), 15-21.

Braithwaite, J. (1989). *Crime, shame, and reintegration.* Cambridge, UK: Cambridge University Press.

Braithwaite, J. (1999). Restorative justice: Assessing optimistic and pessimistic accounts. In M. Tonry (Ed.), *Crime and justice: A review of research* (pp. 1-127). Chicago, IL: University of Chicago Press.

Braithwaite, J., & Mugford, S. (1994). Conditions of successful reintegration ceremonies: Dealing with juvenile offenders. *British Journal of Criminology, 34,* 139-171.

Braithwaite, J., & Pettit, P. (1990). *Not just desserts: A republican theory of criminal justice.* Oxford, England: Clarendon Press.

Bushman, B.J., & Baumeister, R.F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75*, 219-229.

Byrne, B. M. (1991). The Maslach Burnout Inventory: Validating factorial structure and invariance across intermediate, secondary, and university educators. *Multivariate Behavioral Research*, *26*, 583-605.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Thousand Oaks, CA: Sage.

Byrne, B. M., Shavelson, R. J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456-466.

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25,* 1-27.

Cohen, A. K. (1955). *Delinquent boys: The culture of the gang.* Glencoe, IL: Free Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cooley, C. H. (1922). *Human nature and the social order*. New York, NY: Scribners.

Cox, R. P. (1996). An exploration of the demographic and social correlates of criminal behavior among adolescent males. *Journal of Adolescent Health, 19,* 17-24.

Dearing, R. L., Stuewig, J., & Tangney, J. P. (2005). On the importance of distinguishing shame from guilt: Relations to problematic alcohol and drug use. *Addictive Behaviors, 30,* 1392-1404.

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science, 16,* 328-335.

Durkheim, E. (1951). *Suicide* (J. A. Spaulding & G. Simpson, Trans.), New York, NY: Free Press.

Finney, S. J., & DiStefano, C. (2006). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269-314). Greenwich, CT: Information Age.

Gailliot, M. T., & Baumeister, R. F. (2006). Self-esteem, belongingness, and worldview validation: Does belongingness exert a unique influence upon self-esteem? *Journal of Research in Personality, 41,* 327-345.

Garfinkel, H. (1956). Conditions of successful degradation ceremonies. *The American Journal of Sociology, 61,* 420-424.

Gendreau, P., Irvine, M., & Knight, S. (1973). Evaluating response set styles on the MMPI with prisoners: Faking good adjustment and maladjustment. *Canadian Journal of Behavioural Science*, *5*, 183-194.

Gilbert, D. (2006). *Stumbling upon happiness.* New York, NY: Random House.

Gilbert, P. (1997). The evolution of social attractiveness and its role in shame, humiliation, guilt, and therapy. *British Journal of Medical Psychology, 70,* 113-147.

Gilbert, P., & McGuire, M. T. (1998). Shame, status, and social roles: Psychobiology and evolution. In P. Gilbert (Ed.), *Shame: Interpersonal behavior, psychopathology, and culture* (pp. 99-125). New York, NY: Oxford University Press.

Gilligan, J. (2003). Shame, guilt, and violence. *Social Research, 70,* 1149-1180.

Grabhorn, R., Stenner, H., Stangier, U., & Kaufhold, J. (2006). Social anxiety in anorexia and bulimia nervosa: The mediating role of shame. *Clinical Psychology and Psychotherapy, 13,* 12-19.

Greenberg, L., & Paivio, S. C. (1997). Varieties of shame experience in psychotherapy. *Gestalt Review, 1,* 205-220.

Hancock, G. (1997). Structural equation modeling methods of hypothesis testing of latent variable means. *Measurement and Evaluation in Counseling and Development, 30,* 91-105.

Hare, R. D. (1985). A checklist for the assessment of psychopathy in criminal populations. In M. H. Ben-Aron, S. J. Hucker, & C. D. Webster (Eds.), *Clinical criminology: The assessment and treatment of criminal behavior* (pp. 157-168). Toronto, Canada: M & M Graphics.

Harris, N. (2003). Reassessing the dimensionality of the moral emotions. *British Journal of Psychology, 94,* 457-473.

Harris, N. (2006). Reintegrative shaming, shame, and criminal justice. *Journal of Social Issues, 2,* 327-346.

Harris, N., & Burton, J. B. (1998). Testing the reliability of observational measures of reintegrative shaming at community accountability conferences and at court. *Australian and New Zealand Journal of Criminology, 31,* 230-241.

Harris, N., & Maruna, S. (2006). Shame, shaming, and restorative justice: A critical appraisal. In D. Sullivan & L. Tiff (Eds.), *Handbook of restorative justice* (pp. 452-460). New York, NY: Routledge.

Hay, C. (2001). An exploratory test of Braithwaite's Reintegrative Shaming Theory. *Journal of Research in Crime and Delinquency, 2001,* 132-153.

Hershberger, S. L. (2006). The problem of equivalent structural models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 13-41). Greenwich, CT: Information Age.

Hong, S., Malik, M. L., & Lee, M.-K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-western sample. *Educational and Psychological Measurement, 63,* 636-654.

Houts, S. S. (1996). Explaining alcoholism treatment efficacy with the theory of reintegrative shaming. *Alcoholism Treatment Quarterly, 13*(4), 25-38.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Jöreskog, K. G. (1971a). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409-426.

Jöreskog, K. G. (1971b). Statistical analysis of sets of congeneric tests. *Psychometrika, 36,* 109-133.

Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.8 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Langan, P. A., & Levin, D. J. (2002). Recidivism of prisoners released in 1994. *Bureau of Justice Statistics, Special Report* (NCJ, 193427). Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Latimer, J., Dowden, C., & Muise, D. (2005). The effectiveness of restorative justice practices: A meta-analysis. *Prison Journal, 85*(2), 127-144.

Latimer, J., & Kleinknecht, S. (2000). The effects of restorative justice programming: A review of the empirical research. Ottawa, Ontario: Department of Justice Canada, Research and Statistics Division.

Leary, M. R. (2005). Sociometer theory and the pursuit of relational value: Getting to the root of self-esteem. *European Review of Social Psychology, 16,* 75-111.

Leary, M. R. (2007). How the self became involved in affective experience. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 38-52). New York, NY: Guilford Press.

Leary, M. R., & Downs, D. L. (1995). Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer. In M. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 123-144). New York, NY: Plenum.

Leary, M. R., Schreindorfer, L. S., & Haupt, A. L. (2003). The role of low self-esteem in emotional and behavioral problems: Why is low self-esteem dysfunctional? In R. M. Kowalski (Ed.), *The interface of social and clinical psychology: Key readings* (pp. 116-128). New York, NY: Psychology Press.

Lewis, H. B. (1971). *Shame and guilt in neurosis.* New York, NY: International Universities Press.

Lindsay-Hartz, J. (1984). Contrasting experiences of shame and guilt. *American Behavioral Scientist, 27,* 689-704.

Lipsey, M. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works: Reduction reoffending – guidelines from research and practice* (pp. 63-78). New York, NY: Wiley.

Llewellyn, J., & Howse, R. (1998). *Restorative justice: A conceptual framework*. Ottawa, Ontario: Law Commission of Canada.

Losoncz, I., & Tyson, G. (2007). Parental shaming and adolescent delinquency: A partial test of Reintegrative Shaming Theory. *The Australian and New Zealand Journal of Criminology, 40,* 161-178.

Lu, H. (1999). Bang Jiao and reintegrative shaming in China's urban neighborhoods. *International Journal of Comparative and Applied Justice*, 23, 115-125.

Luyten, P., Fontaine, R. J., & Corveleyn, J. (2002). Does the test of self-conscious affect (TOSCA) measure maladaptive aspects of guilt and adaptive aspects of shame? An empirical investigation. *Personality and Individual Differences, 33,* 1373-1387.

Lynd, H. (1958). *On shame and the search for identity.* New York, NY: Academic Press.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490-504.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114,* 185-199.

Makkai, T., & Braithwaite, J. (1994). Reintegrative shaming and compliance with regulatory standards. *Criminology, 32,* 361-381.

Marsh, H. W., Hau, K.T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers of overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling, 11,* 320-341.

Marshall, T. F. (1996). The evolution of restorative justice in Britain. *European Journal on Criminal Policy and Research, 4*, 21-43.

Marshall, T. (1999). *Restorative justice: An overview*. London, England: Home Office, Research Development and Statistics Directorate.

Maxwell, G., & Morris, A. (1993). *Family, victims, and culture: Youth justice in New Zealand* . Wellington, NZ: Social Policy Agency and Institute of Criminology, Victoria University of Wellington.

Maxwell, G., & Morris, A. (2002). The role of shame, guilt, and remorse in restorative justice processes for young people. In E. Weitekamp & H.-J. Kerner (Eds.), *Restorative justice: Theoretical foundations* (pp. 267-284). Cullompton, UK: Willan.

McAlinden, A. M. (2005). The use of "shame" with sexual offenders. *British Journal of Criminology, 2005,* 373-394.

McGee, R., & Williams, S. (2000). Does low self-esteem predict health compromising behaviors among adolescents? *Journal of Adolescence, 23,* 569-582.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.

Merton, R. K. (1957). *Social theory and social structure,* Glencoe, IL: Free Press.

Miethe, T. D., Lu, H., & Reese, E. (2000). Reintegrative shaming and recidivism risks in drug court: Explanations for some unexpected findings. *Crime Delinquency, 46,* 522-541.

Miller, W. B. (1958). Lower-class culture as a generating milieu of gang delinquency. *Journal of Social Issues, 14,* 5-9.

Moore, D. B., & McDonald, J. M. (2000). *Transforming conflict.* Sydney, Australia: Transformative Justice Australia PTY LTD.

Mulaik, S. A., (2000). Objectivity and other metaphors of structural equation modeling. In R. Cudeck, S. Du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 59-78). Lincolnwood, IL: Scientific Software International.

Murphy, K., & Harris, N. (2007). A test of reintegrative shaming theory in the white-collar crime context. *British Journal of Criminology, 47,* 900-917.

Nisbett, E. R., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

Nugent, W., Williams, M., & Umbreit, M. (2004). Participation in victim-offender mediation and the prevalence of subsequent delinquent behavior: A meta-analysis. *Research on Social Work Practice*, *14*(6), 408-416.

Osberg, T. M., & Shrauger, J. S. (1990). The role of self-prediction in psychological assessment. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 8, pp. 97-120). Hillsdale, NJ: Erlbaum.

Oser, C. B. (2006). The criminal offending- self-esteem nexus: Which version of the self-esteem theory is supported? *The Prison Journal, 86,* 344-363.

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

Reckless, W. C. (1967). *The crime problem* (4th ed.). New York, NY: Appleton-Century-Crofts.

Retzinger, S. M., & Scheff, T. J. (1996). *Strategy for community conferences: Emotions and social bonds.* In B. Galaway & J. Hudson (Eds.), *Restorative justice: International perspectives* (pp. 315-336). Monsey, NY: Criminal Justice Press/Willow Tree Press.

Robins, R. W., Noftle, E. E., & Tracy, J. L. (2007). Assessing self-conscious emotions: A review of self-report and nonverbal measures. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 443-467). New York, NY: Guilford Press.

Rosenberg, M. (1979). *Conceiving the self*. New York, NY: Basic Books.

Sabini, J., & Silver, M. (1997). In defense of shame: Shame in the context of guilt and

embarrassment. *Journal for the Theory of Social Behaviour, 27*(1), 1-15.

Scheff, T. J. (1988). Shame and conformity: The deference-emotion system. *American Sociological Review, 53,* 395-406.

Scheff, T. J. (1990). Socialization of emotions: Pride and shame as causal agents. In T. D. Kemper (Ed.), *Research agendas in the sociology of cmotions* (pp. 281-304). Albany, NY: State University of New York Press.

Schretlen, D., & Arkowitz, H. (1990).A psychological test battery to detect prison inmates who fake insanity or mental retardation. *Behavioral Sciences and the Law*, *8*, 75-84.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton-Mifflin.

Sherman, L. W., Braithwaite, J., Strang, H., & Barnes, G. C. (2000). Reintegrative shaming experiments (RISE) in Australia, 1995-1999 [Computer file]. ICPSR version. College Park, MD: University of Maryland [producer]. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research [distributor], 2001.

Simpson, J. A., & Weiner, E. S. C. (Eds.). (1989). *The Oxford English dictionary* (2nd ed.). Oxford, England: Oxford University Press.

Sommer, K. L., & Baumeister, R. F. (2002). Self-evaluation, persistence, and performance following implicit rejection: The role of trait self-esteem. *Personality and Social Psychology Bulletin, 28,* 926-938.

Sorbom, D. (1989). Model modification. *Psychometrika, 54,* 371-384.

SPSS Inc. (2008). SPSS for windows: Version 17. SPSS, Chicago, IL.

Stoeber, J., Harris, R. A., & Moon, P. S. (2006). Perfectionism and the experience of pride, shame, and guilt: Comparing healthy perfectionists, unhealthy perfectionists, and non-perfectionists. *Personality and Individual Differences, 43,* 131-141.

Sutherland, E. H., & Cressey, D. R. (1978). *Criminology* (10th ed.). New York, NY: Lippincott.

Tangney, J. P., & Dearing, R. L. (2002). *Shame and guilt.* New York, NY: Guilford Press.

Tangney, J. P., Wagner, P. E., & Gramzow, R. (1989). *The test of self-conscious affect (TOSCA).* Fairfax, VA*:* George Mason University.

Tannenbaum, F. (1938). *Crime and the community,* New York, NY: Columbia University Press.

Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent variable means. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 119-169). Greenwich, CT: Information Age.

Tittle, C. R., Bratton, J., & Gertz, M. G. (2003). A test of a micro-level application of shaming theory. *Social Problems, 50,* 592-617.

Tomkins, S. S. (1963). *Affect/imagery/consciousness. The negative affects* (Vol. 2.). New York, NY: Springer.

Tomkins, S. S. (1987). Shame. In D. L. Nathanson (Ed.), *The many faces of shame* (pp. 133-161). New York, NY: Guilford Press.

Trzesniewski, K. H., Moffitt, T. E., Poulton, R., Donnellan, M. B., Robins, R. W., & Caspi, A. (2006). Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. *Developmental Psychology, 42,* 381-390.

Ttofi, M. M., & Farrington, D. P. (2008). Reintegrative shaming theory, moral emotions, and bullying. *Aggressive Behavior, 34,* 352-368.

Tyler, T. R., Sherman, L., Strang, H., Barnes, G. C., & Woods, D. (2007). Reintegrative shaming, procedural justice, and recidivism: The engagement of offenders' psychological mechanisms in the Canberra RISE drinking-and-driving experiment. *Law & Society Review, 41,* 553-584.

Umbreit, M. S., Coates, R. B., & Vos, B. (2002). The impact of restorative justice conferencing: A multi-national perspective. *The British Journal of Community Justice, 2,* 21-48.

Vagg, J. (1998). Delinquency and shame: Data from Hong Kong. *British Journal of Criminology, 38*, 247–64.

Van Ness, D., & Strong, K. H. (1997). *Restoring justice.* Cincinnati, OH: Anderson.

Van Stokkom, B. (2002). Moral emotions in restorative justice conferences: Managing shame, designing empathy. *Theoretical Criminology, 6*(3), 339-360.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281-324). Washington, DC: American Psychological Association.

Young, J. W. (1971). The function of theory in a dilemma of path analysis. *Journal of Applied Psychology, 62,* 108-110.

Zehr, H. (1985). *Retributive justice, restorative justice*. Elkhart, IN: Mennonite Central Committee US Office of Criminal Justice.

Zehr, H. (1990). *Changing lenses: A new focus for crime and justice*. Scottsdale, PA: Herald Press.

Zhang, S. X. (1995). Measuring shame in an ethnic context. *British Journal of Criminology, 35,* 248-262.

Zhang, L., & Zhang, S. (2004). Reintegrative shaming and predatory delinquency. *Journal of Research in Crime and Delinquency, 2004,* 433-453.

APPENDICES

Appendix A:

Covariance Matrix for Offenders Assigned to Conference

Covariance Matrix for Offenders Assigned to Conference

	AR1	FAR2	FAR3	FAR4	FAR5	SS1
FAR1	0.56					
AR2	0.31	0.54				
FAR3	0.36	0.32	0.96			
FAR4	0.28	0.31	0.51	0.87		
FAR5	0.24	0.17	0.28	0.22	0.48	
SS1	-0.13	-0.08	-0.20	-0.09	-0.15	0.68
SS2	-0.11	-0.13	-0.19	-0.10	-0.17	0.19
SS3	-0.16	-0.10	-0.21	-0.15	-0.18	0.25
SS4	-0.08	-0.05	-0.17	-0.11	-0.14	0.22
RS1	0.01	0.01	0.13	0.09	0.05	-0.05
RS2	0.11	0.10	0.15	0.14	0.06	-0.04
RS3	0.06	0.04	0.09	0.07	0.07	-0.11
RS4	0.01	0.03	0.06	0.06	0.05	-0.04
RS5	0.09	0.08	0.06	0.04	0.11	-0.03
SEE1	0.06	0.10	0.02	0.06	0.05	0.03
SEE2	0.02	0.03	0.02	-0.01	-0.01	-0.02
SEE3	0.02	0.05	0.04	0.02	0.00	-0.07
SEE4	0.02	0.02	-0.01	0.02	0.01	-0.02
SEE5	0.04	0.03	0.05	0.08	0.04	-0.06
FS1	0.11	0.06	0.12	0.14	0.12	-0.05
FS2	0.09	0.07	0.10	0.13	0.11	-0.06
FS3	0.14	0.06	0.14	0.13	0.11	-0.08
RES1	0.09	0.09	0.05	0.13	0.00	0.09
RES2	0.17	0.14	0.14	0.29	0.00	0.08
RES	0.15	0.14	0.16	0.19	0.08	0.08
SG1	0.09	0.07	0.05	0.12	0.02	-0.04
SG2	0.13	0.11	0.14	0.24	0.10	-0.01
SG3	0.15	0.09	0.09	0.20	0.09	-0.03
SG4	0.15	0.08	0.16	0.22	0.06	-0.06
SG5	0.18	0.07	0.15	0.20	0.13	0.00
E1	0.09	0.07	0.04	0.12	-0.01	0.08
E2	-0.01	0.02	0.02	0.03	-0.09	0.08
E3	-0.05	-0.05	-0.09	0.04	-0.06	0.05
E4	-0.02	0.01	-0.03	0.07	-0.08	0.11
U1	-0.11	-0.14	-0.19	-0.17	-0.16	0.14
U2	-0.07	-0.12	-0.17	-0.10	-0.09	0.12

(*table continues*)

	SS2	SS3	SS4	RS1	RS2	RS3
SS2	0.63					
SS3	0.19	0.71				
SS4	0.15	0.37	0.60			
RS1	-0.11	0.04	-0.01	0.98		
RS2	-0.07	0.01	0.03	0.34	1.03	
RS3	-0.07	0.01	-0.05	0.25	0.21	0.63
RS4	-0.06	0.07	0.03	0.48	0.39	0.31
RS5	-0.05	-0.02	-0.05	0.30	0.39	0.25

	SS2	SS3	SS4	RS1	RS2	RS3
SEE1	-0.02	-0.05	-0.01	0.04	0.03	0.04
SEE2	0.00	0.00	-0.01	0.05	0.01	0.02
SEE3	-0.03	-0.01	-0.03	0.01	-0.01	0.03
SEE4	0.01	0.01	-0.04	0.00	-0.03	-0.01
SEE5	-0.04	-0.03	-0.05	0.04	0.04	0.06
FS1	-0.09	-0.07	-0.05	0.12	0.17	0.27
FS2	-0.06	-0.10	-0.08	0.15	0.15	0.23
FS3	-0.09	-0.11	-0.05	0.13	0.18	0.27
RES1	-0.09	-0.11	-0.03	0.11	0.19	0.00
RES2	-0.10	-0.14	-0.14	0.31	0.30	0.21
RES3	-0.09	-0.14	-0.04	0.09	0.18	-0.01
SG1	-0.03	0.02	-0.01	0.11	0.14	0.13
SG2	-0.06	0.00	0.03	0.15	0.29	0.23
SG3	-0.07	0.00	0.06	0.17	0.25	0.21
SG4	-0.03	0.00	0.06	0.22	0.35	0.25
SG5	-0.10	-0.11	-0.01	0.22	0.21	0.29
E1	0.01	0.10	0.12	0.08	0.26	0.16
E2	-0.01	0.12	0.17	0.08	0.24	0.12
E3	0.10	0.11	0.08	0.01	0.11	0.07
E4	0.09	0.12	0.18	0.12	0.20	0.11
U1	0.18	0.19	0.14	-0.03	-0.06	-0.05
U2	0.10	0.12	0.12	-0.07	-0.03	-0.04

(*table continues*)

	RS4	RS5	SEE1	SEE2	SEE3	SEE4
RS4	1.30					
RS5	0.43	0.97				
SEE1	0.07	0.09	0.32			
SEE2	0.02	0.04	0.19	0.52		
SEE3	0.01	-0.02	0.15	0.19	0.50	
SEE4	0.00	0.02	0.10	0.07	0.13	0.31
SEE5	0.04	0.02	0.20	0.16	0.28	0.12
FS1	0.22	0.17	0.07	0.00	0.06	-0.02
FS2	0.18	0.19	0.05	0.01	0.04	-0.02
FS3	0.19	0.21	0.05	-0.01	0.05	-0.01
RES1	0.02	0.16	0.10	0.06	0.08	0.05
RES2	0.11	0.32	0.07	0.03	0.03	-0.05
RES3	0.12	0.14	0.15	0.13	0.07	-0.01
SG1	0.15	0.12	0.04	0.01	0.00	0.03
SG2	0.23	0.20	0.01	-0.14	-0.13	-0.01
SG3	0.18	0.21	0.05	-0.04	-0.07	0.04
SG4	0.27	0.28	-0.02	-0.14	-0.17	-0.03
SG5	0.21	0.28	-0.01	-0.10	-0.16	0.00
E1	0.25	0.21	-0.07	-0.17	-0.16	-0.03
E2	0.18	0.08	-0.10	-0.15	-0.17	-0.05
E3	0.06	0.06	-0.06	-0.08	-0.06	-0.03
E4	0.10	0.11	-0.08	-0.10	-0.13	-0.03
U1	0.00	-0.05	-0.04	-0.02	-0.03	0.01
U2	-0.04	-0.05	-0.05	-0.06	-0.09	-0.03

	SEE5	FS1	FS2	FS3	RES1	RES2
SEE5	0.50					
FS1	0.07	1.01				
FS2	0.07	0.56	0.87			
FS3	0.08	0.64	0.59	0.88		
RES1	-0.02	0.09	0.05	0.03	2.64	
RES2	-0.06	0.13	0.18	0.08	2.03	3.82
RES3	0.01	0.02	-0.04	-0.01	1.62	1.28
SG1	-0.02	0.08	0.09	0.11	0.30	0.32
SG2	-0.06	0.20	0.20	0.23	0.31	0.40
SG3	-0.05	0.19	0.18	0.24	0.44	0.53
SG4	-0.16	0.24	0.24	0.29	0.47	0.57
SG5	-0.09	0.32	0.34	0.39	0.21	0.25
E1	-0.16	0.20	0.19	0.28	0.37	0.56
E2	-0.17	0.17	0.18	0.23	0.19	0.36
E3	-0.06	0.04	0.11	0.11	-0.02	0.03
E4	-0.12	0.02	0.08	0.08	0.10	0.28

(*table continues*)

U1	-0.05	-0.08	-0.06	-0.07	-0.07	-0.10
U2	-0.08	-0.01	-0.04	0.01	-0.11	-0.18

	RES3	**SG1**	**SG2**	**SG3**	**SG4**	**SG5**
RES3	1.95					
SG1	0.16	0.55				
SG2	0.19	0.35	1.30			
SG3	0.26	0.39	0.74	1.04		
SG4	0.30	0.40	0.91	1.02	1.52	
SG5	0.06	0.32	0.60	0.65	0.79	1.47
E1	0.16	0.27	0.63	0.57	0.76	0.52
E2	0.09	0.16	0.41	0.42	0.65	0.39
E3	-0.09	0.07	0.16	0.11	0.23	0.18
E4	-0.03	0.13	0.31	0.38	0.57	0.32

	E1	**E2**	**E3**	**E4**	**U1**	**U2**
U1	-0.07	-0.03	-0.01	-0.06	0.01	-0.01
U2	-0.12	-0.03	0.05	0.02	0.11	0.05
E1	1.24					
E2	0.83	1.48				
E3	0.20	0.24	0.71			
E4	0.55	0.75	0.40	1.16		
U1	0.04	0.06	0.09	0.10	0.41	
U2	0.09	0.07	0.11	0.09	0.19	0.55

Note. FAR = Perceived Fairness, SS = Stigmatization, RS = Reintegration, SEE = Self-Esteem, FS = Family Support, RES = Offender Responsibility, SG = Shame-Guilt, E = Embarrassment-Exposure, and U = Unresolved Shame.

Appendix B:

Covariance Matrix for Offenders Assigned to Court

Covariance Matrix for Offenders Assigned to Court

	FAR1	FAR2	FAR3	FAR4	FAR5	SS1
FAR1	1.12					
FAR2	0.71	1.07				
FAR3	0.61	0.51	1.26			
FAR4	0.63	0.53	0.71	1.38		
FAR5	0.42	0.46	0.51	0.54	0.87	
SS1	-0.28	-0.30	-0.27	-0.28	-0.33	0.80
SS2	-0.37	-0.39	-0.39	-0.39	-0.36	0.39
SS3	-0.36	-0.36	-0.39	-0.33	-0.39	0.43
SS4	-0.34	-0.39	-0.35	-0.30	-0.35	0.33
RS1	0.03	0.10	0.18	0.06	0.15	-0.05
RS2	0.03	0.15	0.20	0.15	0.18	-0.13
RS3	0.13	0.20	0.28	0.30	0.25	-0.22
RS4	-0.01	0.08	0.10	-0.02	0.05	-0.03
RS5	0.10	0.09	0.23	0.18	0.23	-0.08
SEE1	0.02	0.04	-0.02	-0.04	0.02	-0.05
SEE2	0.08	0.04	0.06	0.00	0.03	-0.03
SEE3	0.05	0.09	0.07	0.01	0.02	-0.04
SEE4	0.03	0.04	-0.03	-0.04	0.00	0.02
SEE5	0.07	0.08	0.01	0.00	0.03	-0.03
FS1	0.14	0.17	0.15	0.04	0.09	-0.12
FS2	0.13	0.14	0.22	0.12	0.15	-0.05
FS3	0.02	0.07	0.12	-0.03	0.08	-0.05
RES1	0.00	-0.05	-0.09	-0.03	-0.09	-0.07
RES2	0.01	-0.05	0.06	0.18	-0.03	-0.03
RES3	0.05	0.02	0.03	0.05	0.04	-0.20
SG1	0.04	0.05	0.02	0.04	0.02	-0.02
SG2	-0.12	-0.06	-0.12	-0.10	-0.06	0.04
SG3	0.00	0.04	0.03	0.00	0.03	-0.04
SG4	-0.03	0.03	0.00	-0.04	0.02	-0.06
SG5	0.03	0.02	0.07	0.06	0.08	-0.05
E1	-0.10	-0.03	-0.11	-0.02	-0.06	0.04
E2	-0.11	-0.17	-0.09	-0.06	-0.09	0.13
E3	-0.31	-0.30	-0.24	-0.21	-0.20	0.19
E4	-0.23	-0.16	-0.17	-0.13	-0.14	0.16
U1	-0.43	-0.48	-0.50	-0.52	-0.40	0.41
U2	-0.21	-0.34	-0.26	-0.23	-0.20	0.27

(*table continues*)

	SS2	SS3	SS4	RS1	RS2	RS3
SS2	0.78					
SS3	0.43	0.94				
SS4	0.35	0.57	0.75			
RS1	0.01	0.00	-0.05	1.07		
RS2	-0.05	-0.10	-0.06	0.46	1.27	
RS3	-0.16	-0.14	-0.14	0.50	0.47	1.31
RS4	-0.01	0.05	0.09	0.58	0.46	0.50
RS5	-0.05	-0.09	-0.02	0.40	0.58	0.45

	SS2	SS3	SS4	RS1	RS2	
SEE1	-0.02	-0.01	-0.04	0.04	0.05	-0.03
SEE2	-0.04	-0.01	-0.04	0.00	-0.03	-0.05
SEE3	-0.04	-0.08	-0.07	0.00	-0.04	-0.02
SEE4	0.03	0.04	-0.01	0.02	0.03	-0.01
SEE5	-0.02	-0.07	-0.09	0.01	-0.06	-0.01
FS1	-0.08	-0.11	-0.09	0.07	0.04	0.16
FS2	-0.07	-0.12	-0.11	0.25	0.17	0.34
FS3	-0.04	-0.05	-0.07	0.23	0.20	0.23
RES1	-0.17	-0.23	-0.10	-0.16	0.09	-0.08
RES2	-0.01	-0.06	-0.10	0.10	0.22	0.31
RES3	-0.08	-0.21	-0.18	-0.14	0.06	-0.18
SG1	0.00	0.02	0.03	0.10	0.14	0.20
SG2	0.13	0.19	0.09	0.17	0.17	0.21
SG3	0.09	0.10	0.10	0.27	0.23	0.30
SG4	0.08	0.11	0.06	0.24	0.27	0.37
SG5	0.03	0.06	0.07	0.29	0.15	0.41
E1	0.14	0.12	0.10	0.22	0.17	0.35
E2	0.18	0.20	0.17	0.17	0.14	0.19
E3	0.27	0.33	0.28	0.01	-0.04	0.02
E4	0.25	0.32	0.20	0.23	0.09	0.20
U1	0.47	0.46	0.42	-0.01	-0.18	-0.21
U2	0.25	0.32	0.26	-0.01	-0.09	-0.07

(*table continues*)

	RS4	RS5	SEE1	SEE2	SEE3	SEE4
RS4	1.55					
RS5	0.52	1.11				
SEE1	0.09	0.01	0.35			
SEE2	0.00	0.04	0.16	0.52		
SEE3	-0.01	0.03	0.13	0.22	0.53	
SEE4	-0.02	0.02	0.09	0.13	0.14	0.33
SEE5	0.02	0.01	0.16	0.22	0.31	0.16
FS1	0.15	0.09	0.01	-0.01	0.06	0.01
FS2	0.24	0.20	0.04	0.02	0.04	0.01
FS3	0.24	0.21	0.09	0.03	0.07	0.05
RES1	-0.31	-0.02	0.01	0.02	0.09	0.07
RES2	-0.04	0.17	-0.05	-0.13	0.03	0.04
RES3	-0.28	0.08	0.03	0.03	-0.01	0.03
SG1	0.10	0.06	0.02	0.02	0.00	0.06
SG2	0.12	0.13	0.01	-0.05	-0.08	0.07
SG3	0.19	0.18	0.00	-0.03	-0.05	0.08
SG4	0.19	0.16	0.00	-0.06	-0.03	0.09
SG5	0.29	0.23	-0.03	0.00	-0.02	0.02
E1	0.15	0.16	-0.02	-0.16	-0.09	0.01
E2	0.10	0.12	-0.05	-0.12	-0.11	0.00
E3	0.06	-0.02	-0.04	-0.14	-0.15	-0.03
E4	0.07	0.05	-0.05	-0.14	-0.15	0.03
U1	0.00	-0.08	0.00	-0.06	-0.04	0.04
U2	-0.03	-0.08	-0.02	-0.09	-0.05	0.01

	SEE5	FS1	FS2	FS3	RES1	RES2
SEE5	0.54					
FS1	0.00	0.91				
FS2	0.01	0.50	0.84			
FS3	0.02	0.45	0.52	0.82		
RES1	0.02	-0.06	0.00	0.04	3.60	
RES2	-0.07	0.02	0.22	0.08	2.24	4.97
RES3	-0.03	0.08	0.11	0.16	2.52	1.80
SG1	0.05	0.05	0.08	0.11	0.05	0.29
SG2	0.01	0.01	0.05	0.11	-0.08	0.14
SG3	0.03	0.09	0.15	0.20	0.07	0.29
SG4	0.00	0.16	0.18	0.27	0.09	0.38
SG5	-0.02	0.19	0.29	0.30	-0.11	0.07
E1	-0.08	0.11	0.18	0.16	0.23	0.44
E2	-0.10	0.03	0.10	0.12	0.18	0.31
E3	-0.13	-0.02	0.01	0.06	-0.18	-0.13
E4	-0.11	0.08	0.07	0.10	-0.02	0.21
U1	-0.02	-0.07	-0.05	0.01	-0.29	-0.20
U2	-0.05	-0.04	-0.04	0.02	-0.20	-0.11

(*table continues*)

	RES3	SG1	SG2	SG3	SG4	SG5
RES3	3.20					
SG1	-0.03	0.86				
SG2	-0.09	0.58	1.39			
SG3	-0.02	0.69	0.92	1.36		
SG4	-0.06	0.68	1.00	1.24	1.62	
SG5	-0.17	0.34	0.52	0.62	0.68	1.18
E1	0.21	0.40	0.85	0.81	0.99	0.50
E2	0.17	0.38	0.80	0.79	0.98	0.50
E3	-0.05	0.08	0.31	0.26	0.35	0.27
E4	-0.04	0.37	0.69	0.68	0.84	0.43
U1	-0.24	-0.04	0.17	0.10	0.10	0.11
U2	-0.18	0.03	0.13	0.12	0.15	0.22

	E1	E2	E3	E4	U1	U2
E1	1.54					
E2	1.24	1.80				
E3	0.59	0.71	1.25			
E4	1.02	1.15	0.76	1.69		
U1	0.19	0.18	0.34	0.34	1.03	
U2	0.22	0.20	0.31	0.32	0.54	0.91

Note. FAR = Perceived Fairness, SS = Stigmatization, RS = Reintegration, SEE = Self-Esteem, FS = Family Support, RES = Offender Responsibility, SG = Shame-Guilt, E = Embarrassment-Exposure, and U = Unresolved Shame.

CURRICULUM VITAE

Elizabeth Dansie

Department of Psychology
Utah State University, 2810 Old Main Hill
Logan, Utah 84322-2810

527 Eastlake Ave E, Apt 7
Seattle, Washington, 98109
Tel: (435) 650-0805
Email: elizabeth.dansie@gmail.com

EDUCATION

2009 - Present	Doctoral Candidate (Graduation: December 2010) Experimental and Applied Psychological Science Program: Research and Evaluation Methodology Emphasis Utah State University, Logan, Utah
2006-2009	M.S. Psychology Utah State University, Logan, Utah. Thesis Title: *An Empirical Investigation of the Adaptive Nature of Shame*
2002-2006	Bachelor's of Science (Summa Cum Laude). Psychology (Biology minor). Southern Utah University, Cedar City, Utah.
2002	East Carbon High School

RESEARCH/EMPLOYMENT EXPERIENCE

July 2010 – Present	Research Scientist II, University of Washington: Center for Clinical and Epidemiological Research
January 2010 – June 2010	Research Associate, Purdue University: Military Family Research Institute

Summer 2009-Spring 2009	Data analyst and program evaluator, National Center for Engineering and Technology Education, Utah State University
Summer 2009	Data analyst, National Substitute Teaching Institute, Logan, Utah
Spring 2008-Fall 2009	Statistical consultant, Office of Methodological and Data Sciences, Utah State University Provide statistical consulting services for student/faculty research projects and theses/dissertations.
Fall 2008-Fall 2009	Graduate Research Assistant, Department of Psychology, Utah State University. Collect and analyze data for the Multi-Sensory Cognition Lab.
Fall 2005-Spring 2006	Research Assistant, Department of Psychology, Southern Utah University Assisted with collection, analysis, and presentation of data for research on self-conscious emotions.
Fall 2004-Spring 2005	Research Assistant, Department of Psychology, Southern Utah University Tested participants and assisted with data analysis for research study on parietal lobe functioning.

TEACHING EXPERIENCE

Fall 2010	Online Instructor, Psychology 3500, Research Methods Utah State University
Fall 2009	Instructor, Psychology 3500, Research Methods Utah State University
Spring 2009	Instructor, Psychology 2800, Psychological Statistics, Utah State University

Fall 2006-Spring 2008	Teaching Assistant, Department of Psychology, Utah State University Tutored students in Introductory Psychology, taught class when instructor was absent, taught lab sections, and grader.
Fall 2005-Spring 2006	Teaching Assistant, Department of Psychology, Southern Utah University Tutored students in Introductory Psychology. Taught Psychology of Gender, Introductory Psychology, Cognitive Psychology, and Motivation and Emotion when instructor was absent.
Spring 2004-Fall 2004	Teaching Assistant, Department of Psychology, Southern Utah University Tutored students in Experimental Analysis of Behavior and assisted with final grade evaluation.
Summer 2004-Fall 2004	Teaching Assistant, Department of Psychology, Southern Utah University Tutored students in Statistics of Psychology and grader.

PUBLICATIONS

Dansie, E. J., & Fargo, J. D. (2009). Individual and community predictors of fear of criminal victimization: Results from a national sample of urban US citizens. *Journal of Crime Prevention and Community Safety, 11,* 124-140.

Dansie, E. J., Ferguson, T. J., & Elison, J. (in preparation). The shame display in interpersonal encounters: Saying "sorry" is not quite enough. *European Journal of Social Psychology.*

Elison, J., & Dansie, E. J. (in press). Humiliation. Invited essay for *The Encyclopedia of Adolescence.*

Elison, J. & Dansie, E. J., Ferguson, T. J. (in review). More perils of partialing: The case of the TOSCA. *Multivariate Behavioral Research.*

PROFESSIONAL PRESENTATIONS

Dansie, E. J., & Fargo, J. D. (2010, February). *Individual and community predictors of fear of criminal victimization.* Poster presented at the Academy of Criminal Justice Sciences conference, San Diego, California.

Dansie, E. J., Drollette, E., Frye, A., & Jordan, K. (2009). *Multisensory stimuli enhance infant numerical representation only when present during both learning and testing*. Poster presented at the Jean Piaget Society Conference, Park City, Utah.

Dansie, E. J., & Fargo, J. D. (2009, April). *Individual and community predictors of fear of criminal victimization.* Paper presented at the Graduate Student Research Symposium, Utah State University.

Dansie, E. J. (2008, April). *The shame display in interpersonal encounters: Saying 'sorry' is not quite enough.* Paper presented at the Graduate Student Research Symposium, Utah State University.

Dansie, E. J. (2008, April). *The shame display in interpersonal encounters: Saying 'sorry' is not quite enough.* Paper presented at the Rocky Mountain Psychological Association convention, Boise, ID.

Dansie, E. J. (2008, April). *How I deceive thee, let me count the ways.* Paper presented at the Rocky Mountain Psychological Association convention, Boise, ID.

Elison, J., Dansie, E. J., Ferguson, T. J., & Webb, R. (2008, April). *More perils of partialing.* Paper presented at the Rocky Mountain Psychological Association convention, Boise, ID.

Elison, J., Nowill, J., Corser, G., & Dansie, E. J (2008, April). *Revision of the shame and guilt scale.* Paper presented at the Rocky Mountain Psychological Association convention, Boise, ID.

Elison, J., Ferguson, T. J., Dansie, E. J., & Haycock, M. (2007, April). *Perils of partialing.* Paper presented at the Rocky Mountain Psychological Association convention, Denver, CO.

Elison, J., & Dansie, E. (2007, April). *An MDS analysis of shame, guilt, and empathy.* Paper presented at the Rocky Mountain Psychological Association convention, Denver, CO.

Dansie, E. (2007, February). *Investigating the predictive validity of the Compass of Shame Scale*. Paper presented at the Utah Conference for Undergraduate Research, University of Utah, UT.

Dansie, E. (2006, April). *Investigating the predictive validity of the Compass of Shame Scale.* Paper presented at the Rocky Mountain Psychological Association convention, Park City, UT.

Elison, J., Pulos, S., & Dansie, E. (2006, April). *Factor structure of the Compass of Shame Scale.* Paper presented at the Rocky Mountain Psychological Association convention, Park City, UT.

Elison, J., & Dansie, E. (2006, April). *Rejection sensitive: What's "normal"?* Paper presented at the Rocky Mountain Psychological Association convention, Park City, UT.

Elison, J., & Dansie, E. (2006, April). *Self-conscious emotions: Reasons versus experience*. Paper presented at the Rocky Mountain Psychological Association convention, Park City, UT.

Dansie, E., Rott, J., Rawlins, M., Miller, M., Crowell, K., & Kynaston, A. (2006). *Shame and guilt: recent research*. Symposium at the Rocky Mountain Psychological Association convention, Park City, UT.

Dansie, E., Jones, M., Gurr, M., & Juntunen, S. (2005, April). *Relationship of functionality in the severe and persistently mentally ill with attendance at Oasis Clubhouse*. Paper presented at the Rocky Mountain Psychological Association convention, Phoenix, AZ.

HONORS AND AWARDS

2011 Post-Doctoral Fellowship, University of Washington
2010 Research featured in *Research Matters,* annual publication from Utah State University.
2009 1st Place Best Paper Graduate Student Research Symposium ($150)
2009 Borg Scholarship Recipient (Utah State University Psychology Department Research Productivity Award; $2,500)
2008 Invited presentation, Southern Utah University. *An empirical investigation of the adaptive nature of shame.*
2008 Psi Chi Best Paper Award Winner at the Rocky Mountain Psychological Association Conference ($300)

2007 Selected to present paper for Utah Conference for Undergraduate Research
2006 Humanities and Social Sciences Student Scholar of the Year
2006 Southern Utah University Psychology Department Research Day Best Paper
2005 Rondthaler Award (Awarded for highest departmental G.P.A.; $3,000)
2005 McCray-Cloward Scholarship (Scholarly Activity; $500)
2004 Southern Utah University Academic Scholarship ($1,500)
2003 Southern Utah University Academic Scholarship ($1,500)
2003 AmeriCorps Scholarship ($1,000)
2002 Graduated as Valedictorian from East Carbon High School

CURRENT RESEARCH

-Using multiple group structural equation modeling to compare various constructs of Reintegrative Shaming Theory.
-Using confirmatory factor analysis and structural equation modeling to determine which factors are predictive of fear of criminal victimization using a nationally representative sample of U.S. citizens.
-Investigating the implications of misusing the statistical technique partialing in research on shame and guilt.

PROFESSIONAL DEVELOPMENT

Affiliations
Member, Psi Chi Honor Society
Member, Alpha Chi Honor Society

Statistical Training
ANOVA
Regression
Multivariate Methods
Factor Analysis
Structural Equation Modeling
Bivariate Twin Analysis

Software Programs
SPSS, M-Plus, MSAccess, Lisrel, AMOS, Excel, SAS

Service

2006 Organized undergraduate research symposium on shame and guilt for Rocky Mountain Psychological Association conference.
2004 Psychology Club Vice President
2003 Volunteered at Oasis Clubhouse to assist with the severe and persistently mentally ill (450 hours).

REFERENCES

Jamison D. Fargo, PhD, MS.Epi
Center for Clinical Epidemiology & Biostatistics
University of Pennsylvania School of Medicine
Philadelphia, PA 19014-6021
Phone: (215) 746-4177

Christopher Johnson, PhD
Department of Psychology
Utah State University
Logan, Utah, 84321
Phone: (435) 633-3211

Steve Barney, PhD
Department of Psychology
Southern Utah University
Cedar City, Utah, 84720
Phone: (435) 586-7955

Leslie Jones, PhD
Department of Psychology
Southern Utah University
Cedar City, Utah, 84720
Phone: (435) 586- 7863

Kerry Jordan, PhD
Department of Psychology
Utah State University
Logan, Utah, 84321
Phone: (435) 797-1111