

A COMPARISON OF PERSONAL ATTRIBUTE AND SCENARIO
BASED SHAME MEASURES

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

Glenna S. Rousseau

Certificate of Approval:

Phil Lewis
Professor
Psychology

Laura Vernon, Chairperson
Assistant Professor
Psychology

Chris Correia
Assistant Professor
Psychology

Stephen L. McFarland
Dean
Graduate School

A COMPARISON OF PERSONAL ATTRIBUTE AND SCENARIO
BASED SHAME MEASURES

Glenna S. Rousseau

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Auburn, Alabama
December 16, 2005

A COMPARISON OF PERSONAL ATTRIBUTE AND SCENARIO
BASED SHAME MEASURES

Glenna S. Rousseau

Permission is granted to Auburn University to make copies of this thesis at its discretion,
upon request of individuals or institutions and at their expense. The author reserves all
publication rights.

Signature of Author

Date

VITA

Glenna S. Rousseau is the daughter of John Rousseau and Florence Rousseau. Born on April 28, 1978 in Baton Rouge, Louisiana, Glenna completed her B.S. at the University of Georgia in May 2000 majoring in Psychology and graduating with honors. She is currently pursuing her doctoral degree in Adult Clinical Psychology at Auburn University.

THESIS ABSTRACT
A COMPARISON OF PERSONAL ATTRIBUTE AND SCENARIO
BASED SHAME MEASURES

Glenna S. Rousseau

Master of Science, December 16, 2005
(B.S. University of Georgia, 2000)

65 Typed Pages

Directed by Laura Vernon, Ph.D.

Shame has become an increasingly important topic in the study of emotion and psychopathology; however, the best way to measure it remains unclear. Three current, prominent shame measures include the Experience of Shame Scale (ESS), Internalized Shame Scale (ISS), and Test of Self-Conscious Affect (TOSCA). The former two we categorize as personal attribute measures, which assess shame via global negative self-statements. We refer to the latter as a scenario-based measure because participants rate the likelihood of various responses to specific situations. Participants (N=205) completed these three shame measures, measures of guilt and anger, and of mood, anxiety, and Posttraumatic Stress Disorder symptoms. Convergent validity was demonstrated as all three shame measures were significantly correlated, with the two personal attribute measures more closely associated with each other than with TOSCA

Shame. Divergent validity was not as clearly supported. Several shame scales were more closely associated with guilt and anger scales than with other shame measures. Measurement type (personal attribute, scenario-based) played a larger role in the size of the correlations than did the emotion being measured. Multiple regression analyses revealed that the ISS was the only emotion measure to independently account for a significant portion of the variance of all three measures of psychopathology: mood, anxiety, and PTSD scores. Additionally, a series of confirmatory factor analyses failed to find sufficient model fit for the shame measure data. Therefore, shame as a construct is not completely solidified or measured in a straightforward manner.

ACKNOWLEDGEMENTS

Thanks to Laura Vernon for consistently offering her knowledge and support. Without Dr. Vernon's motivation and expertise, this project would not have been possible.

Style manual or journal used APA Publication Manual (5th edition)

Computer software used Microsoft Word 2002

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
I. INTRODUCTION.....	1
II. METHOD.....	16
III. RESULTS.....	23
IV. DISCUSSION.....	37
V. REFERENCES.....	47

LIST OF TABLES

1. Means and Standard Deviations for Emotions Measures By Gender.....	23
2. Convergent and Divergent Validity for Emotion Measures.....	24
3. Bivariate Correlations between Psychopathology and Emotion	27
4. Psychopathology and Emotion Regressions.....	28
5. Goodness of Fit Indicators for Confirmatory Factor Analysis for ESS	31
6. Goodness of Fit Indicators for Confirmatory Factor Analysis for ISS	33
7. Goodness of Fit Indicators for Confirmatory Factor Analysis for TOSCA	35

LIST OF FIGURES

1. ESS 3-Factor CFA Model... ..	32
2. ISS 1-Factor CFA Model	34
3. TOSCA 1-Factor Shame Only CFA Model	36

I. INTRODUCTION

Within the last 15 years, the study of shame has shifted from a primarily theoretical approach to a more empirical approach. As a result of this shift, there has been a need for empirically validated measures of shame and a number of measures have been developed to assess trait shame from different perspectives. Utilizing recently developed measures of shame, clinical psychology has begun to examine shame and the role it plays in the acquisition and maintenance of various forms of psychopathology. The current study examined the psychometric properties of three self-report shame measures, including their convergent and divergent validity, their structure, and their psychopathology correlates.

Shame Theories

Some emotion researchers have included shame in their list of primary or fundamental emotions, which are considered the basic building blocks of all other emotions (e.g., Izard, 1997; Tomkins, 1984). Shame is part of a set of emotions, along with guilt, pride, and embarrassment, that have been defined as self-conscious emotions (Tangney, 1999), as well as moral emotions that motivate self-regulatory processes (e.g., Damon, 1988). Shame is typically thought to be characterized by a negative evaluation of the global self (Tangney, Wagner, & Gramzow, 1992b). It is regarded as a self-focused/characterological emotion, sometimes associated with events in which the

individual feels that he/she has little control (Roseman, Antoniou, & Jose, 1992). Recent conceptualizations of shame are similar to Janoff-Bulman's (1979) concept of characterological self-blame, in which an individual finds fault in him/herself following a negative event. Shame is described as an acutely painful emotion, involving self-scrutiny and imagery of how one's defective self appears to others, and therefore causing a desire to escape or hide (Tangney, 1995). Shame is thought to be characterized by feelings of being worthless, small, and powerless, and a sense of shrinking (Roseman, Wiest, & Swartz, 1994; Tangney et al., 1992b). These feelings are often accompanied by physical reactions, such as blushing, and non-verbal behaviors, such as averting the gaze and bowing the head (c.f., Keltner & Buswell, 1996).

Distinguishing Shame from Guilt and Anger. Many definitions of shame have grown out of attempts to distinguish it from guilt. Initially, definitions of guilt encompassed characteristics that are now ascribed to shame; however, researchers such as Janoff-Bulman (1979) began to report two sets of diverging results within the study of self-blaming emotions. There appeared to be an emotional response corresponding with the thought "I am a bad person" and a different, distinct response corresponding with the thought "I have done a bad thing" (e.g., Lewis, 1971). Current emotion theories characterize the former as a shameful response and the latter as a guilty response. In accordance with this characterological view of shame and behavioral view of guilt, individuals experiencing guilt are thought to feel that they have more control over the future because they need only change their behavior to produce a different outcome. However, experiencing shame is posited to result in more hopeless feelings because the self-deprecating thoughts and related behaviors feel permanent, pervasive, and

uncontrollable (Janoff-Bulman, 1979). Hence, guilt is generally thought to be less painful than shame and to generate a more empathetic and adaptive response, encouraging people to engage in corrective action, leading to reconciliation of the negative emotion (e.g., Lewis, 1971; Tangney, 1995). Shame, on the other hand, tends to lead to feelings of powerlessness, which lead to urges to avoid others (Gramzow & Tangney, 1992).

Whether guilt or shame is experienced is, recently, thought to be based largely on the individual's subjective view of an event and to a lesser degree on objective aspects of the event (Tangney, 1992; Tangney, Miller, Flicker, & Barlow, 1996). In examining the antecedents of shame, Tangney (1992) found that both objective situational variables, such as whether the situation was public or private, and subjective appraisals, such as having a sense of one's own shortcomings, were important. For example, finding that one has toilet paper trailing off of one's shoe might be amusing in private but shaming in public. On the other hand, even if objective situation variables are held fairly constant, one sexual assault survivor might feel guilty about having been in a parking garage alone late at night whereas another woman in similar circumstances might feel that the sexual assault resulted because of something dirty, ugly, and shameful within herself.

Another emotion in which both objective situational variables and subjective appraisals play a role is anger. The relationship between anger and shame has been relatively neglected in the emotion literature and often the relationship between anger and guilt or a combination of guilt and shame is examined instead (e.g., Tangney, Wagner, Fletcher, & Gramzow, 1992a). However, there has been theoretical speculation about the relationship between anger and shame and a few empirical studies examining their

association. It has been posited that although shame can follow anger, the emotions usually occur in the opposite order with the shameful individual initiating anger in an attempt to regain a feeling of control (e.g., Nathanson, 1987, Retzinger, 1987; Scheff, 1987). Lewis (1971) reported a link between shame and “humiliated fury,” an angry reaction that is initially directed at the self and then in a defensive reaction directed at the rejecting other. Tangney et al. (1992a) have reported that two measures of shame-proneness were both positively correlated with indices of anger and hostility measured by the Trait Anger Scale (Spielberger, Jacobs, Russell, & Crane, 1983); however, guilt appeared unrelated to anger. Tangney et al. (1992a) posit that anger is engendered by shame as a way to alleviate some of the pain and self-deprecation experienced with shame. The fact that the experience of guilt does not bring about the same anger response may suggest that measuring the presence of anger could be helpful in differentiating between shame and guilt.

Prominent Shame Measures

The growing empirical literature on shame has been accompanied by a number of new self-report measures that operationalize shame in different ways. Currently, trait or dispositional shame is primarily measured in two manners: situational or scenario-based measures and measures based on personal attributes. Situational or scenario-based approaches to shame measurement instruct participants to imagine their responses to specific hypothetical shame-inducing scenarios, such as those in the Test of Self-Conscious Affect (TOSCA; Tangney, Wagner, Gramzow, 1989). The personal attribute approach, on the other hand utilizes shame-related statements about one’s character, such as those in the Internalized Shame Scale (ISS; Cook, 1989). Participants then endorse

such descriptions of their personality or character at varying levels of agreement or intensity. Andrews and Hunter's (1997) Externalized Shame Scale (ESS) would also fit into the personal attributes measurement category because the interview-style items question participants about how they think and feel in general, not in relation to a specific situation. The development of each of the ISS, ESS and TOSCA Shame measures will now be reviewed.

The Internalized Shame Scale (ISS). The current version of the ISS has been used since 1989 and was developed by Cook to study shame in individuals with psychological disorders. Both non-clinical and clinical samples (several chemical dependency treatment programs, two inpatient programs, and two outpatient programs) were used to obtain reliability data for this measure. The alpha reliability coefficients of the scale in both a non-clinical and clinical population are quite strong, .95 and .96, respectively. A subset of the non-clinical population was retested on the ISS after an interval of seven weeks, resulting in an adequate test-retest correlation of .84 for the shame items. Modest convergent validity has been demonstrated by correlations of the ISS shame measure with another shame measure, the Personal Feelings Questionnaire (PFQ)- shame subscale, $r=.63$. In regards to divergent validity, Cook and Wong (1993) reported a correlation of .56 between ISS shame and trait anger as measured by the State-Trait Anger Expression Inventory (STAXI-2). Interestingly, this association is not much weaker than the PFQ, which is purported to measure the same construct as the ISS. No examination of the structure of the underlying model could be found.

Cook (1996) argues that shame should be measured using a trait format because it is related to many shaming situations that have been experienced over time and

“internalized” as part of an individual’s self-concept. Thus, he posits that items must be globally related to the self in order to access shame. There is some support for this assertion. Allan, Gilbert, and Goss (1994) conducted a study similar to the present one with questionnaires that are not as well-known. They compared shame scales based on global negative beliefs (e.g., Other As Shamer Scale (OAS; Goss, Gilbert, Allan, 1994), the ISS) with scales that attempt to measure shame responses to specific events, such as the Dimensions of Conscience Questionnaire (DCQ; Gilbert, Pehl, & Allan, 1994). They found that personal attribute scales show stronger associations with measures of psychopathology, such as the Beck Depression Inventory (BDI; Beck & Steer, 1987) and the General Health Questionnaire-28 (Goldberg & Williams, 1991), than the DCQ. Personal attribute items, such as those in the ISS, resemble self-statements; thus, an individual that endorses an item is likely identifying with the trait that it represents.

By developing the ISS utilizing inpatient and outpatient populations, Cook intended it to be efficacious in a clinical setting. He utilized the responses of clinical, as well as non-clinical populations to decide which items defined and measured the construct of shame. However, the ISS has been only lightly researched (e.g., Wong & Cook, 1992; Allan et al., 1994) relative to the TOSCA, but at a similar level of research as the ESS.

The Experience of Shame Scale (ESS). Andrews and Brown (1993) support the interview as an additional method of shame measurement. Andrews endorses direct questions about whether an individual has felt ashamed of specific personal attributes or behaviors, rather than putting them in a statement, such as Cook (1994) or employing hypothetical scenarios such as those used by Tangney et al. (1989). Andrews notes that

shame interviews have been effective in prospective studies predicting depression (Andrews, 1995) and PTSD symptoms (Andrews, Brewin, Rose, & Kirk, 2000). Therefore, Andrews and Hunter (1997) created the ESS as a self-report paper-and-pencil version of their shame interview. Typically, self-report measurement is thought to have less predictive validity than an interview format. Andrews et al. (2002) propose that the ESS measures a specific predisposition to experience shame instead of state shame or general trait negative affect. The ESS is a new measure and thus has not yet been widely used.

Andrews et al. (2002) report a strong internal consistency (Cronbach's alpha) of the scale, 0.92. They also found a test-retest reliability of .83 after 11 weeks. The ESS total scale and its component subscales were all significantly correlated with the TOSCA Shame scale (total scale, $r=.61$), which suggests convergent validity. Discriminant validity was demonstrated by the weak correlation between the ESS and the TOSCA Guilt scale ($r=.23$).

Andrews et al. (2002) hypothesized that a three-factor model aligned with their three bodily, behavioral, and characterological shame subscales would best fit the ESS and compared this to a two-factor model, which included bodily shame as one factor and combined behavioral and characterological for the second factor. They found that the three-factor model provided the best fit (normed fit index of .91, relative fit index of .89).

Test of Self-Conscious Affect (TOSCA-3). Tangney (1992, 1996) asserts that guilt and shame are two separate constructs, but that individuals have difficulty distinguishing them conceptually. Thus, the TOSCA's scenario-based measurement format asks participants to imagine their responses to different situations without requiring them to

conceptually discriminate between guilt and shame (Lindsay-Hartz, 1984). Tangney's original scenario-based measure, the Self-Conscious Affect and Attribution Inventory (SCAAI; Tangney, Burggraf, Hamme, & Domingos, 1988) consists of a series of negatively and positively valenced scenarios with multiple response options, two of which are representative of "shame proneness" and two of "guilt proneness." Although the original version of the TOSCA (Tangney, Wagner, & Gramzow, 1989) was based on the SCAAI, it consists of entirely new scenarios. The TOSCA included a number of improvements upon the SCAAI. First, in an attempt to improve the measure's ecological validity Tangney used items that were subject generated, rather than experimenter generated. The original TOSCA and later versions (e.g., the TOSCA-3; Tangney, J.P., Dearing, R., Wagner, P.E., & Gramzow, R., 2000) were created from written accounts provided by college students and community volunteers describing experiences that generated shame, guilt, and pride. Secondly, the TOSCA items were intended to apply not only to college students, but to all adults. Thus, a subset of TOSCA items were drawn from a pool of affective, cognitive, and behavioral responses provided by adults not attending college. Since it was created to distinguish between shame and guilt proneness, the TOSCA, in recent years, has become the most widely used empirical measure of guilt and shame.

Estimates of internal consistency (Cronbach's alpha) for the TOSCA subscales are consistently reported near 0.76 for shame, 0.66 for guilt, (Fontaine, Luyten, DeBoek, & Corveleyn, 2001; Luyten, Fontaine, & Corveleyn, 2002; Tangney, 1992). In regards to test-retest reliability, Tangney and colleagues (1992a) reported a shame scale correlation of .85 over a 3-5 week period. Ferguson and Crowley (1997b) reported a modest

correlation ($r=.39$) between the TOSCA Shame scale and the Personal Feelings Questionnaire (PFQ-2) (Harder & Zalma, 1990), indicating some convergent validity. In reference to divergent validity, the same study reported a correlation of .26 between TOSCA Shame and PFQ-2 Guilt. Additionally, STAXI Trait Anger was modestly correlated with TOSCA Shame ($r=.17$, $p<.05$) indicating that these emotions likely co-occur sometime and in some cases anger may be present as a method to alleviate the personal pain related to shame (Tangney et al., 1992a).

To examine the internal structure of the TOSCA, Fontaine and colleagues (2001) conducted a series of Confirmatory Factor Analyses (CFAs). They reported encountering some difficulties, such as the complex nature of the measure with four separate scores within each item, some items presenting as very skewed, and the cross-loading of items. When they conducted an Exploratory Factor Analysis (EFA), the researchers found 23 items loaded highly on multiple scales. They referred to these cross-loaded items as non-differentiating. Of importance for the present investigation, ten guilt items were found to load highly on both the guilt and shame factors and three shame items loaded onto the guilt factor at a significant level. Then, item parcels were constructed using all 55 TOSCA items, with differentiating and non-differentiating items from each hypothesized scale placed in separate parcels (e.g., two separate parceled shame indicants, one composed of differentiating items and one of non-differentiating items). Fontaine and colleagues then compared two models. One model was based on Tangney's hypothesized four-factor model with each item parcel loading onto only one of the four factors aligned with each subscale (i.e., shame, guilt, externalization, and unconcern). The second model consisted of the 12 differentiating parcels loaded on a single factor and

the 8 non-differentiating parcels being allowed to load on the two factors they were previously found to cross-load on. When these two models were tested on a separate sample, the second model provided the only acceptable fit of the data (RMSEA=.05; SRMR=.04; $\chi^2=ns$). Obviously, the first set of analyses would have been more rigorous had it used an independent sample to test the models, rather than using the same sample from which the parcels had been constructed. Most important for our purposes, the model of Tangney's TOSCA subscales did not provide an acceptable fit for the data.

Currently, it is unclear whether situational self-report instruments such as the TOSCA and personal attribute instruments such as the ISS are examining the same aspects of shame. Cook (1996), the author of the ISS, asserts that it is not possible for a situational measure to accurately assess trait-like aspects of shame as this type of measure examines only responses to a specific situation. He posits that shame is something that is internalized and thus requires global items about the self to tap into the characterological variants of shame. Although Tangney (1996), the author of the TOSCA, agrees that situational measures have constraints concerning the range of shame-inducing situations that are possible to present, she asserts that situational measures are valid methods for assessing trait characteristics of shame. Tangney argues that the TOSCA consists of broadly applicable situations in which almost any individual can imagine him/herself, thus encouraging the application of traits to the interpretation of the situations. Tangney (1996) posits a possible confound with measures that focus on personal attributes, suggesting that an individual may become more defensive when responding to such items relative to behavioral items regarding a specific situation. Additionally, Tangney argues that situational measures are better able to distinguish between shame and guilt because

responses to measures such as the TOSCA indicate whether the individual found fault in a specific behavior he/she had performed (guilt) or blamed a characterological aspect of him/herself (shame) without requiring the respondents to understand the definitions of shame and guilt.

Despite such arguments regarding the strengths of these different approaches, strong positive correlations ($r=.61$, $p<.001$) have been reported between the TOSCA Shame and the ESS (Andrews, Qian & Valentine, 2002). However, without an examination of the correlates of each, it is not clear whether the ESS and the TOSCA may be measuring different aspects of shame. Further, neither measure has been compared to the recently developed ESS. The current study examined the associations among the TOSCA, ISS, and ESS and their correlates.

Shame and Psychopathology

Shame and Depression. As more studies are conducted in the area of shame, a clear link has been established between shame and the presence of psychopathology. However, questions remain concerning the nature of this relationship. The cross-sectional study has been used often to establish a significant association between shame and depression (e.g., Allan, Gilbert, & Goss, 1994; Cook, 1996; Tangney, Wagner, & Gramzow, 1992). For example, a number of studies have reported significant positive associations of the ISS with various measures of depression, including the Depression Scale of the Multiple Affect Adjective Check List-Revised (Zuckerman & Lubin, 1985), the Severe Depression Scale of the General Health Questionnaire (Goldberg & Hillier, 1979), and the BDI in both clinical and non-clinical samples (e.g., Allan, Gilbert, & Goss, 1994; Weiner, 1993). Other studies have reported similar significant positive

correlations between the TOSCA Shame scale and the BDI (Fontaine, Luyten, DeBoeck, & Corveleyn, 2001; Luyten, Fontaine, & Corveleyn, 2002). Woien et al. (2003) reported that TOSCA Shame was also significantly correlated with the depression subscale of the Hopkins Symptom Checklist (Derogatis, Lipman, Rickets, Uhlenhuth, & Covi, 1974). Tangney et al. (1992b) conducted a series of hierarchical regression analyses, reporting that the TOSCA Shame scale accounted for a significant amount of variance in scores on the BDI and the depression subscale of the Symptom Checklist-90 (SCL-90; Derogatis, Lipman, & Covi, 1973).

Although there is a growing body of research documenting cross-sectional associations between shame and depression, few studies have examined shame's role in the onset of depression. Andrews et al. (2002) conducted a longitudinal study examining shame, guilt, and depression at two time points eleven weeks apart. Shame was measured by the ESS and the TOSCA, guilt was measured by the TOSCA, and depression was measured by the SCL-90. Andrews and her colleagues found the ESS and TOSCA Shame and Guilt to be significantly associated with depression at time one, but only the two shame measures were significant at time two. Interestingly, in a multiple regression analysis, when time one depression was controlled for, the time one ESS score continued to make an independent contribution to time two depression, whereas the TOSCA Shame and Guilt subscales did not. Andrews et al. argue that these results suggest that the relationship between ESS and depression is not purely due to shared general negative affectivity. They propose instead that the ESS measures a specific trait proneness to shame that contributes to depression. Additional evidence demonstrating converging results with another measure designed to assess trait shame, as

well as a demonstration of divergent results with trait measures of other additional negative emotions, would allow for more confidence in these claims

Shame and Post Traumatic Stress Disorder (PTSD). There is still much to learn about the relationship between shame and depression, and links between shame and PTSD are even less well understood. Not surprisingly, various studies (e.g., MacKinnon, 1979; Armsworth & Holaday, 1993) have demonstrated that shame is common and can be quite intense following trauma. There is some evidence to suggest that shame may play a role in the development and/or maintenance of PTSD. Leskela, Dieperink, and Thuras (2002) reported that the shame-proneness subscale of the TOSCA was significantly positively correlated with PTSD symptom severity as measured by the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993), although TOSCA Guilt was not. Although predictive studies have not been conducted with trait measures of shame, such as the TOSCA, one longitudinal study utilized a single-item retrospective measurement of state shame at the time of the trauma. Andrews, Brewin, Rose, and Kirk (2000) reported that the shame item was an independent predictor of PTSD symptom severity at one month and six-months following the trauma. In fact, at the six-month point, shame was the only variable of a host of demographic, sociocultural, emotional, and trauma severity variables to be significantly associated with PTSD symptom level. These results hint at the potential importance of shame in PTSD and underscore the need for additional PTSD research utilizing standardized, empirically validated measures of both state and trait shame.

Shame and Substance Abuse. Shame and PTSD have also been studied within the context of other types of psychopathology, such as substance abuse. Wong and Cook

(1993) compared ISS trait shame levels in a group of individuals with PTSD relative to individuals with a substance abuse diagnosis. They reported that the PTSD group scored significantly higher on the ISS than the substance abuse group. Interestingly, most studies (e.g., Janikowski & Glover, 1995; Lisak, 1995) that report a link between shame and substance abuse involve a population that has experienced a traumatic event.

Substance abuse has been described as a self-medication coping strategy to deal with shame, especially in women (e.g., Gomberg, 1989). The current study will evaluate the relationships among trait shame, trauma, and substance abuse.

Goals of the Present Study

The current study used three different self-report shame measures, the ESS, ISS, and TOSCA-3, each of which examines shame in a different manner: interview style questions, descriptive self-statements, and hypothetical scenarios, respectively. The goal of the current study was to examine the factor structure of these instruments, test whether the shame scales are associated with each other, are differentially related to two discriminant negative emotions, guilt and anger, and are differentially related to various forms of psychopathology. It is expected that females will report higher levels of shame than males on all three measures of shame, replicating previous findings of gender differences reported in the shame literature (e.g., Ferguson & Crowley, 1997; Lutwak & Ferrari, 1996).

We also explored associations among the shame measures. It was expected that the ISS shame subscale would be more strongly associated with the ESS than the TOSCA-3 Shame. We expected TOSCA-3 Shame to be equally associated with the total scale scores on the ESS and the ISS. The creators of the TOSCA, ISS, and ESS all assert

that their instruments measure dimensional, trait-like aspects of shame. If this is the case, then scores on all three shame scales should be highly correlated. Additionally, if the measures are assessing the same aspects of shame, they should all differ from other negative emotions, such as anger and guilt, in a similar manner.

The current study also explored associations among the shame and psychopathology measures. We expected the ESS and ISS to be more strongly associated with all measures of psychopathology than the TOSCA-3 because the former measures consist of items which are of similar format to the psychopathology questionnaires. In reference to discriminant emotions, we expected a measure of guilt based on feelings and attributions in response to a specific trauma, the Trauma-Related Guilt Inventory (TRGI; Kubany, Haynes, Abueg, Manke, Brennan, & Stahura, 1996), to be more highly positively correlated with the TOSCA-3 Guilt scale than the ESS and the ISS Shame scale. However, it was predicted that the ESS and ISS will be more strongly associated than TOSCA-3 Shame with the TRGI because of similar measure type and with the STAXI-2 because the STAXI-2 is also a personality-based measure.

In order to examine the structure of the three shame measures, Confirmatory and Exploratory Factor Analysis were conducted. The current study examined several different theoretically sound models for each shame measure to determine the best model fit for the data. It was expected that the best model for the ESS would be a three-factor model parallel to its three subscales and that the best ISS model would have two factors, shame and self-esteem. It was hypothesized that a one-factor model of TOSCA Shame would provide the best fit and that the Guilt scale would not create a cohesive factor, similar to the findings of Fontaine et al. (2001).

II. METHOD

Participants

There were 203 participants ranging from age 19 to 26 years old ($M=20.4$). The majority of the sample (72%) was female and Caucasian (78%), with 14% African-American and 4% Asian-American. Most participants were from a southern region of the U.S.: 62% from the East South Central region (Alabama, Kentucky, Mississippi, and Tennessee) and 22% from the South Atlantic region (Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia).

Demographic Questionnaire

A demographic questionnaire assessed participants' personal and historical characteristics, including gender, ethnicity, age, and level of education. It also included state of origin, based on U.S. Census Bureau divisions (Cohen, 2001; Cohen, Vandello, Puente, & Rantilla, 1999).

Emotion Measures

The Experience of Shame Scale (ESS). The ESS (Andrews et al., 2002) was developed from an interview measure designed to assess shame (Andrews & Hunter, 1997). It is a 25-item questionnaire consisting of three subscales: Characterological Shame, Behavioral Shame, and Bodily Shame. For the current study, analyses were conducted with the total scale only. Participants are instructed to respond by considering

feelings they have experienced in the past year. Each item (e.g., “Have you ever tried to cover up or conceal any of your personal habits?”) is rated on a 4-point scale, ranging from 1=not at all to 4=very much. Similar to past estimates of internal consistency (Andrews et al., 2002), the Cronbach’s alpha of the total scale was .93 in the present study. Andrews and colleagues (2002) reported that the ESS total scale and its component subscales were all significantly correlated with the TOSCA Shame scale (total scale, $r=.61$) which suggests construct validity.

The Internalized Shame Scale (ISS). The ISS (Cook, 1993) is a 30-item self-report questionnaire. Twenty-four items create a shame subscale designed to explore negative global evaluations of the self (e.g. “I feel like I am never quite good enough,” “I think others are able to see my defects”). The remaining six items are positively worded self-esteem items (e.g. “I feel I have a number of good qualities”), which will not be used in the analyses of the present study. Participants rate the frequency of each experience on a 5- point scale, from 1=never to 5=almost always. The ISS was constructed using a large clinical ($n=370$) and student sample ($n=645$). For the non-clinical participants, there was a test-retest correlation of 0.84 for the shame scale at seven weeks. Consistent with previous reports (Cook, 1993), for the current study, the internal consistency (Cronbach’s alpha) for the shame scale was .95.

Test of Self-Conscious Affect (TOSCA-3). The TOSCA-3 (Tangney et al., 2000) presents a series of 16 scenarios (11 negative and 5 positive) that people are likely to encounter in day-to-day life (e.g., situation- “While out with a group of friends, you make fun of a friend who is not there;” responses to be rated- “You would think ‘It was all in fun; it’s harmless,” “You would feel small... like a rat,” You would think that perhaps

that friend should have been there to defend himself,” You would apologize and talk about that person’s good points”). Participants rate the probability of acting out each of 4 or 5 possible responses provided for each scenario on a 5-point scale, ranging from 1=not likely to 5=very likely. Each response to a scenario represents an action consistent with guilt, shame, externalization, or detachment/unconcern, which are considered the most commonly experienced reactions to guilt and shame inducing situations (Tangney, 1995). Estimates of internal consistency (Cronbach’s alpha) for the TOSCA Guilt and Shame subscales respectively were .71 and .72 in the present study. Although these alphas are not high, they are consistent with past reliability reported in the literature (Fontaine, Luyten, DeBoek, & Corveleyn, 2001; Luyten, Fontaine, & Corveleyn, 2002; Tangney, 1992).

Trauma-Related Guilt Inventory (TRGI). The TRGI (Kubany, Haynes, Abueg, Manke, Brennan, & Stahura, 1996) is a 32-item inventory with three scales and three subscales. None of the three scales, Distress, Global Guilt, and Guilt Cognitions, nor the subscales, were analyzed separately in the current study. Only the total score was used. Participants respond to all items on a 5-point scale indicating how “true” that item is for the participant, 1=extremely true to 5=not at all true. Before completing the questionnaire, participants are instructed to take a few moments to reflect on their thoughts and feelings related to a single traumatic experience. The current internal consistency for the total scale was strong, .93, consistent with previous alpha coefficients reported for the scales ranging from .90 to .94 (Kubany et al., 1996).

State-Trait Anger Expression Inventory (STAXI-2). The STAXI-2 (Spielberger, 1988) is comprised of 44 items that measure state anger, trait anger, and anger

expression. Participants rate each item on a 4-point scale, indicating the intensity of anger on the state anger scale, the frequency they feel angry on the trait scale, and the frequency they engage in different types of expressive behavior on the expression scale. This inventory consists of six scales that measure anger: State Anger, Trait Anger, Anger-in, Anger-out, Anger Control, and Anger Expression. Trait Anger, with an internal consistency of .84, was the only scale utilized in the present analyses.

Psychopathology Measures

Life Events Checklist (LEC). This measure includes a list of 17 potentially traumatic events (e.g., sexual assault, physical assault, transportation accident). Participants indicate whether they have experienced each event, have witnessed the event happening to someone else, or have learned about the event happening to someone close to them. Participants are then instructed to think about the event they consider to be the most traumatic and to answer questions addressing DSM-IV-TR Criterion A1, such as whether someone's life was in danger during the event, and Criterion A2, whether they experienced fear, helplessness, or horror during the event.

PTSD Checklist (PCL). The PCL (Weathers, Litz, Herman, Huska, & Keane, 1993) consists of 17 items corresponding to the 17 DSM-IV PTSD symptoms, such as distressing dreams and hypervigilance. While considering their most traumatic event from the LEC, participants rate the degree to which they have been bothered by each symptom during the previous month, on a scale from 1=not at all to 5=extremely. Thus, the PCL total score provides a dimensional index of PTSD symptom severity. The checklist can also be used categorically; those participants scoring above a cutoff score (either 50 or 44) are considered to have PTSD. In the current study, this measure was

used both categorically with a cutoff score of 44 and as a dimensional index. Scores on the PCL have been demonstrated to be closely related to results from two widely utilized diagnostic interviews, the Structured Clinical Interview for DSM-III-R (SCID; Spitzer, Williams, Gibbon, & First, 1990) and the Clinician Administered PTSD Scale (CAPS; Blake, Weathers, Nagy, Kaloupek, Gusman, Charney, & Keane, 1995; Blanchard, Hickling, Taylor, & Forneris, 1995). Diagnostic agreement between the PCL and the SCID (which produces categorical diagnoses: PTSD is present or absent) is high, pairwise kappa = .64 (Weathers et al., 1993). Scores on the PCL and the CAPS (which produces dimensional ratings of PTSD symptom severity, which can be made categorical through the use of several standard cutoffs) are highly correlated, $r = .93, p < .0001$ (Blanchard et al., 1995; 1996). For the current study, the Cronbach's alpha was .91.

Survey of Smoking Behaviors. This questionnaire contains 17 self-report items that assess participants' current and past smoking status. The items are based on a smoking measure developed by Fava, Velicer, and Prochaska (1995) that included questions regarding history and frequency of smoking as well as motivation to quit. Participants choose a statement that best describes their "smoking status" (e.g., "I am a daily smoker"). For frequency, participants report how many days in the last 30 they have smoked cigarettes and how many cigarettes on average they have smoked in the last 30 days. The variable utilized in current analyses was average number of cigarettes smoked in the previous 30 days.

Drug Use Inventory. This measure asks participants to report their use of a range of licit and illicit drugs, including tobacco products, heroine, and cocaine. Questions include the age of first use, total years used, and frequency of current use for each of the

13 substances listed. Frequency is measured by the number of days in the last 28 that the participant has used each substance, which was used as a dimensional variable in the current study. Research supports the validity of self-reported drug use, including tobacco use, particularly when confidentiality is assured and the assessment is done in a clinical or research setting (Patrick, Cheadle, Thompson, Diehr, Koepsell, & Kinne, 1994; Sobell, Toneatto, & Sobell, 1994).

Mood and Anxiety Symptom Questionnaire (MASQ). This instrument consists of 90 items reflecting Watson and Clark's (1991) tripartite model of anxiety and depression. It is defined in terms of common symptoms relating to general distress, anxiety-specific symptoms of hyperarousal, and depression-specific symptoms of low positive affect/loss of interest (anhedonia). This instrument consists of five subscales: General Distress: Mixed, General Distress: Depression, General Distress: Anxiety, Anxious Arousal, and Anhedonic Depression. Participants rate the extent of each symptom during the past week on a scale from 1= not at all to 5= extremely. In the present study, the two subscales utilized were Anhedonic Depression (22 items; "felt really bored") and Anxious Arousal (17 items; "heart was racing"), as these represent two distinct aspects of depression and anxiety. The Anhedonic Depression and Anxious Arousal scales had alphas of .93 and .88, respectively.

Procedures

Participants who completed the above questionnaires were compensated with extra credit in psychology classes. To avoid order effects, questionnaires were presented in one of four semi-randomized orders. Due to the nature of the questionnaires, the LEC,

in which participants identify their most upsetting trauma, always preceded the PCL, in which participant's report PTSD symptoms relating to the most upsetting trauma as identified on the LEC. The order of presentation of all other questionnaires was randomized.

II. RESULTS

Gender Differences

To examine potential gender differences among the emotion measures, we conducted a 2 (gender: female vs. male) X 6 (emotion measure: ESS vs. ISS Shame vs. TOSCA Shame vs. TOSCA Guilt vs. TRGI vs. STAXI) MANOVA and found significant results $F(6,180)=3213, p<.01$. We found significant between-subjects main effects for TOSCA Shame, $F(1,186)=8.9; p<.01$, and TOSCA Guilt, $F(1,186)=19.9; p<.01$, with women reporting more shame and guilt than men. No other measures demonstrated significant gender differences. The means and standard deviations for the emotion measures are presented by gender in Table 1.

Table 1
Means and Standard Deviations for Emotion Measures by Gender

	possible range	Males (N=56)		Females (N=146)	
		<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>
ESS	25-100	53.56	14.05	53.84	14.04
ISS	0-96	26.43	14.10	29.17	18.92
TOSCA- Shame	16-80	42.21	9.41	46.97	9.17
TOSCA- Guilt	16-80	61.14	7.17	66.21	6.06
TRGI	32-160	65.84	24.47	65.83	21.92
STAXI	10-40	19.39	4.69	18.45	5.24

Convergent Validity

Table 2 presents the correlations among the three shame measures (TOSCA-Shame, ESS, and ISS-Shame). Not surprisingly, participants' scores on all three shame measures, and their subscales, were significantly correlated with one another ($ps < .01$). Although the measures were all significantly associated, the strengths of these relationships varied and followed the predicted pattern. The ISS and the ESS Total score, both of which are personal attribute measures, were more highly related to one another (.63) than to the TOSCA Shame scale, a hypothetical scenario based measure. The TOSCA Shame scale was moderately correlated with the ESS Total score (.37) and the ISS (.48). Even when the ESS was examined at the subscale level, the personal attribute measures were more closely related to one another than to the hypothetical scenario measure.

Table 2

Convergent and Divergent Validity for Emotion Measures

	ESS Total	ESS Charecterological	ESS Behavioral	ESS Bodily	ISS Shame	TOSCA Shame	TOSCA Guilt	TRGI Total
ESS Total								
ESS Charecterological	.90***							
ESS Behavioral	.88***	.66***						
ESS Bodily	.66***	.48***	.42***					
ISS Shame	.63***	.60***	.50***	.47***				
TOSCA Shame	.37***	.27***	.34***	.34***	.45***			
TOSCA Guilt	.04	-.05	.06	.15*	.06	.41***		
TRGI Total	.25**	.20**	.22**	.21**	.35***	.05	-.09	
STAXI Trait Anger	.46***	.40***	.41***	.33***	.45***	.21**	-.05	.26***

Note. ** $p < .01$
 *** $p < .001$

Divergent Validity

To investigate whether the three shame measures were quantifying negative emotion in general, or shame specifically, the shame measures were compared to two guilt measures (TOSCA-Guilt and TRGI) and one anger measure (STAXI). As shown in Table 2, the TOSCA Guilt scale was not significantly correlated with the ISS or the ESS Total score, but was significantly associated with the TOSCA Shame scale (.41). In contrast, the TRGI displayed the opposite pattern and was significantly related to the ISS and ESS (.25, .35, respectively), but not the TOSCA Shame scale. The pattern of associations between the shame and guilt measures suggests correlations among similar measure types. The scenario-based guilt measure (TOSCA Guilt) was associated with the scenario-based shame measure (TOSCA Shame) and not the personal attribute shame measures (ISS, ESS). In contrast, the personal attribute guilt measure (TRGI) was associated with the personal attribute shame measures and not with the scenario-based shame measure. Unlike the guilt measures, which were differentially associated with the shame measures, the anger measure was significantly related to all shame measures ($p < .01$): TOSCA-Shame (.21), ISS (.45), and ESS (.46).

Psychopathology Correlates

To further study the similarities and differences among the shame measures, as well as their possible clinical utility, their psychopathological correlates were examined. Correlations among the shame measures and measures of PTSD, depressive and anxious symptoms, and substance use were calculated. Table 3 presents the correlations of the shame measures with the PCL and the five subscales of the MASQ. The PCL was

significantly correlated with the ESS Total score (.15) and the ISS Shame scale (.33), but not the TOSCA Shame scale.

There were also significant associations among the shame measures and the anxiety and depression symptom measures. The TOSCA Shame scale was significantly correlated with four of the five MASQ subscales. Interestingly, similar to our finding with the PCL, TOSCA Shame was not significantly related to the MASQ Anxious Arousal subscale. All five subscales of the MASQ were significantly related to the ESS Total score and ISS Shame scale. In general, these two personal attribute measures had higher correlations with the psychopathology measures than did the hypothetical scenario measure, the TOSCA. The ISS consistently demonstrated the highest associations with the psychopathology measures.

The trend for the ISS to be most highly associated with psychopathology measures was continued with substance use. So few individuals reported recent use of most of the substances assessed, that only tobacco, alcohol, and marijuana use was reported with sufficient frequency for data analysis. Of these three, the only substance to significantly correlate with any shame measure was tobacco (cigarettes). The ISS and the ESS were both associated with cigarette use, whereas TOSCA Shame was not significantly correlated with any substance use. Given the well-known comorbidity between PTSD and alcohol use disorders, associations among alcohol use and the emotion measures were examined separately among those with PTSD. These analyses revealed a significant association ($r=.6$; $p<.05$) between TOSCA Shame and the number of days the participant had drunk alcohol in the last 28 days. These results may suggest that shame is a mediator for PTSD and alcohol use.

Psychopathology correlations were also conducted with the measures of anger and guilt. As presented in Table 3, both the STAXI and the TRGI were significantly correlated with the PCL and all five subscales of the MASQ. In contrast, TOSCA guilt was only significantly correlated with MASQ Anhedonic Depression, and surprisingly this was in a negative direction, suggesting that lower levels of trauma-related guilt were associated with higher levels of anhedonic depression. In regards to substance use, cigarettes were the only substance significantly correlated with Trait Anger. Interestingly, alcohol use was significantly associated with trauma related guilt (TRGI), but not general guilt proneness as measured by the TOSCA Guilt scale. Marijuana was not significantly correlated with any emotion measure.

Table 3

Bivariate Correlations Between Psychopathology and Emotion

	ESSTotal	ISSShame	TOSCAShame	STAXI	TOSCA Guilt	TRGI
PCLTotal	.15*	.33**	.11	.33**	.08	.41**
MASQAnhedonic Depression	.38**	.66**	.22**	.34**	-.14*	.33**
MASQGeneral Distress Depression	.49**	.77**	.30**	.45**	.05	.3**
MASQGeneral Distress Mix	.37**	.58**	.21**	.44**	.04	.25**
MASQGeneral Distress Anxiety	.46**	.52**	.23**	.43**	.09	.22**
MASQAnxious Arousal	.31**	.45**	.13	.36**	.03	.21**
Alcohol Use in Last 28 Days	-.01	.04	.07	.14	.00	.19*
Cigarette Use in Last 30 Days	.14*	.21**	.04	.27**	-.06	.22**

Note. * $p < .05$.

** $p < .01$

Since PTSD, anxiety, and depression symptoms were generally related to the emotion measures, we next examined whether the emotion measures contributed to

psychopathology scores independent of one another. We selected three measures indicative of distinct aspects of psychopathology: the PCL, MASQ Anxious Arousal, and MASQ Anhedonic Depression. We conducted three multiple regression analyses, each with a single form of psychopathology as the dependent variable, and the six emotion measures as the predictor variables. Table 4 shows that emotion accounted for a significant portion of the variance in psychopathology scores (for the PCL, MASQAD, and MASQAA, 28%, 48%, and 25%, respectively). All predictor variables with the exception of TOSCA shame were significantly independently associated with PCL scores. In contrast, only ISS Shame and STAXI Anger were significantly independently associated with MASQ Anxious Arousal. ISS Shame and TOSCA Guilt were related to anhedonic depression above and beyond the effects of other measures of emotion. Notably, the ISS shame scale was the only emotion measure to independently account for a significant portion of the variance of the scores of all three forms of psychopathology.

Table 4

Psychopathology and Emotion Regressions

	β						Model	
	ESS	ISS	TOSCA Shame	STAXI	TOSCA Guilt	TRGI	R	F
PCL	-.21*	.25**	-.04	.25**	.14*	.32***	.28	11.39***
MASQAA	0	.41***	.07	.13*	.10	.03	.25	9.95***
MASQAD	-.07	.67***	0	.04	-.14*	.10	.48	27.03***

Note. * $p < .05$

** $p < .01$

*** $p < .001$

Since our PTSD measure, the PCL, can also be used categorically, we also conducted a 2 (PTSD: presence vs. absence) X 6 (emotion measure: ESS vs. ISS Shame

vs. TOSCA Shame vs. TOSCA Guilt vs. TRGI vs. STAXI) MANOVA. The overall model was significant, $F(6,180)=1730, p<.01$. Follow-up contrasts revealed significant between-subjects main effects for ISS Shame, $F(1,186)=7.6; p<.01.$, STAXI Trait Anger, $F(1,186)=13.2; p<.01.$, and TRGI, $F(1,186)=14; p<.01$, with the PTSD group reporting higher emotion on each measure than the non-PTSD group. There were no other significant group differences. These main effects are not surprising based on the high correlations of the ISS Shame and STAXI Trait Anger with the PCL dimensional score. Due to the significant correlations of the ESS with the PCL, however, it would be expected that this measure would also have shown a significant main effect on PTSD presence, but they did not.

Factor Structure

In order to examine the factor structure of the three shame measures, CFAs were performed. Models were examined by AMOS (5.0) maximum likelihood factor analysis (Arbuckle, 2003). Less than .15% of data was missing. In order to obtain goodness of fit indices in AMOS, SPSS (11.5) linear interpolation was used to impute missing data. All analyses were run without missing data and are reported in order to include goodness of fit indices. The models were evaluated by a variety of fit measures that are classified as absolute, relative, and population discrepancy. The measure of absolute fit used in this study was the Goodness of Fit Index (GFI). The measure of relative fit, comparing the hypothesized model to the null model, was the Comparative Fit Index (CFI; Bentler, 1990). Lastly, the population discrepancy measure was the Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993).

For the ESS, we examined Andrews' proposed three-factor model, and compared it to a one-factor model and the null model. The models were constructed using all 25 items as indicants. ESS CFA fit indices are displayed in Table 5. The model with the best fit was the three-factor model, shown in Figure 1. In the three-factor model, all measured variables loaded on the expected factor at a statistically significant level, $p < .05$. However, the model chi-square was significant, supporting the null model. The large sample size in the present study likely contributed to the significance of the chi-square since it is sensitive to large sample sizes (Schumacker & Lomax, 1996).

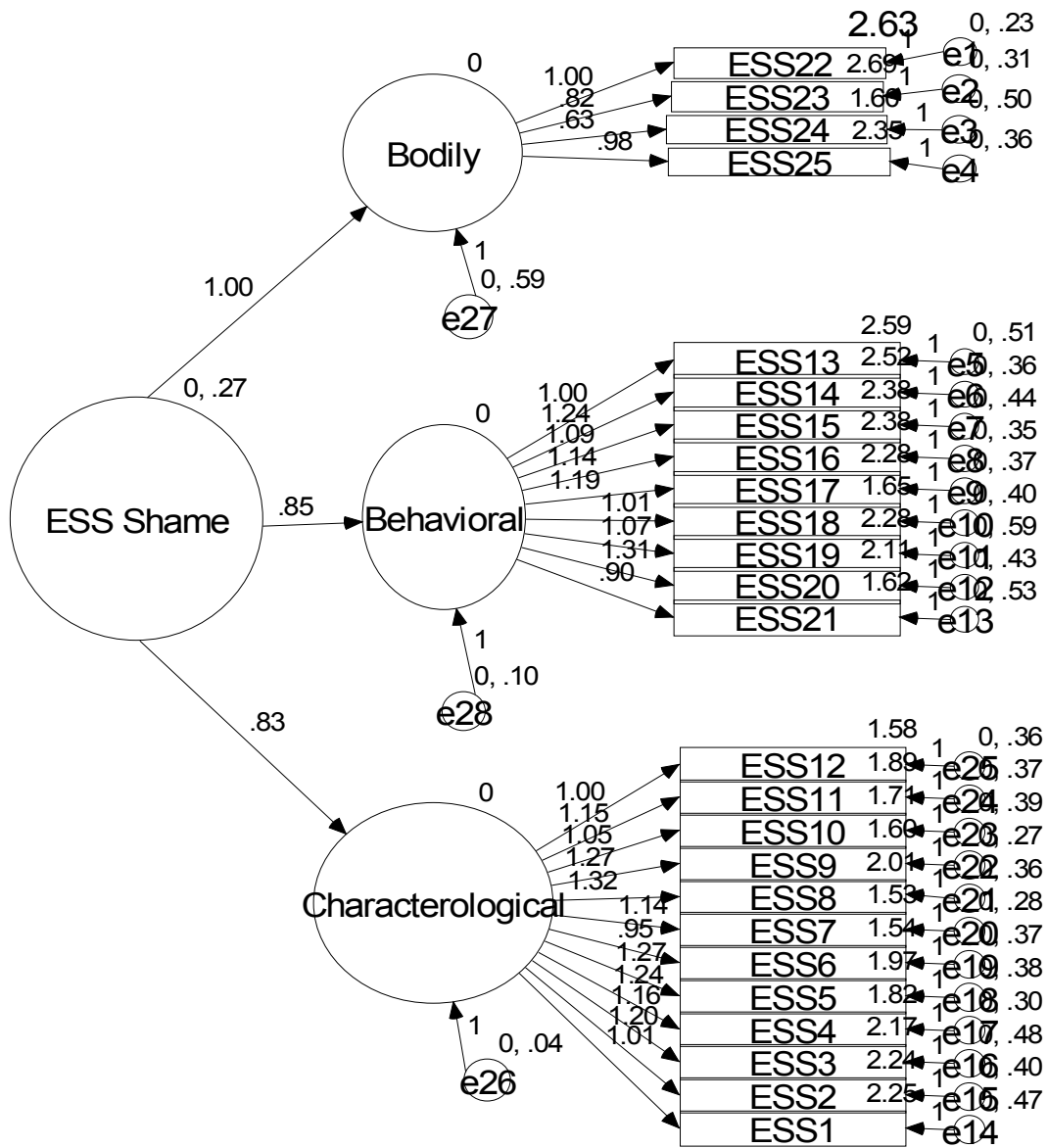
There are differing opinions about acceptable values for model fit (e.g., Browne & Cudeck, 1993; Hu & Bentler, 1999; Steiger, 1990; Vandenberg & Lance, 2000). Relatively lenient model fit cutoffs of GFI and CFI greater than .90 and a RMSEA below .10 have been suggested recently (Guarino, Shannon, & Ross, 2001; Kline, 1998). Even using these liberal cutoffs, the fit statistics do not support any of the models tested. However, it should be recognized that fitting even strong theoretical models can be a problem when a relatively high number of indicators are used for each latent variable (Floyd & Widaman, 1995; Kishton & Widaman, 1994; Quintana & Maxwell, 1999). It may have been the case that model fit would have been improved had the 25 items been parceled to create fewer indicants. However, as tested, none of the models provided acceptable fit for the ESS data.

Table 5
*Goodness-of-Fit Indicators for Confirmatory Factor Analysis
 for ESS*

Model	GFI	chi-sq	df	CFI	RMSEA
One-factor	.62	1361.1***	275	.62	.14
Three-factor	.72	896.2***	272	.78	.11
Null	.22			0	.22

Note. *** $p < .001$

Figure 1
ESS 3-Factor CFA Model



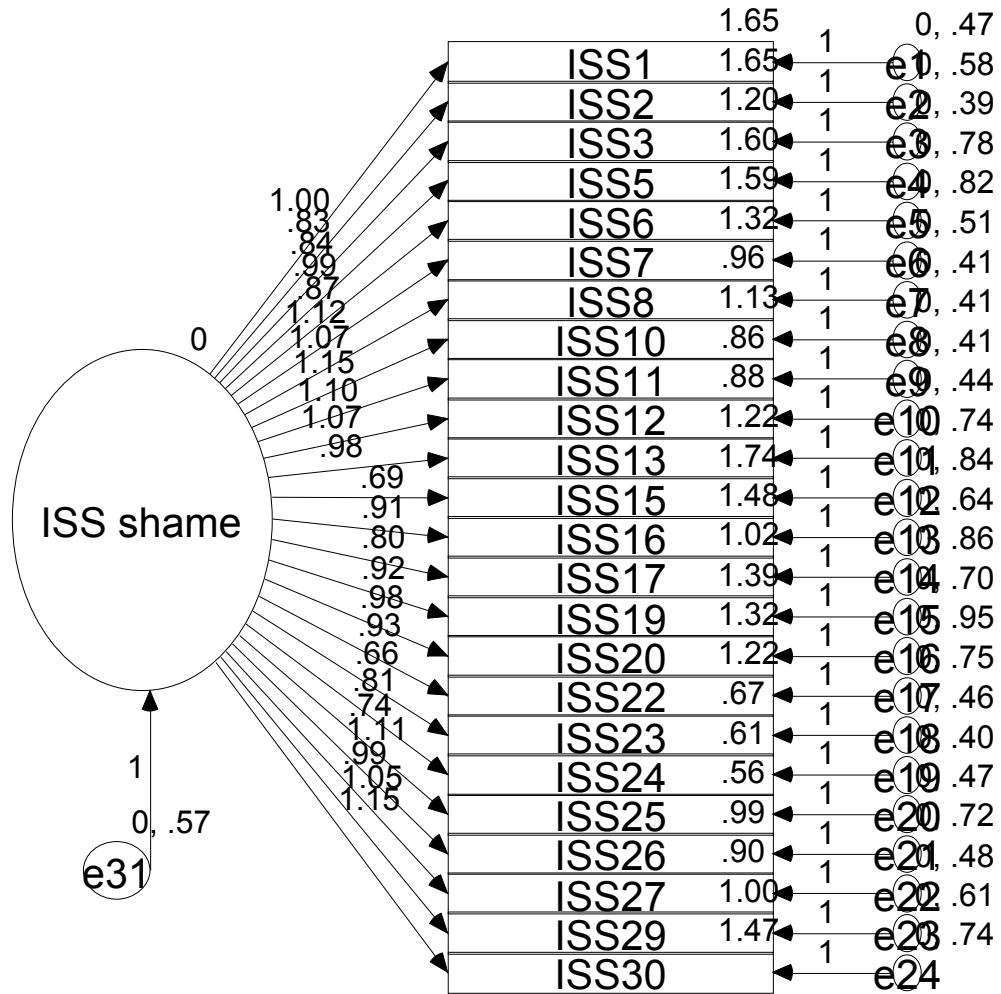
For the ISS, we first examined a two-factor model, which can be presumed based on the two subscales of the questionnaire, shame and self-esteem, using all 30 items as indicants. The second model tested was a one-factor model using the 24 shame subscale items as indicants. These models were compared to a null model presuming no structure. ISS CFA fit indices are displayed in Table 6. No model provided acceptable fit for the data. In both the two-factor and one-factor models, all measured variables loaded on the expected factor at a statistically significant level, $p < .05$. As with the ESS, however, the chi-square was significant and the fit indices failed to provide sufficient support for the models tested.

Table 6
*Goodness-of-Fit Indicators for Confirmatory Factor Analysis
 for ISS*

Model	GFI	chi-sq	df	CFI	RMSEA
One-factor shame subscale	.72	821.1***	252	.81	.11
Two-factor	.72	1068.2***	404	.83	.09
Null	.14			0	.21

Note. *** $p < .001$

Figure 2
ISS 1-Factor CFA Model



Finally, three CFAs were conducted on the TOSCA data to compare shame as a single factor, shame and guilt as two separate factors, and shame and guilt as one factor. The three CFA models of the TOSCA were a one-factor model of shame, a one-factor model of shame and guilt combined as a single self-conscious emotion factor, and a two-factor model of shame and guilt separate, which are shown in Table 7. The model with the best fit was the one-factor model of shame shown in Figure 3. In this model, all measured variables loaded on the expected factor at a statistically significant level, $p < .05$. The chi-square was significant, however, and the fit indices failed to provide sufficient support for the hypothesized one-factor model of shame.

Table 7
*Goodness-of-Fit Indicators for Confirmatory Factor Analysis
 for TOSCA*

Model	GFI	chi-sq	df	CFI	RMSEA
One-factor shame subscale	N/A	180***	104	.82	.06
Two-factor	.79	838***	463	.64	.06
Null	.14			0	.10

Note. *** $p < .001$

Figure 3
TOSCA 1-FactorShame only CFA Model



IV. DISCUSSION

As predicted based on previous research (e.g., Ferguson & Crowley, 1997; Lutwak & Ferrari, 1996), the present study found gender differences in self-reported shame and guilt. Women reported higher levels of shame and guilt than men; however, contrary to predictions, significant differences were only noted when these emotions were measured with the TOSCA. Interestingly, most previous studies citing a gender difference used the TOSCA, or its earlier form, the SCAAI, to measure guilt and shame. Recently, there has been a trend in the literature suggesting that shame and guilt gender differences are due to gender role ideals and norms (Efthim, Kenny, & Mahalik, 2001) and gender role socialization for women (Gross & Hansen, 2000). Consistent with these views, our results may reflect responses to the situational measure, the TOSCA, in a gender stereotypic manner. It is notable that the ISS and ESS, personal attribute measures of shame, did not reveal different levels of shame between men and women.

The difference in gender variation between measure type may be due in part to social desirability. Most of the items for the ESS and ISS are general and may be easily applicable to either gender. However, the TOSCA items present very specific situations in which the participant must picture him/herself reacting. Male participants may have been less likely to endorse the sometimes passive TOSCA shame items (e.g., “You would feel like you wanted to hide,” “You would think: ‘I’m inconsiderate’”). Additionally, because the respondent rates the likelihood of four or five responses to each situation,

they are likely working to be internally consistent within each situation item set. For example, the males may be more likely to endorse masculine responses, such as those offered by the externalizing items, and hence would give complementarily low probability ratings for the shame responses, regardless of their actual likelihood to perform the response. Such a gender difference might not occur if the sets of responses were not in “competition” with one another. For example, it would be interesting to see whether a gender difference occurred if participants were only asked to rate the probability of reacting to a situation with a single response, the shame response.

As hypothesized, the scores on all three shame measures were related to one another; the current research is the first to compare all three of these measures. The only previous study to compare a subset of the current shame measures was an investigation of the construct validity of the ESS. Andrews and colleagues (2002) found that the ESS total score was significantly associated with TOSCA Shame. The current study replicates this association. In regards to a comparison between the ISS and TOSCA, there have not yet been any published studies reporting correlations of these two measures.

The ESS and ISS have not been previously compared in the literature. In the present research these two measures were more closely related to one another than to the TOSCA Shame scale. This was likely due to a difference in measure type, as the ISS and ESS are both personal attribute measures that tap the degree to which a participant identifies with characteristics related to shame. The ISS and ESS may speak to participants’ ability to recognize their own feelings and interpret their past actions related to shame, making judgments about their emotional style. The TOSCA, on the other hand,

requires participants to put themselves in a particular situation and then evaluate how they would behave, think, and feel.

The three shame measures were also compared to measures of two divergent emotions, guilt and anger. When applying Campbell and Fiske's (1959) classic multitrait/multimethod matrix, none of the shame measures provided a clear case for construct validity. Campbell and Fiske (1959) suggested that to meet the parameters for divergent validity the correlations within a trait must be higher than those for a particular method. Therefore, it is preferable to have fairly low interscale correlations within each method type. This is not the case in the current study. Since guilt is often categorized with shame as a self-conscious emotion, moderate correlations between the two emotion measure types were expected. Instead, our guilt-shame correlation findings followed the measure type division. The TRGI, a personal attribute measure, was only significantly related to the two personal attribute shame measures, the ESS and ISS. TOSCA Guilt, on the other hand, was only significantly correlated with TOSCA Shame.

Ferguson and Crowley (1997b) also reported such method variance in another measure of shame and guilt. They found higher correlations between the PFQ-2 (Harder & Zalma, 1990) Shame and Guilt scales than between the PFQ-2 Shame scale and the TOSCA Shame scale. However, it may not only be a function of method variance, but also of the TOSCA itself. The presentation and factor model of the TOSCA may be sufficiently different from those of other measures as to suggest a relatively distinct measure of shame.

Interestingly, our findings and those of Ferguson and Crowley (1997b) of associations among guilt and shame based on measure type do not match the findings

reported by two previous investigations. Kubany and colleagues (1996) reported a significant association between the TRGI and TOSCA Shame, which was not replicated in the present study. Further, Andrews et al. (2002) found a significant correlation between the ESS and the TOSCA Guilt scale, which we did not replicate. Although Andrews et al. reported somewhat different findings than the current study, the two are similar in terms of the relative strength of associations reported. Andrews et al. reported a weaker association between the ESS and TOSCA Shame than the one reported between TOSCA Guilt and TOSCA Shame, reflecting the same kind of measure type findings in the present study. However, since both Andrews et al. and the current study employed only a single scenario-based measure of shame and guilt, the TOSCA, the strength of conclusions about measure type must be tempered by the fact that these two emotion measures are subscales of the same instrument, likely inflating their association. Strong associations between TOSCA Shame and a different scenario-based guilt measure would strengthen measure type conclusions.

Recent research and theory in the emotion literature has also begun to suggest associations between shame and anger (e.g., Nathanson, 1987, Retzinger, 1987; Scheff, 1987). Consistent with previous research (Tangney et al., 1992a; Fontaine et al., 2001), the present research found significant associations between measures of shame and trait anger. Such correlations across measure type are likely due to the co-existence of shame and anger and the fact that they are often elicited by similar situations. Additionally, it is thought that shame and anger can result from one another. Tangney et al. (1992a) suggest that anger is created in some circumstances to deal with the pain and extreme negative self-evaluation associated with shame. With the exception of the Tangney et al.,

(1992a) study, the current study was the first to compare responses on the STAXI, a personal attribute measure of anger, to any of the three shame measures. Although the STAXI was significantly associated with all three shame measures, it was correlated at a much higher level with the two personal attribute measures, consistent with our measure type pattern of findings.

Considering that shame, guilt, and anger are all negative emotions, it was expected that there would be some significant associations among them. Nonetheless, it was predicted that the correlations among the shame measures would be higher than their associations with measures of other emotions, suggesting some degree of convergent and discriminant validity in a multitrait/multimethod matrix. However, this was not the case. Associations between shame and divergent emotion measures were in some cases greater than or equal to the strength of associations between shame measures. For example, the TOSCA Shame scale was somewhat more closely related to TOSCA Guilt than to the ESS. Such results do not reflect clear divergent validity because associations tended to be somewhat higher among the same measure type rather than among those instruments measuring the same emotion.

Due to shame's status as a negative emotion associated with negative self-evaluation, it was hypothesized that it would be correlated with various types of psychopathology. As predicted, PTSD was significantly related to the two personal attribute shame measures. Surprisingly, it was not associated with the TOSCA Shame scale. Also, all of the emotion measures except TOSCA Shame were independently associated with the PCL scores. This exception of the TOSCA Shame was unexpected and cannot be explained by method variance because the TOSCA Guilt scale was

independently associated with the PCL. Interestingly, the ISS was more strongly associated with PTSD than the ESS. This difference may be due to the fact that the ESS was created and normed on college students, whereas the ISS was normed on inpatient and outpatient populations in addition to college students.

All three shame measures were significantly associated with the depression measure, similar to results reported by Andrews and colleagues (2002). The ISS was particularly strongly related to the MASQ Anhedonic Depression scale. One likely explanation for this strong association may be that the ISS consists of negative self-statements (e.g., “I see myself striving for perfection only to continually fall short”) likely tapping into hopelessness and low self-esteem associated with depression. Another possibility is that the ISS items, and to some extent those of the ESS, represent more severe, pathological aspects of shame that are more likely to be associated with psychopathology relative to the TOSCA items. However, it was surprising that the ESS was not independently associated with the Anhedonic Depression scale considering that Andrews and colleagues (2002) found discriminate results when they conducted a multiple regression analysis for depressive symptoms with ESS and TOSCA Shame as predictor variables.

Associations of shame and anxious arousal have not been previously examined in the shame literature. The current study demonstrated that the ISS and ESS are significantly associated with MASQ Anxious Arousal, whereas TOSCA Shame is not. The ISS was the only shame measure to be independently associated with Anxious Arousal. This difference is likely in part due to measure type; the MASQ is largely a personal attribute measure. Additionally, as with our depression findings, the ISS may be

tapping into a more severe form of shame than the TOSCA Shame scale. Furthermore, when completing the TOSCA, respondents must rate their likelihood to perform four or five behaviors for each scenario. During this process, they may be comparing the responses and be reluctant to identify strongly with the most negative, pathological response which is usually shame.

No significant associations were found between drug and alcohol use and shame. Our failure to find such associations may be due, at least in part, to low power resulting from the low substance use base rates in our sample. A very small percentage of participants in our sample reported using any substance besides alcohol and marijuana, which is comparable to the national averages reported for college students (O'Malley & Johnston, 2002).

The factor analyses performed on the three shame measures indicated that none of the models tested provided an adequate fit of the data. Given that all three measures include a relatively high number of items, it was not entirely surprising that model fit was not acceptable. It would be enlightening to perform similar analyses with items parcels to create fewer indicants. Further, the model chi-squares were all significant, supporting the null models, although this was likely due to the large sample size. Nonetheless, the present analyses do provide some information regarding the relative fit of various models for the data from the three measures. The three-factor model of the ESS, as suggested by Andrews and colleagues (2002), provided the best fit. Similarly, in the Andrews and colleagues (2002) study this model provided a marked improvement over the independence model, but did not represent a perfect model for the data either as the chi-square was significant in their study as well. Based on our results, it appears that such a

division of their shame items into subscales may be useful, as it represented a better fit than a one-factor model lumping all items together. It would be beneficial to the understanding of shame as a construct to further examine these divisions.

The ISS one-factor and two-factor models had comparable fit, suggesting that the self-esteem subscale may be superfluous when measuring shame, adding nothing to the model. There were no CFAs examining the structure of the ISS in the literature to which the current study's analyses could be compared.

Previous research has examined the factor structure of the TOSCA. In the current study, the one-factor model of the Shame scale alone was the best fit of the data relative to the two-factor model of the Shame and Guilt scales as separate factors and a model combining these subscales into a single self-conscious emotion latent variable. These results may indicate that shame as measured by the TOSCA is clouded by the presence of guilt items. Fontaine et al.'s (2001) report that the TOSCA Guilt items had much overlap with shame is supported by our findings. Fontaine and colleagues reported numerous items that did not discriminate guilt from shame and instead referred to a range of self-conscious experiences, such as avoiding eye contact, feeling isolated, and counter-factual thinking. These items may be better accounted for by shame due to its global characterological focus, but the findings of the present study could not confirm this.

Many of the limitations of the current study are related to the sample used. It was comprised of college students in psychology courses at a southern university. The majority of the participants were white females. Therefore, the data may not generalize to all populations. Further, this weakens our examination of gender differences. It may be the case that we did not find gender differences on the ISS and ESS because we had a

relatively small male sample. In the future, researchers may benefit from extending this protocol to more diverse populations, including inpatient and outpatient populations, to further examine the measurement of shame and the role that shame plays in psychopathology. Another limitation of this study was its self-report format. Future research would benefit from the inclusion of alternative measures of emotion and psychopathology, such as an interview or behavioral observation. A third limitation of the current study is related to the numerous emotion questionnaires, many of which were related to self-conscious emotion, that the participants completed in one sitting. This method may have led to increased shared variance and participants may have gotten into a response set. However, this makes our findings of the distinctiveness of the personal attribute and scenario-based shame measures even more impressive, since this method would have been expected to inflate their associations and make divergent validity more difficult to observe.

This study investigated three measures of shame and their associations with each other, as well as their discriminant validity as demonstrated by correlations with anger and guilt measures. Construct validity was not clearly demonstrated due to a method effect. Although all three measures are based on a similar characterological, negative self-view definition of shame, it appears that current shame self-report instruments may not be measuring the same construct. The ESS and ISS may be assessing the same trait, but their association may equally well be attributed to method variance. Additionally, this may be the result of the inherent difficulty of attempting to measure an emotion that by its very nature participants wish to conceal and avoid. Such a hypothesis may be partially supported by the fact that none of the factor structure models for the three shame

measure were an adequate fit of the data. This study also explored the relationships of these shame measures with forms of psychopathology and thus their clinical utility. The ISS and ESS were most closely related to forms of psychopathology.

V. REFERENCES

- Allan, S. Gilbert, P., & Goss, K. (1994). An exploration of shame measures-II Psychopathology. *Personality and Individual Differences, 17*, 719-722.
- Andrews, B. (1995). Bodily shame as a mediator between abusive experiences and depression. *Journal of Abnormal Psychology, 104*, 669-675.
- Andrews, B., Brewin, C.R., Rose, S., & Kirk, M. (2000). Predicting PTSD symptoms in victims of violent crime: The role of shame, anger, and childhood abuse. *Journal of Abnormal Psychology, 109*, 69-73.
- Andrews, B., & Brown, G.W. (1993). Self-esteem and vulnerability to depression: The concurrent validity of interview and questionnaire measures. *Journal of Abnormal Psychology, 102*, 565-572.
- Andrews, B. & Hunter, E. (1997). Shame, early abuse and course of depression in a clinical sample: A preliminary study. *Cognition and Emotion, 11*, 373-381.
- Andrews, B., Qian, M., Valentine, J. (2002). Predicting depressive symptoms with a new measure of shame: The Experience of Shame Scale. *British Journal of Clinical Psychology, 41*, 29-42.
- Antonovsky A. (1987). *Unraveling the mystery of health. How people manage and stay well*. San Francisco: Jossey-Bass.
- Antonovsky, A. (1993). The structure and properties of the sense of coherence scale. *Social Science Medicine, 36*, 725-733.

- Beck, A.T. & Steer, R.A. (1987). *Beck depression inventory manual*. San Antonio, TX: The Psychological Corporation.
- Blake, D.D., Weathers, F.W., Nagy, L.M., Kaloupek, D.G., Gusman, F.D., Charney, D.S., & Keane, T.M. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress, 8*, 75-90.
- Blanchard, E. B, Hickling, E. J., Taylor, A. E., Forneris, C. A., et al. (1995). Effects of varying scoring rules of the Clinician-Administered PTSD Scale (CAPS) for the diagnosis of post- traumatic stress disorder in motor vehicle accident victims. *Behavior Research & Therapy, 33*, 471-475.
- Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behavior Research & Therapy, 34*, 669-673.
- Breslau, N., Davis, G.C., Andreski, P., & Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry, 48*, 216-222.
- Campbell, D. & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 105*, 456-466.
- Cohen, D. (2001). Cultural variation: Considerations and implications. *Psychological Bulletin, 127*, 451-471.
- Cohen, D., Vandello, J., Puente, S., & Rantilla, A. (1999). “When you call me that, smile!” How norms for politeness, interaction styles, and aggression work together in Southern culture. *Social Psychology Quarterly, 62*, 257-275.
- Cook, D. (1989). *The Internalized Shame Scale*. University of Wisconsin, Stout.

- Cook, D. (1994). *The Internalized Shame Scale*. New York: Multi-Health Systems Inc.
- Cook, D. (1996). Empirical studies of shame and guilt: The internalized shame scale. In D. Nathanson (Ed.), *Knowing Feeling* (pp. 132-165). New York: Norton.
- Cook, D. & Wong, M. (1993). Shame and its contribution to PTSD. *Journal of Traumatic Stress, 5*, 557-562.
- Damon, W. (1988). *The moral child: Nurturing children's natural moral growth*. New York: Free Press.
- Derogatis, L.R., Lipman, R.S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale- Preliminary report. *Psychopharmacology Bulletin, 9*, 13-28.
- Fava, J. L., Velicer, W. F., & Prochaska, J. O. (1995). Applying the trans-theoretical model to a representative sample of smokers. *Addictive Behaviors, 20*, 189-203.
- Ferguson, T. & Crowley, S. (1997a). Gender differences in the organization of guilt and shame. *Sex Roles, 37*, 19-44.
- Ferguson, T. & Crowley, S. (1997b). Measure for measure: A multitrait-multimethod analysis of guilt and shame. *Journal of Personality Assessment, 69*, 425-441.
- Gilbert, P., Pehl, J., & Allan, S. (1994). The phenomenology of shame and guilt: An empirical investigation. *British Journal of Medical Psychology, 67*, 23-36.
- Goldberg, D. & Williams, P. (1991). *A user's guide to the General Health Questionnaire*. Windsor, England: NFER-Nelson.
- Gomberg, E. L. (1989). Shame and guilt issues among women alcoholics. *Alcoholism Treatment Quarterly, 4*, 139-155.
- Goss, K., Gilbert, P., & Allan, S. (1994). An exploration of shame measures-I: The Other As Shamer Scale. *Personality and Individual Differences, 17*, 713-717.

- Gramzow, R. & Tangney, J.P. (1992). Proneness to shame and the narcissistic personality. *Personality & Social Psychology Bulletin*, *18*, 369-376.
- Harder, D.W. & Lewis, S.J. (1987). The assessment of shame and guilt. In J.N. Butcher & C.D. Spielberger (Eds.), *Advances in Personality assessment* (pp 89-114). Hillsdale NJ: Lawrence Erlbaum Associates.
- Harder, D. & Zalma A. (1990). Two promising shame and guilt scales: A construct validity comparison. *Journal of Personality Assessment*, *55*, 729-745.
- Izard, C. (1977). *Human Emotions*. New York: Plenum Press.
- Janikowski, T. & Glover, N. (1995). Incest and substance abuse: Implications for treatment professionals. *Journal of Substance abuse Treatment*, *11*, 177-183.
- Janoff-Bulman, R. (1979). Characterological versus behavioral self-blame: Inquiries into depression and rape. *Journal of Personality and Social Psychology*, *37*, 1798-1809.
- Keltner, D. & Buswell, B. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition & Emotion*, *10*, 155-171.
- Kubany, E., Haynes, S., Abueg, F., Manke, F., Brennan, J., & Stahura, C. (1996). Development and Validation of the Trauma-Related Guilt Inventory (TRGI). *Psychological Assessment*, *8*, 428-444.
- Leskela, J., Dieperink, M., & Thuras, P. (2002). Shame and Posttraumatic Stress Disorder. *Journal of Traumatic Stress*, *15*, 223-226.
- Lewis, H.B. (1971). *Shame and guilt in neurosis*. New York: International Universities Press.

- Lindsay-Hartz, J. (1984). Contrasting experiences of shame and guilt. *American Behavioral Scientist*, 27, 689-704.
- Lisak, D. (1994). The psychological impact of sexual abuse: Content analysis of interviews with male survivors. *Journal of Traumatic Stress*, 7, 525-548.
- Lutwak, N. & Ferrari, J. (1996). Moral affect and cognitive processes: Differentiating shame from guilt among men and women. *Personality and Individual Differences*, 21, 891-896.
- Luyten, P., Fontaine, 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*, 3, 1373-1387.
- Nathanson, D.L. (1987). A timetable for shame. In D.L. Nathanson (Ed.), *The many faces of shame* (pp 1-63). New York: Guilford.
- O'Malley, P. & Johnston, L. (2002). Epidemiology of alcohol and other drug use among American college students. *Journal of Studies on Alcohol*, 14, 23-39.
- Patrick, D.L., Cheadle, A., Thompson, D.C., Diehr, P., Koepsell, T., Kinne, S. (1994). The validity of self-reported smoking: A review and meta-analysis. *American Journal of Public Health*, 84, 1086-1093.
- Retzinger, S. (1987). Resentment and laughter: Video studies of the shame-rage spiral. In H. Lewis (Ed). *The role of shame in symptom formation*. (pp. 151-181).
- Roseman, I.J., Antoniou, A.A., & Jose, P.E. (1992). Appraisal determinants of emotions: Revising current theories. Unpublished Manuscript. Univeristy of California at Los Angeles.

- Roseman, I.J., Weist, C., & Swartz T.S. (1994). Phenomenology, behaviors, and goals differentiate discrete emotions. *Journal of Personality and Social Psychology*, 67, 206-221.
- Scheff, T. (1987). The shame-rage spiral: A case study of an interminable quarrel. In H. Lewis (Ed). *The role of shame in symptom formation*. (pp. 109-149).
- Sobell, L.C., Toneatto, T., & Sobell, M.B. (1994) Behavioral assessment and treatment planning for alcohol, tobacco, and other drug problems: Current status with an emphasis on clinical applications. *Behavior Therapy*, 25, 533-580.
- Spielberger, C. D. (1999). *Professional Manual for the State-Trait Anger Expression Inventory-2 (STAXI-2)*. Lutz, FL: Psychological Assessment Resources.
- Spielberger, C.D., Jacobs, G., Russell, S., & Crane, R.S. (1983). Assessment of anger: The State-Trait Anger Scale. In J. N. Butcher & C.D. Spielberger (Eds.), *Advances in personality assessment* (pp. 161-189). Hillsdale, NJ: Erlbaum.
- Spitzer, R. L., Williams, J. B., Gibbon, M., & First, M. B. (1990). *User's guide for the structured clinical interview for DSM-III-R: SCID*. New York, NY: Columbia University.
- Tangney, J.P. (1992). Situational determinants of shame and guilt in young adulthood. *Personality and Social Psychology Bulletin*, 18, 199-206.
- Tangney, J. (1995). Recent advances in the empirical study of shame and guilt. *American Behavioral Scientist*, 38, 1132-1145.
- Tangney, J.P. (1996). Conceptual and methodological issues in the assessment of shame and guilt. *Behavioral Therapy and research*, 34, 741-754.

- Tangney, J.P. (1999). The self-conscious emotions: Shame, guilt, embarrassment and pride. In T. Dalgleish, & M. Power (Eds). *Handbook of cognition and emotion*. (pp. 541-568).
- Tangney, J.P., Burggaf, S.A., Hamme, H., & Domingos, B. (1988). Assessing individual differences in proneness to shame and guilt: The Self-Conscious Affect and Attribution Inventory. Poster presented at the Annual meeting of the Eastern Psychological Association, Buffalo, NY.
- Tangney, J.P., Dearing, R., Wagner, P.E., & Gramzow, R. (2000). *The Test of Self-Conscious Affect – 3 (TOSCA-3)*. George Mason University, Fairfax VA.
- Tangney, J.P., Miller, R.S., Flicker, L., & Barlow, D.H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *70*, 1256-1269.
- Tangney, J.P., Wagner, P., Fletcher, C., & Gramzow, R. (1992a). Shame into anger? The relation of shame and guilt to anger and self-reported aggression. *Journal of Personality and Social Psychology*, *62*, 669-675.
- Tangney, J.P., Wagner, P., & Gramzow, R. (1989). *The Test of Self-Conscious Affect*. Fairfax, VA: George Mason University.
- Tangney, J.P., Wagner, P., & Gramzow, R. (1992b). Proneness to shame, proneness to guilt and psychopathology. *Journal of Abnormal Psychology*, *101*, 469-478.
- Tomkins, S.S. (1984). Affect theory. In D.R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 163-195). Hillsdale, NJ: Erlbaum.
- Vrana, S., & Lauterbach, D. (1994). Prevalence of traumatic events and post-traumatic

- psychological symptoms in a nonclinical samples of college students. *Journal of Traumatic Stress, 7*, 289-302.
- Watson, D. & Clark, L.A. (1991). The Mood and Anxiety Symptoms Questionnaire. Unpublished Manuscript, University of Iowa, Department of Psychology, Iowa City.
- Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993). The PTSD checklist: reliability, validity, and diagnostic utility. Paper presented at the Annual Meeting of the International Society for Traumatic Stress Studies, San Antonio, TX, October.
- Weiner, K. (1993). Internalized shame and family dysfunction in adult children of alcoholics. Doctoral Dissertation, University of Pittsburgh.
- Wong, M.R. & Cook, D. (1993). Shame and its contribution to PTSD. *Journal of Traumatic Stress, 5*, 557-562.