Shame, guilt, and personality judgment

Jo Ann Abe

Department of Psychology, Southern Connecticut State University, 501 Crescent St., New Haven, CT 06515, USA

Abstract

The purpose of this study was to shed further light on the intrapersonal and interpersonal correlates of shame- and guilt-proneness by examining their relations with self- and peer-ratings of the five-factor model (FFM) of personality. Shame- and guilt-proneness were assessed using a scenario-based and a checklist measure. Consistent with findings from previous research, the self-rating data yielded support for the view that shame-proneness may be associated with more maladaptive patterns than guilt-proneness. However, peer-ratings of personality failed to corroborate these findings. Both scenario-based and checklist measures of shame-proneness were associated with a tendency to underestimate one's Agreeableness relative to peer-ratings. Several possible interpretations of these findings are discussed in light of the extant research on personality judgment.

2003 Elsevier Science (USA). All rights reserved.

1. Introduction

Over the past decade, research on shame and guilt has flourished. A major impetus for the resurgence of interest in these emotions is that shame, which was once identified as the “sleeper in psychopathology” (Lewis, 1987), has in recent years come to be seen as “the bedrock of psychopathology” (Miller, 1996). The growing interest in shame has, in turn, led to renewed debates over the similarities and differences between shame and guilt (see Tangney & Fischer, 1995). Most researchers concur that shame and guilt are both common human emotions that everyone
experiences on occasion, and even though unpleasant, serve adaptive functions under some circumstances. However, there is considerable debate as to the implications that the proneness to experiencing these emotions has for psychological adjustment and interpersonal behaviors. Some researchers claim that although shame-proneness is associated with maladaptive patterns, guilt-proneness is associated with adaptive outcomes, especially in the interpersonal realm (e.g., Baumeister, Stillwell, & Heatherton, 1995; Tangney, 1995; Tangney, Burgraff, & Wagner, 1995). Others claim that shame- and guilt-proneness are equally maladaptive and that both of these emotions are associated with psychological symptoms and problems with interpersonal functioning (e.g., Harder, 1995; Jones, Kugler, & Adams, 1995; O’Connor, Berry, & Weiss, 1999).

The purpose of the present study was to shed further light on the intrapersonal and interpersonal correlates of shame- and guilt-proneness by examining their relations with self- and peer-ratings of the five-factor model (FFM) of personality. According to the FFM, most of the traits studied by psychologists and used by people to describe themselves and others can be captured by five broad personality dimensions: Agreeableness, Conscientiousness, Extraversion, Openness/Intellect/Intellect, and Neuroticism/Emotional Stability. The FFM was used in this study for two reasons. First, this personality taxonomy is considered by many personality psychologists to be one of the most comprehensive (e.g., Wiggins, 1996). Second, the FFM has been used extensively in previous studies comparing self vs. observer ratings of personality (e.g., Funder, 1999; Gostling, John, Craik, & Robins, 1998; John & Robins, 1993; Watson, Hubbard, & Wiese, 2000). The rationale for including both self- and peer-ratings in this study was that research on personality judgment has found that the self has unique advantages and disadvantages as a judge of personality (e.g., Funder, 1999; John & Robins, 1993, 1994). The self has access to internal thoughts and feelings and other ‘privileged information,’ which is not available to the external observer. On the other hand, the self is also more ego-involved than others in its assessment of itself and has greater difficulty viewing itself objectively. A growing body of research reveals that peer-ratings of personality may offset some of the limitations of inherent in self-ratings of personality and thereby serve as an informative complement to them.

1.1. Conceptualization and measurement of shame- and guilt-proneness

A clear understanding of the implications shame- and guilt-proneness have for intrapersonal and interpersonal functioning has been complicated by the fact that these emotions have been conceptualized and measured in different ways (Andrews, 1998; Ferguson & Crowley, 1997; Gilbert, 1998; Harder, 1995; Kugler & Jones, 1992; Tangney, 1996). One longstanding theoretical tradition, which has its origin in anthropological literature (e.g., Benedict, 1946), has maintained that the primary distinction between shame and guilt is the locus of negative evaluation (Ausbel, 1955; Buss, 1980; Harder, 1995; Hogan & Cheek, 1983; Piers and Singer, 1971; Smith, Webster, Parrot, & Eyre, 2002; Wallbott & Scherer, 1995). According to this view, guilt arises from remorse over violation of one’s internalized conscience,
whereas shame arises from imagined or actual exposure of some impropriety or shortcoming.

Another major theoretical tradition, which has recently gained ascendance, maintains that the primary distinction between shame and guilt is the focus of negative self-evaluation (e.g., Ferguson & Stegge, 1995; Lindsay-Hartz, de Rivera, & Mascaro, 1995; Miller, 1996; Tangney, 1995; Tangney et al., 1995). According to this view, guilt involves regret or remorse over specific behaviors, whereas shame involves a global condemnation of the self. Because the entire self is found wanting in the shame experience, proneness to experiencing this emotion is likely to promote intense self-focus and to motivate extremes of either “seething, resentful, and retaliative” anger toward others or withdrawal from the interpersonal context (Tangney et al., 1995). By contrast, because the self remains intact in the guilt experience, proneness to experiencing this emotion is likely to foster sympathetic concern for others and to motivate reparative behaviors.

Researchers not only differ in terms of their conceptualization of shame- and guilt-proneness, but also in their methods of assessing these emotions. Some researchers have used checklist measures, which ask respondents to make global judgments of how often they experience shame- or guilt-related affective, cognitive and/or behavioral responses in their daily life, devoid of specific contexts (Harder & Zalma, 1990; Izard, Libero, Putnam, & Haynes, 1993; Kugler & Jones, 1992). Other researchers have used scenario-based measures, which ask respondents to rate the likelihood of experiencing shame or guilt-related affective, cognitive, and/or behavioral responses in potentially shame or guilt-inducing situations (e.g., Tangney, Wagner, & Gramzow, 1989). Even though both types of measures assess proneness to experiencing shame and guilt, checklist measures are likely to tap into a more chronic and, therefore, maladaptive form of these emotions than scenario-based measures (Andrews, 1998; Ferguson & Crowley, 1997; Quiles & Bybee, 1997; Tangney, 1996). In particular, the most widely used scenario-based measure of guilt—the Test of Self-Conscious Affect (TOSCA, Tangney et al., 1989)—incorporates reparative behaviors in its conceptualization and is thus likely to assess a more adaptive form of guilt than the checklist measures of guilt-proneness.

1.2. Research on intrapersonal and interpersonal correlates of shame- and guilt-proneness

Given the diversity of conceptualizations and measures of shame- and guilt-proneness, the extant research on the intrapersonal and interpersonal correlates of these emotions has, not surprisingly, yielded somewhat inconsistent findings across different studies. In general, studies using scenario-based measures have found that guilt-proneness is weakly associated with measures of psychological symptoms (Tangney, 1991; Tangney, Wagner, & Gramzow, 1992b). Guilt-proneness as assessed by scenario-based measures has been positively correlated with empathy and perspective-taking (Leith & Baumesiter, 1998; Tangney, 1991) as well as adaptive strategies of dealing with interpersonal conflicts (Tangney, Wagner, Barlow, Marschall, & Gramzow, 1996). However, in contrast to studies using scenario-based measures,
studies using global checklist measures have found robust relations between guilt-proneness and various psychological symptoms as well as interpersonal problems such as loneliness and dissatisfaction with relationships (Harder, Cutler, & Rockart, 1992; Jones & Kugler, 1993; O’Connor et al., 1999).

Compared to the research on guilt, research on shame-proneness has generally yielded more consistent findings across the different studies. Shame-proneness as assessed by different types of measures has been robustly associated with intense self-focus, a strong tendency towards self-derogation, as well as with a host of psychological symptoms (Harder, 1995; Hoblitzelle, 1987; Tangney, Wagner, Fletcher, & Gramzow, 1992a, 1992b, 1995; Wicker, Payne, & Morgan, 1983). In terms of its interpersonal correlates, shame-proneness has been negatively correlated with empathy (Leith & Baumesiter, 1998; Tangney, 1991) and positively correlated with the use of maladaptive strategies of dealing with interpersonal conflicts (Tangney et al., 1996).

In keeping with the foregoing findings, previous research using self-ratings of the FFM of personality has found that measures of shame- and guilt-proneness generally show the most robust relations with the personality dimensions of Agreeableness, Extraversion, and Neuroticism (Abe, Abe, & Huebner, 1999; Einstein & Lanning, 1998; Harder & Greenwald, 1999). These studies have consistently found that both scenario-based and checklist measures of shame- and guilt-proneness are positively correlated with Neuroticism. Scenario-based measures of guilt were also consistently found to be positively correlated with Agreeableness, whereas checklist measures of shame were found to be negatively correlated with Agreeableness (Abe et al., 1999; Einstein & Lanning, 1998; Harder & Greenwald, 1999). Both checklist measures of guilt and measures of shame-proneness have been negatively correlated with Extraversion (Abe et al., 1999; Harder & Greenwald, 1999). In some of the studies, checklist measures of shame-proneness were also found to be inversely related with Conscientiousness and Openness/Intellect, but these findings have been less consistent across the different studies.

1.3. Present study

To gain a better understanding of the intrapersonal and interpersonal correlates of shame- and guilt-proneness, this study investigated their relations with self- and peer-ratings of the FFM. In the present study, shame- and guilt-proneness were assessed using the Test of Self-Conscious Affect (TOSCA; Tangney et al., 1989) and the guilt and shame subscales of the Differential Emotions Scale (DES-IV; Izard et al., 1993). The TOSCA is a scenario-based measure that is based on the view that shame involves a global condemnation of the entire self, whereas guilt involves remorse or regret over specific behaviors. The DES is a checklist measure of discrete emotions that incorporates aspects of both the external vs. internal as well as the global vs. specific distinctions between shame and guilt in its conceptualization of these emotions. Self- and peer-ratings of personality were obtained using the Goldberg Five Factor Markers (Goldberg, 1992). The self- and peer-ratings of personality were also used to investigate whether the measures of shame- and guilt-proneness were
associated with distinct patterns of discrepancies between self- vs. peer-ratings of personality dimensions.

Based on theory and past research, this study hypothesized that the measures of shame- and guilt-proneness would show the most robust relations with self-ratings of Agreeableness, Extraversion, and Emotional Stability. All of the shame and guilt subscales were expected to be negatively correlated with Emotional Stability. However, TOSCA guilt was predicted to be positively correlated with self-ratings of Agreeableness, whereas DES shame was predicted to be negatively correlated with self-ratings of Agreeableness. The measures of shame-proneness and DES guilt were also expected to be inversely related with self-ratings of Extraversion.

Given that the bulk of the extant research on shame- and guilt-proneness has been based almost exclusively on self-report measures, the predictions involving the relations between shame and guilt with self vs. peer discrepancies in ratings of personality are offered somewhat speculatively. TOSCA guilt taps into adaptive reactions to daily transgressions and has been positively correlated with social desirability bias (Tangney et al., 1996). TOSCA guilt was, therefore, expected to be associated with a tendency to overestimate oneself on highly evaluative traits such as Agreeableness relative to peer-ratings. By contrast, shame-proneness has been associated with a tendency toward self-focus and self-derogation, which in turn, have been associated with a tendency to underestimate oneself on various characteristics (Carver & Scheier, 1981; Swann, Wenzlaff, Krull, & Pelham, 1992). Shame-proneness was, therefore, expected to be associated with a tendency to underestimate oneself on various personality dimensions relative to peer-ratings. Some support for the latter prediction comes from a recent study on emotion expressivity, which found that self-reports of shame are positively correlated with peer-ratings of Agreeableness, but not with self-ratings of Agreeableness (Trieweiler, Eid, & Lischetzke, 2002).

2. Method

2.1. Participants

The participants in this study were 97 students enrolled in psychology classes at a state university. The majority of the students (82%) were between the ages of 18 and 26 and 24% of the sample was male. In terms of ethnic composition, 90% were White, 4% were African-American, and 6% were Latino.

2.2. Procedure

The participants were asked to complete self-ratings of their emotion and personality characteristics and to also refer a friend who knows them well and who would be willing to rate the participant’s personality traits. Peer-ratings of 44 participants were obtained by mailing the personality questionnaire directly to the friend with a cover letter explaining the purpose of the study and asking the rater to mail back the questionnaire directly to the experimenter in a pre-addressed, stamped envelope.
The peer-ratings for the remaining 43 participants were collected toward the end of a semester. To ensure that the questionnaires would be received before the semester ends, the participants were asked to give an envelope containing the cover letter and questionnaire to a friend. The raters were allowed the option of mailing back the questionnaire to the experimenter or placing it in a preaddressed, sealed envelope, and signing over the back flap and giving it to the participant to return to the experimenter. The response rate from the peers in the first sample was approximately 70% and the response rate in the second sample was 100%. Only participants with both self- and peer-ratings of personality were included in the study. The peer-ratings of personality obtained using the different methods did not systematically differ either in terms their of mean levels or their pattern of relations with the shame and guilt measures and thus the two samples were combined. Both groups of raters were asked to provide their name, address, and phone number and to sign an informed consent form. They were informed that the information they provide would not be disclosed to anyone and instructed not to discuss their ratings with their friends. The number of years that the participant and raters knew one another ranged from 6 months to 20 years, with a mean of 8 years and a median of 5 years. Students received course credit for participating in the study. One of the participants responded randomly to the questionnaires and was thus eliminated from the study.

3. Measures

3.1. Test of self-conscious affect (TOSCA, Tangney et al., 1989)

The TOSCA is a widely used scenario-based measure of shame and guilt. It consists of a series of brief scenarios (10 negative and 5 positive) and associated affective, cognitive, and behavioral responses and which are designed to assess shame- and guilt-proneness. A sample scenario for the TOSCA is “You are out with a group of friends and you make fun of a friend who is not there.” A sample shame response is: “You would feel small...like a rat.” A sample guilt response is: “You would apologize and talk about the person’s good traits.” In this study, the alphas for the shame and guilt subscales were .80 and .61, respectively. The correlation between the subscales was .36.

3.2. Differential emotions scale (DES-IV; Izard et al., 1993)

The DES is a checklist measure of discrete emotions, which asks respondents to rate their emotion experiences on a 5-point Likert scale ranging from “Never or Hardly Ever” to “Very Often.” The shame and guilt subscales consist of three items each. A sample guilt item is: “Feel like you ought to be blamed for something you did.” A sample shame item is “Feel like people always look at you whenever anything goes wrong.” The alpha reliabilities for the shame and guilt subscales in this sample were .87 and .81, respectively. The correlation between the subscales was .57.
Two previous studies (Abe et al., 1999; Abe, Abe, Sakamoto, Takahashi, & Kawashima, 1998) have found that the DES shame and guilt subscales show a moderate to high level of convergence with corresponding subscales of the Personal Feelings Questionnaire (PFQ-2; Harder & Zalma, 1990). The major advantage that the DES offers over the PFQ is that the shame and guilt subscales of the DES tend to be less highly intercorrelated with one another than the PFQ subscales.¹

3.3. Five Factor Markers (Goldberg, 1992)

The Five Factor Markers consists of a list of 100 trait adjectives and asks respondents to rate on a scale of 1–9 how characteristic each of the traits is of a person. Previous research reveals that the five markers show a high degree of convergence with the corresponding dimensions of the NEO-PI dimensions. In this study, the reliabilities for the Agreeableness, Conscientiousness, Extraversion, Openness/Intellect, and Emotional Stability scales ranged from .81 to .90 for the self-ratings, and .87 to .92 for peer-ratings of personality traits.

4. Results

4.1. Preliminary analyses

4.1.1. Relations between TOSCA and DES subscales

Consistent with findings from previous research that checklist measures of guilt-proneness tend to show a higher level of convergence with measures of shame-proneness than with scenario-based measures of guilt-proneness, DES guilt was more strongly correlated with TOSCA shame ($r = .52, p < .01$, two-tailed) than TOSCA guilt ($r = .31, p < .05$, two-tailed). Also in keeping with previous research, DES shame and TOSCA shame were moderately correlated with one another ($r = .50, p < .01$, two-tailed), but DES shame was not correlated with TOSCA guilt ($r = .01, n.s.$).

4.1.2. Relations between self- and peer-ratings of personality

Self- and peer-ratings of all of the personality dimensions were significantly correlated with one another and ranged from .20 for Agreeableness to .36 for Emotional Stability ($ps < .05$ to .01, two-tailed). The magnitude of the correlations for some of the personality dimensions were somewhat lower than those reported in previous studies using the NEO-PI (Watson et al., 2000). The finding that the correlation between self-ratings and peer-ratings was the lowest for Agreeableness is consistent with previous reports that self–peer agreement tends to be lowest for highly evaluative traits (Funder & Colvin, 1998; Jensen-Campbell & Graziano, 2001; John &

¹ In three previous studies, the correlations between the PFQ-2 shame and guilt subscales exceeded .70 (Abe et al., 1999; Abe et al., 1998; Averill, Diefenbach, Stanley, Breckenridge, & Lusby, 2002).
Robins, 1993). However, the finding that self–peer agreement was highest for Emotional Stability is somewhat at variance with previous reports that self–peer agreement typically tends to be higher for visible traits such as Extraversion.

Pairwise comparisons revealed mean levels of self- and peer-ratings of Agreeableness, Conscientiousness, and Emotional Stability did not significantly differ ($t_s = 0.03, 0.57, 0.58$, respectively). However, the mean level of peer-ratings of Openness/Intellect was significantly higher than self-ratings ($t = 2.04$, $p < .05$, two-tailed) and peer-ratings of Extraversion showed a trend toward being higher than self-ratings ($t = 1.86$, $p < .07$, two-tailed, $d = .23$).

### 4.2. Main analyses

The bivariate correlations between the shame and guilt subscales with the self-ratings of personality, peer-ratings of personality, and self-criterion residuals are shown in Table 1.

#### 4.2.1. Relations between shame and guilt with self-ratings of personality

The patterns of relations between the shame and guilt subscales with self-ratings of personality dimensions were similar to those obtained in previous studies. TOSCA guilt was positively correlated with Agreeableness and weakly negatively correlated

### Table 1

Bivariate correlations of TOSCA and DES shame and guilt subscales with Goldberg dimensions

<table>
<thead>
<tr>
<th></th>
<th>TOSCA</th>
<th>DES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shame</td>
<td>Guilt</td>
</tr>
<tr>
<td><strong>Self-ratings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-.04</td>
<td>.24**</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.02</td>
<td>.14</td>
</tr>
<tr>
<td>Extraversion</td>
<td>-.28**</td>
<td>-.04</td>
</tr>
<tr>
<td>Openness/Intellect</td>
<td>-.09</td>
<td>.13</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>-.33***</td>
<td>-.18*</td>
</tr>
<tr>
<td><strong>Peer-rating</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>.18*</td>
<td>.08</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.08</td>
<td>.02</td>
</tr>
<tr>
<td>Extraversion</td>
<td>-.16</td>
<td>-.22*</td>
</tr>
<tr>
<td>Openness/Intellect</td>
<td>.07</td>
<td>.00</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>-.05</td>
<td>.00</td>
</tr>
<tr>
<td><strong>Self-criterion residuals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-.08</td>
<td>.23*</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-.01</td>
<td>.14</td>
</tr>
<tr>
<td>Extraversion</td>
<td>-.24*</td>
<td>.03</td>
</tr>
<tr>
<td>Openness/Intellect</td>
<td>-.11</td>
<td>.13</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>-.34***</td>
<td>-.18*</td>
</tr>
</tbody>
</table>

* $p < .05$.
** $p < .01$.
*** $p < .001$. 
with Emotional Stability. The other shame and guilt subscales were robustly negatively correlated with Emotional Stability. In addition, TOSCA shame was negatively correlated with Extraversion and DES shame was negatively correlated with Agreeableness, Conscientiousness, and Extraversion. The patterns of findings for the self-rating data are consistent with the view that shame-proneness may be associated with maladaptive patterns, whereas context-specific guilt as assessed by TOSCA guilt may be associated with adaptive interpersonal patterns.

4.2.2. Relations between shame and guilt with peer-ratings of personality

A different pattern of results emerged, however, from the peer-rating data. In sharp contrast to the self-rating data, TOSCA shame was positively correlated with Agreeableness and TOSCA guilt was negatively correlated with Extraversion. Furthermore, DES shame was negatively correlated only with peer-ratings of Conscientiousness. The pattern of relations between the measures of shame- and guilt-proneness with peer-ratings of personality, therefore, tended to be somewhat weak and inconsistent.

4.2.3. Self-criterion residuals

One of the methods used in this study to assess the discrepancy between self- and peer-ratings of personality dimensions involved computing self-criterion residuals (SCR; Gostling et al., 1998; John & Robins, 1994). The SCR is computed by regressing self-ratings of a construct on peer- or observer-ratings of the same construct and saving the self-report standardized residual as a separate variable. Because all the self-report variance shared with the peer rating has been removed, the residual scores index the degree and direction of the discrepancy. Higher or positive residual scores indicate a tendency to overestimate oneself relative to peer-ratings. Lower or negative scores indicate a tendency to underestimate oneself relative to peer-ratings.

The SCRs revealed that all of the shame and guilt subscales were associated with a tendency to underestimate oneself on Emotional Stability relative to peer-ratings. TOSCA guilt was associated with a tendency to overestimate oneself on Agreeableness relative to peer-ratings, whereas DES shame was associated with a tendency to underestimate oneself on Agreeableness, Conscientiousness, and Extraversion relative to peer-ratings. These findings are consistent with the prediction of this study that TOSCA guilt may be associated with a tendency to overestimate oneself on highly evaluative traits, whereas shame-proneness may be associated with a tendency toward negative self-evaluation.

4.2.4. Hotelling’s t tests

The foregoing analyses suggested some intriguing differences in the correlations obtained between the measures of shame- vs. guilt-proneness with self- and peer-ratings of personality dimensions as well as in the correlations obtained between the measures of shame- and guilt-proneness with the self-criterion residuals. To test whether the strength of the correlations between the shame vs. guilt subscales with self- and peer-ratings of personality as well as the correlations between shame and guilt subscales with self- vs. peer-ratings of personality were indeed significantly
different from one another, a series of Hotelling’s $t$ tests for the significance of difference between two dependent correlations were conducted (Cohen & Cohen, 1983). To minimize the number of $t$ tests performed, the shame and guilt comparisons were conducted only for the subscales that belonged to the same measure (e.g., TOSCA shame vs. TOSCA guilt). The rationale for this is that the TOSCA and DES are based on different formats and conceptualizations and, therefore, comparing the shame and guilt subscales across the measures would constrain the interpretability of the findings.

The first set of Hotelling’s $t$ tests examined whether TOSCA shame vs. TOSCA guilt and DES shame vs. DES guilt were differentially correlated with self-ratings of personality. The $t$ tests revealed that TOSCA shame vs. TOSCA guilt were differentially correlated with self-ratings of Agreeableness ($t = 2.50, p < .01$), Extraversion ($t = 2.14, p < .05$), and Openness/Intellect ($t = 1.92, p < .05$). DES shame vs. DES guilt were also differentially correlated with self-ratings of Agreeableness ($t = 4.01, p < .001$). These findings indicate that chronic shame was correlated negatively with Agreeableness to a significantly greater degree than was chronic guilt. Context-specific guilt, on the other hand, was correlated positively with Agreeableness, and to a significantly greater degree than was context-specific shame. Moreover, context-specific guilt correlated significantly more strongly with Openness/Intellect, yet less strongly with Extraversion than did context-specific shame.

The second set of $t$ tests examined whether TOSCA shame vs. TOSCA guilt and DES shame vs. DES guilt were differentially correlated with peer-ratings of personality. In contrast to the self-rating data, none of the tests yielded significant results, which indicates that peer-ratings of personality dimensions were not differentially correlated with context-specific shame vs. guilt or with chronic shame vs. guilt.

Finally, the last set of Hotelling’s $t$ tests examined whether self vs. peer ratings of personality dimensions were differentially correlated with the measures of shame- and guilt-proneness. Given that the $t$ tests constituted a more stringent test of self–peer discrepancies than the self-criterion residuals, it is not surprising that these tests yielded fewer significant results. The $t$ tests revealed that self- vs. peer-ratings of Agreeableness were differentially correlated with TOSCA shame and DES shame ($ts = 1.72$ and $2.01$, respectively, $ps < .05$). Furthermore, self vs. peer-ratings of Emotional Stability were differentially correlated with TOSCA shame, DES shame, and DES guilt ($ts = 3.51, 2.42$, and $2.43$, $ps < .01$). These findings indicate that whereas both DES guilt and the measures of shame-proneness were associated with a significant discrepancy between self- and peer-ratings of Emotional Stability, only the measures of shame-proneness were associated with a significant discrepancy between self- vs. peer-ratings of Agreeableness.

4.3. Additional analyses

The failure to find robust differences in the relations between the measures of shame- and guilt-proneness with the peer-ratings of personality may have been due (a) to the overlap in the shame and guilt subscales or, (b) to the low validity of the peer-ratings. Additional analyses were conducted to examine these possibilities.
4.3.1. Partial correlations

In the present study, as in previous studies, the shame and guilt subscales showed a substantial positive correlation with one another, presumably reflecting the fact that these two emotions share a number of key features such as dysphoric affect and negative self-evaluation (Tangney, 1995). Thus, following the data analytic strategy used by previous researchers (e.g., Averill et al., 2002; Harder & Greenwald, 1999) partial correlations were computed, to factor TOSCA shame and guilt out from each other and DES shame and guilt out from each other. It should be noted that although this data analytic strategy may reveal the unique contribution of shame and guilt and clarify the relations among variables, it may also remove some of the meaningful variance associated with the emotions and hence should be interpreted with caution (Funder, 1999; Harder, 1995).

Although in general, the partial correlations yielded similar results as the bivariate correlations, several notable changes emerged for the self-ratings of personality and the self-criterion residuals. Factoring out TOSCA guilt, resulted in TOSCA shame becoming associated with a tendency to underestimate one’s Agreeableness relative to peer-ratings ($r = -.18$, $p < .05$). On the other hand, partialing out TOSCA shame, resulted in TOSCA guilt no longer showing relations with self-reports of Emotional Stability ($r = -.06$, n.s.) or with the tendency to underestimate one’s Emotional Stability ($r = .07$, n.s.). Factoring out TOSCA shame also resulted in TOSCA guilt becoming positively correlated with self-reports of Openness/Intellect ($r = .17$, $p < .05$) and with a tendency to overestimate one’s Openness/Intellect ($r = .18$, $p < .05$). With respect to the DES subscales, partialing out DES shame resulted in DES guilt becoming positively correlated with self-reports of Agreeableness ($r = .28$, $p < .01$) and also becoming associated with a tendency to overestimate one’s Agreeableness relative to peer-ratings ($r = .25$, $p < .01$). The partial correlations for the peer-ratings revealed that factoring out the shared variance between the shame and guilt subscales, reduced two of the correlations to marginal significance ($rs = .16$, $ps = .059$), but did not change their patterns of relations with the personality dimensions. These analyses indicate that the failure to find robust differences in the relations between the measures of shame- vs. guilt-proneness with the peer-ratings of personality cannot be attributed to the overlap in the shame and guilt subscales.

4.3.2. Additional DES emotion subscales

Another possible explanation for the weak and inconsistent relations between the shame and guilt subscales with the peer-ratings of personality is that the peer-ratings of personality in this study simply had low validity. To examine this possibility, the correlations between four additional DES emotion subscales and the peer-ratings of personality were also computed. DES contempt was negatively correlated with peer-ratings of Agreeableness, $r = -.18$, $p < .05$. DES shyness was negatively correlated with peer-ratings of Extraversion, $r = -.24$, $p < .01$, and DES sadness and anger were negatively correlated with peer-ratings of Emotional Stability, $rs = -.17$ and $-.18$, respectively, $ps < .05$. Although the magnitude of the correlations between the four DES emotion subscales and peer-ratings of personality were not strong,
they were higher in magnitude than those obtained for the measures of shame- and
guilt-proneness, and the patterns of relations were theoretically coherent. These find-
ings suggest that the weak and inconsistent relations between the measures of shame-
and guilt-proneness with the peer-ratings of personality cannot be attributed to the
low validity of the peer-ratings of personality.

5. Discussion

The findings from this study cast a different light on the extant research on shame-
and guilt-proneness. Consistent with findings from previous studies, the self-rating
data yielded some support for the view that shame-proneness may be associated with
less adaptive patterns than guilt-proneness. However, the peer-rating data, failed to
corroborate these findings. This study, furthermore, revealed that some of the shame
and guilt subscales were associated with a discrepancy between self- vs. peer-ratings
of personality dimensions. Several possible interpretations of these findings as well as
their implications for future research are discussed in light of the extant research on
personality judgment.

5.1. Self-ratings of personality

The patterns of relations between the measures of shame- and guilt-proneness
with self-ratings of personality were similar to those obtained in previous studies
and were also supportive of the view that shame-proneness may be associated with
less adaptive patterns than guilt-proneness. As expected, all of the shame
and guilt subscales were negatively correlated with Emotional Stability, though
in the case of TOSCA guilt, the correlation was reduced to zero, once the contri-
bution with TOSCA shame was factored out. TOSCA guilt was positively corre-
lated with Agreeableness, and once the contribution of TOSCA shame was
factored out, it also became positively correlated with Openness/Intellect. Quite
unexpectedly, DES guilt became positively correlated with Agreeableness, once
the contribution of DES shame was controlled. In contrast to the measures of
guilt-proneness, TOSCA shame was negatively correlated with Extraversion and
DES shame was negatively correlated with Agreeableness, Conscientiousness,
and Extraversion.

Although the self-rating data were generally supportive of the view that shame-
proneness is associated with less adaptive patterns than guilt-proneness, it must also
be acknowledged that the patterns of relations between the shame and guilt subscales
with the self-ratings of personality dimensions showed substantial variation across
the measures. Unlike DES guilt, TOSCA guilt was positively correlated with Agree-
ableness, even before the contribution of the shame subscale was factored out. In ad-
dition, unlike TOSCA guilt, DES guilt remained robustly negatively correlated with
Emotional Stability, even after the contribution of the shame subscale was
controlled. Furthermore, whereas TOSCA shame was negatively correlated with
only Extraversion and Emotional Stability, DES shame was, in addition to these
personality dimensions, also negatively correlated with Agreeableness and Conscientiousness. These patterns of findings are consistent with the view that checklist measures are likely to tap into more chronic levels of shame- and guilt-proneness than the scenario-based measures, and hence are more likely to be associated with maladaptive patterns (e.g., Ferguson & Crowley, 1997; Quiles & Bybee, 1997).

When the strength of the correlations between the self-ratings of personality with TOSCA shame vs. guilt and DES shame vs. guilt were compared, more robust differences between shame- and guilt-proneness were obtained using the TOSCA rather than the DES. DES shame vs. guilt were differentially correlated with only Agreeableness, whereas TOSCA shame vs. guilt were differentially correlated with Agreeableness, Extraversion, and Openness/Intellect. These pattern of findings are congruent with those obtained in previous studies and lend further support to the view that perhaps because TOSCA guilt incorporates reparative behaviors in its conceptualization, it is particularly likely to tap into adaptive reactions to daily transgressions.

5.2. Peer-ratings of personality

The peer-rating data, however, yielded a different pattern of findings from the self-ratings of personality. Consistent with the self-rating data, DES shame was negatively correlated with peer-ratings of Conscientiousness. However, in sharp contrast to the self-rating data, TOSCA shame was positively correlated with peer-ratings of Agreeableness and TOSCA guilt was negatively correlated with peer-ratings of Extraversion. These findings run counter to the view that shame-proneness is associated with destructive modes of conflict resolution, whereas TOSCA guilt is associated with relationship-enhancing functions (e.g., Tangney, 1995). Given that only three significant correlations emerged out of the twenty that were computed between the measures of shame- and guilt-proneness with peer-ratings of personality, one cannot rule out the possibility of chance findings. Nonetheless, it remains problematic that, in contrast to the self-rating data, the tests of significance of difference between the correlations revealed that neither TOSCA shame vs. guilt nor DES shame vs. guilt were differentially correlated with peer-ratings of any of the five personality dimensions. Previous research reveals that peer-ratings represent a valid method of assessing overt personality characteristics provided that the raters have had an adequate opportunity to observe the targets in everyday life (e.g., Funder, 1999; John & Robins, 1993; Kolar, Funder, & Colvin, 1996). In the present study, the mean duration of time the peers knew the participants was 8 years and the median duration was 5 years; hence most of the raters presumably had acquired sufficient information to provide valid ratings of the targets. This study, furthermore, revealed that compared to the measures of shame and guilt, the DES emotion subscales for contempt, shyness, anger, and sadness showed more theoretically coherent patterns of relations with the peer-ratings of personality dimensions. These findings suggest that the weak and inconsistent relations between the measures of shame- and guilt-proneness cannot be simply attributed to the low validity of the peer-ratings.
5.3. Self–peer discrepancies

The findings from this study also revealed that the shame and guilt subscales were associated with distinct patterns of discrepancies in self- vs. peer-ratings of personality. The self-criterion residuals revealed that all of the shame and guilt subscales were associated with underestimating one's Emotional Stability relative to peer-ratings, though in the case of TOSCA guilt, once the shared variance with TOSCA shame was controlled, the correlation reduced to zero. TOSCA guilt was associated with overestimating one's Agreeableness relative to peer-ratings and also became associated with a tendency to overestimating one's Openness/Intellect, once the contribution of TOSCA shame was removed. Interestingly, DES guilt became associated with a tendency to overestimate one's Agreeableness, once the contribution of DES shame was removed. In contrast to the measures of guilt-proneness, the measures of shame-proneness were associated with a tendency to underestimate oneself in several domains relative to peer-ratings. TOSCA shame was associated with a tendency to underestimate one's Extraversion and also became associated with a tendency to underestimate one's Agreeableness, once TOSCA guilt was controlled. Furthermore, DES shame was associated with underestimating one's Agreeableness, Conscientiousness, and Extraversion relative to peer-ratings. Taken together, these findings are consistent with the predictions of this study that TOSCA guilt may be associated with a tendency to overestimate oneself on highly evaluative traits, whereas shame-proneness may be associated with a pervasive tendency toward negative self-evaluation.

When more rigorous tests were used to examine the relations between the shame and guilt subscales with self–peer discrepancies in personality ratings, however, only few of the foregoing findings held up. The tests revealed that TOSCA shame, DES shame, and DES guilt were associated with a significant self–peer discrepancy with respect to Emotional Stability. In addition, both measures of shame-proneness were also associated with a significant self–peer discrepancy with respect to Agreeableness. The latter set of findings are broadly consistent with recent reports that self-ratings of emotion expressions of shame are positively correlated with peer-ratings, but not with self-ratings of Agreeableness (Trieweiler et al., 2002).

The question that arises is when there is a significant discrepancy between self- vs. peer-ratings of personality, whether more credence or weight should be given to the self- or to the peer-ratings. The extant research on personality judgment suggests that this may depend on some extent on the personality dimension under consideration. A growing body of research has found that the self has unique advantages and disadvantages as a judge of personality (Funder, 1999; John & Robins, 1993, 1994). The self has access to internal thoughts and feelings and other ‘privileged information,’ which is not available to the external observer. Thus, it is not surprising that self-ratings of Neuroticism have been found to be more accurate than informant ratings in predicting daily reports of emotional experiences (Spain, Eaton, & Funder, 2000). On the other hand, the self is also more ego-involved than others in its assessment of itself and has greater difficulty viewing itself objectively. Previous research has found that self-reports of Agreeableness are highly susceptible to social
desirability bias (Funder, 1999; John & Robins, 1993), and that of the five personality dimensions, it consistently exhibits the lowest level of self–peer agreement (Jensen-Campbell & Graziano, 2000; John & Robins, 1993; Watson et al., 2000). Studies that have examined the comparative validity of self and other reports of personality have, furthermore, found that self- and peer- ratings of Agreeableness show a low level of convergence and that peer–peer and observer–observer reports generally show a higher level of convergence than self–peer or self-observer reports (John & Robins, 1993; Kolar et al., 1996; Gostling et al., 1998).

The extant research on personality judgment, therefore, indicates that in the case of Emotional Stability, more weight or credence should probably be given to self- than to peer-ratings. Emotional Stability assesses characteristics that are largely internal and private, which may be difficult for observers to rate accurately. Furthermore, in the present study, the discrepancy between self- vs. peer-ratings of Emotional Stability was due to the fact that TOSCA shame, DES shame, and DES guilt were robustly negatively correlated with self-ratings, but not with peer-ratings of Emotional Stability. Similar results were obtained in a recent study, which found that self-ratings of a range of negative emotions (e.g., sad and fear) were robustly positively correlated with self-ratings of Neuroticism, but not with peer-ratings of Neuroticism (Spain et al., 2000). These findings suggest that self–peer discrepancies in ratings of Emotional Stability are not unique to shame- and guilt-proneness, but are rather characteristic of negative emotions in general. Perhaps the most plausible explanation for the discrepancy between self- vs. peer-ratings of Emotional Stability is that with socialization most individuals learn to inhibit the expression of negative emotions in accordance with social ‘display rules’ (Ekman, 1972).

With respect to the personality dimension of Agreeableness, however, the extant research on personality judgment suggests that more credence or weight should probably be given to peer- rather than to self-ratings. In contrast to Emotional Stability, Agreeableness assesses overt social behaviors such as “helpful,” “considerate,” and “kind.” The research on personality judgment indicates that these are characteristics, which may be more objectively assessed by a friend than a target. Furthermore, in contrast to Emotional Stability, the source of the self–peer discrepancy for Agreeableness were in opposite directions for the TOSCA and the DES shame subscales. The discrepancy for TOSCA shame, was due to the fact that TOSCA shame was positively correlated with peer-ratings, but not with self-ratings of Agreeableness. In the case of DES shame, the discrepancy was due to the fact that DES shame was negatively correlated with self-ratings, but not with peer-ratings of Agreeableness. These seemingly inconsistent pattern of findings can be reconciled, if one takes into account that checklist measures assess more chronic levels of shame-proneness than scenario-based measures, but that both types of measures of shame-proneness are associated with a tendency toward self-devaluation (e.g., Harder, 1995; Tangney et al., 1995). Thus, perhaps the most plausible explanation for the finding that shame-proneness is associated with a discrepancy between self- vs. peer-ratings of Agreeableness is that shame-prone individuals do, in fact, underestimate their interpersonal qualities.
However, an alternative explanation for why shame-proneness is associated with a discrepancy between self- vs. peer-ratings of Agreeableness may also be offered. In a recent study, Gross and John (1998) identified five domains related to emotion expressivity. One of the domains, referred to as Masking, involves concealing the expression of one’s feelings for self-presentational purposes. Masking was found to be highly correlated with self- and peer-ratings of the self-conscious emotions. According to Colvin’s (1993) research on judgability, individuals characterized by a high discrepancy in private inner-self and public outer-self tend to have lower self-other agreement on personality traits. Thus, shame-prone individuals may, in fact, experience “seething, resentful, retaliative anger” towards others, but manage to conceal overt manifestations of these feelings from their peers.

Further research needs to be done to clarify the relative merits of the two foregoing explanations because they lead to different implications for clinical practice. If shame-prone individuals have a tendency to underestimate their interpersonal qualities, then an important therapeutic task would be to challenge their irrational beliefs and to bring their perceptions more in line with reality which, in turn, may serve to modify their negative self-image. However, if the discrepancy between self- vs. peer-ratings of Agreeableness is due to the fact that shame-prone individuals engage in high levels of ‘masking’ then an important therapeutic task would be to help them to manage their feelings more effectively. According to Gross and John (1998) individuals who engage in high levels of masking “have the worst of possible worlds (p. 186).” The strategies used by these individuals to regulate their moods are not only counterproductive, but may also come with high physiological and cognitive costs.

The view that shame-proneness is more maladaptive than guilt-proneness has rested in part on studies which have demonstrated that shame-proneness as assessed by a variety of measures is associated with aggression and hostility as well as destructive modes of conflict resolution, whereas guilt-proneness as assessed by scenario-based measures and narrative analysis is associated with empathy and constructive modes of conflict resolution (e.g., Leith & Baumesiter, 1998; Tangney, 1995). A growing body of research has revealed that the core features associated with Agreeableness are the ability to regulate anger, the use of constructive modes of conflict resolution, and the motivation to maintain positive relations with others (e.g., Graziano, Jensen-Campbell, & Hair, 1996; Jensen-Campbell & Graziano, 2001; Tobin, Graziano, Vanman, & Tassinary, 2000). Research that further clarifies the relations between shame- and guilt-proneness with Agreeableness is, therefore, central to understanding the implications that proneness to experiencing these emotions have for psychological functioning.

5.4. Implications for future research

This study has several implications for future research on the correlates of shame- and guilt-proneness. First, the findings from this study underscore the importance of supplementing self-report measures with additional types of outcome measures. The discrepancy in the patterns of findings between the self- and peer-ratings of personality raise the possibility that some of the results from the extant research on shame
and guilt may have been in part an artifact of self-report bias. One method of extending this study would be to obtain peer- or observer-ratings of personality from multiple raters and aggregating them across the different raters. Some researchers have argued that in the absence of a single objective criterion for assessing complex social behaviors, social consensus or the aggregated judgment of others may serve as an acceptable criterion against which to measure the accuracy of self-perception (e.g., Gostling et al., 1998; John & Robins, 1994; Kolar et al., 1996).

Second, the findings from this study also underscore the importance of using multiple measures of shame- and guilt-proneness. Consistent with expectations, the patterns of relations between the shame and guilt subscales with self-reports of personality dimensions showed substantial variation across the TOSCA and DES. In general, the patterns of relations were consistent with the view that checklist measures of shame and guilt may assess more chronic levels of these emotions than scenario-based measures (e.g., Ferguson & Crowley, 1997; Quiles & Bybee, 1997). There is considerable debate in the field as to whether chronic guilt as assessed by checklist measures is distinguishable from shame-proneness (Harder, 1995; Tangney, 1995). In the present study, as in previous studies, the checklist measure of guilt was more robustly correlated with the measures of shame-proneness than with the scenario-based measure of guilt. Furthermore, like the measures of shame-proneness, DES guilt was robustly negatively correlated with self-reports of Emotional Stability. However, unlike the measures of shame-proneness, DES guilt was not associated with a significant self–peer discrepancy with respect to Agreeableness. If these findings are replicated in future studies, they may indicate that even though chronic guilt and shame-proneness are both associated with high levels of dysphoric affect, the tendency to negatively evaluate one's interpersonal qualities is uniquely associated with shame-proneness.

References


