A Confirmatory Factor Analytic Study of Anxiety, Fear, and Behavioral Inhibition System Measures

Andrew J. Cooper, Adam M. Perkins, and Philip J. Corr

Department of Psychology, University of Wales, Swansea, UK

Abstract. Recent revisions to the reinforcement sensitivity theory (RST) of personality have highlighted the distinction between the emotions of fear and anxiety. These revisions have substantial implications for self-report measurement; in particular, they raise the question of whether separate traits of fear and anxiety exist and, if so, their interrelationship. To address this question, the current study used confirmatory factor analytic procedures to examine the convergent and discriminant validity of measures of trait anxiety, fear, and the behavioral inhibition system (BIS). We also examined measurement and structural invariance across gender in 167 males and 173 females who completed the Spielberger State-Trait Anxiety Inventory (STAI), the Carver and White BIS Scale, and the Fear Survey Schedule (FSS). The findings suggested that trait anxiety and the BIS scale are relatively distinct from Tissue Damage Fear (FSS). Further, the final model showed measurement and structural invariance across gender. The implications of the results for future self-report assessment in RST research are discussed.

Keywords: reinforcement sensitivity theory, behavioral inhibition system, fear, trait anxiety, confirmatory factor analysis, measurement invariance

Empirical evidence concerning the structure of negative emotion in human personality is mixed: Anxiety often appears to be distinct from fear in clinical populations (e.g., Geer, 1965; Wolpe & Lang, 1977), whereas studies in the wider population (especially those that entail factor analyses of responses to general personality questionnaires; Caseras, Ávila, & Torrubia, 2003) usually indicate that negative emotions form a single dimension of personality (often labeled neuroticism; Eysenck, 1967; or negative affectivity; Watson, Wiese, Vaidya, & Tellegen, 1999). Gray and McNaughton (2000) bridged the clinical-general divide on this topic by proposing that anxiety (manifested as risk assessment and behavioral inhibition) is elicited by threats that need to be approached, whereas fear (manifested as flight if an escape route is available) is elicited by threats that need not be approached. Anxiety and fear, thus, represent distinct emotional and motivational states in terms of how they relate to defensive behavior. The implications of these revisions to the broader model of personality, now known as reinforcement sensitivity theory (RST; Corr, 2004; McNaughton & Corr, 2004), are yet to be fully explored; how the anxiety and fear states in the revised RST model relate to personality traits is an open question (Corr & McNaughton, 2007). The aim of the current study was to use confirmatory factor analysis (CFA) to examine the convergent and discriminant validity of measures of anxiety, fear, and the behavioral inhibition system (BIS) used in RST research, and to use multiple-group CFA to examine measurement and structural invariance across gender.

Measuring RST Constructs

RST represents the outcome of more than 30 years of research by Jeffrey Gray and colleagues on the causal basis of personality (Corr, 2004; Gray, 1970; Pickering et al., 1997). One key aspect of this earlier work was the elucidation of a personality construct known as punishment sensitivity, based on a BIS. The BIS was originally theorized to mediate sensitivity to conditioned signals of punishment and frustrative nonreward. On the basis that Gray (1970) initially aligned punishment sensitivity with trait anxiety, many early RST studies simply used existing self-report measures of trait anxiety as proxy measures of BIS sensitivity. An example of one such measure would be the Spielberger STAI (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). More recently, however, purpose-built RST measures have been designed and utilized. The BIS scale from the Carver and White (1994) BIS/BAS Scales and the sensitivity to punishment scale from the sensitivity to punishment and sensitivity to reward questionnaire (Torrubia, Ávila, Molto, & Caseras, 2001) are two such measures now widely used. These measures, however, are only fully applicable in the context of the original RST. In the revised RST, what has been termed the flight-fight-freeze system (FFFS) is now presumed to mediate reactions to all aversive stimuli, while the BIS is now presumed to mediate goal conflict and risk assessment (McNaughton & Corr, 2004). Existing RST measures may, thus, conflate systems and

DOI 10.1027/1614-0001.28.4.179

processes that are now presumed to have different causal bases. Very few studies have considered the revised RST, let alone the implications the revisions may have for psychometric assessment.

Exploratory factor analyses of RST-related measures show that BIS-related measures tend to load on one factor (Caseras et al., 2003; Zelenski & Larsen, 1999). This is in contrast to behavioral approach system (BAS) measures, which tend to load across two or three separate factors. These findings would suggest that BIS-related measures tend to share substantial variance. Indeed, previous studies have shown that BIS measures do intercorrelate reasonably highly (Carver & White, 1994; Caseras et al., 2003; Gomez & Gomez, 2005). It should be noted, however, that these previous studies have not included measures of fear. Although a number of well-established questionnaire measures of fear exist, they tend to have been developed for the specific purpose of measuring change in fear responses to phobic stimuli during treatment rather than measuring personality per se. These questionnaires, therefore, typically consist of a list of potential fear-inducing stimuli (e.g., spiders) that are rated according to how much fear they elicit from the respondent. One example of this type of fear questionnaire is the Fear Survey Schedule (FSS; Wolpe & Lang, 1964). A number of versions of the FSS have been developed and there have been several factor analytic studies of this measure (Arrindell et al., 2003; Arrindell et al., 1987). The two most robust of the factors extracted across a range of studies have been labeled as bodily injury/tissue damage fear and social/interpersonal fear (Arrindell, 1980). The content of the FSS is sufficiently different from that of most well-known personality questionnaires (which typically ask general questions such as "are you a worrier?") that it could be argued that the FSS and other fear questionnaires might not represent Gray and McNaughton's (2000) view of fear as a general personality construct caused by threat sensitivity.

This particular issue has been addressed in two recent studies (Perkins & Corr, 2006; Perkins, Kemp, & Corr, 2007). Perkins et al. (2007) examined whether fear questionnaire scores relate meaningfully to performance in a military examination of tactical judgment in combat scenarios. In a sample of 101 officer cadets, they found that total scores on the FSS were significantly and negatively associated with tactical judgment performance and captured variance in performance that was not shared with anxiety (and nor with any of the other major personality constructs measured, including neuroticism). When examining subscales of the FSS, it was found that scores on tissue damage fear related negatively to performance in combat scenarios, while scores on other fear subscales (e.g., animal fear and social fear) turned out to show no significant relationship to performance. Perkins and Corr (2006) examined how a number of RST-related personality measures related to measures of *defensive direction* (e.g., approach vs, avoidance of threat) and *defensive intensity* (e.g., strength of defensive reaction) derived from a set of human defensive scenarios that were modeled on typical rodent defensive reactions. The results showed that the total fear score from the FSS positively correlated with defensive direction away from a source of threat (e.g., avoidance); the BIS scale also positively correlated with defensive direction away from a source of threat. The bivariate relationship between the STAI and defensive direction was not significant, but the STAI did significantly predict approach to threat in a multiple regression model. Overall, the findings suggest that FSS scores are tapping a personality construct of fearfulness in a face-valid manner, and that this construct approximates fairly well to Gray and McNaughton's (2000) conceptualization of fear as sensitivity to threat.

The Current Study

It is important that the implications of the revisions of the RST for self-report assessment are considered. Accordingly, our first aim was to examine the convergent and discriminant validity of the FSS tissue damage fear and social fear scales, the STAI, and the BIS scale. A series of nested CFA models were tested to examine these relationships. Only one previous study (Perkins & Corr, 2006) has considered relationships between these measures, and these data may be limited as only observed scores were considered. As such, the relationships may have been confounded by measurement error. A clearer test of convergent and discriminant validity between these variables may be obtained by examining relationships between latent variables. Further, Perkins and Corr (2006) only considered the total FSS score. As Perkins et al. (2007) demonstrated, in the context of the RST there may be crucial differences between tissue damage fear and social fear. Tissue damage fear may relate more closely to fear as conceptualized by Gray and McNaughton (2000); social fear may relate more closely to anxiety, as the social fear items tend to describe BIS-related conflict situations (e.g., public speaking) that entail approach to threat or conflict.

Having established a suitable CFA model, the second aim of this study was to examine measurement and structural invariance across gender. There have been relatively few gender invariance studies in the RST literature; it is important to establish that measurement models hold across important demographic variables (Brown, 2006; Leone, Perugini, Bagozzi, Pierro, & Mannetti, 2001). Further, previous studies have generally shown that females tend to score more highly than males on measures of negative emotionality, particularly measures of fear (Arrindell et al., 2003; Arrindell et al., 1987). As these differences have been established on observed scores, it is important to control for measurement error by examining differences in latent scores across gender.

Method

Participants

A total of 340 people (167 males and 173 females), aged between 18 and 77 (*M* = 26.06; *SD* = 8.03), were recruited through advertisements at a regional university campus in the United Kingdom and among employees at a local supermarket in the university area. Approximately 15% of the participants received course credit for introductory psychology for participating in the study.

Measures

Trait Anxiety

The STAI Form Y2 (Spielberger et al., 1983) is a 40-item self-report measure of trait and state anxiety. Only trait anxiety was measured in the current study. Items were rated on a 4-point Likert-type response format, with a response of 1 indicating *almost never* and 4 indicating *almost always*. Items are summed to form a total score for trait anxiety. This scale has been used widely as a proxy measure of the BIS in RST research (Torrubia et al., 2001). It has very good reported reliability and validity (Spielberger et al., 1983). The Cronbach's α value in the current study was 0.91.

BIS

The Carver and White (1994) BIS/BAS scales are a measure comprising a BIS scale (seven items) and three BAS scales: reward responsiveness (five items), drive (four items), and fun-seeking (four items). Each item is answered using a 4-point Likert scale, ranging from 1 (*very false for me*) to 4 (*very true for me*). In the current study only the BIS scale was utilized. Previous research has shown the scales have satisfactory internal reliability and construct validity (Carver & White, 1994; Gomez, Cooper, & Gomez, 2005; Heubeck, Wilkinson, & Cologon, 1998). In the current study the BIS scale had a Cronbach's α value of 0.83.

Fear

Fear was assessed using the FSS (Wolpe & Lange, 1964). The FSS has been used in numerous studies over the last 3 decades and is arguably the most reliable and valid measure of fear available (Oei, Cavallo, & Evans, 1987). Different versions of the FSS, ranging in length from 8 items to 108 items, have been developed (Arrindell, 1980). The FSS used here was the standard length version, consisting of a list of 73 items representing specific aversive stimuli such as "worms" or "angry people." Respondents indicated, using a scale of 0 (*no fear*) to 4 (*very much fear*), how much they would be disturbed by each item. Arrindell (1980) derived a five-factor solution from the FSS based on a subset of 52 FSS items. Two of the derived factors were Tissue Damage Fear (composed of 12 items) and Social Fear (composed of 13 items). In the current study, we only consider the Tissue Damage Fear and Social Fear factors derived by Arrindell (1980). In the current study, the Social Fear factor had a Cronbach's α value of 0.89, while the Tissue Damage Fear factor had a Cronbach's α value of 0.88.

Procedure

Following informed consent, all participants completed the BIS, STAI, and FSS measures as part of a battery of personality questionnaires. All participants completed the questionnaires individually. They were informed that the study was aiming to examine relationships between personality traits. Following completion of the questionnaire package, participants were thanked and debriefed.

Data Analysis

All CFA analyses were conducted with the MPlus 4.1 software program (Muthen & Muthen, 2006), using maximum likelihood estimation of the sample covariance matrix. In this study, parcels of summed items were created as observed indicators for each latent factor[1]. Item parceling in structural equation modeling is appropriate in the current study, as CFA models using individual items as observed indicators would have created complex models (e.g., 52 observed indicators and four latent factors) that would arguably have needed a far larger sample size than that available in the current study (Bandalos & Finney, 2001; Little, Cunningham, Shahar, & Widaman, 2002). Further, item parcels are less affected by measurement error and nonnormal distributions, and, thus, reduce computational problems in the testing of structural equation models. Item parceling has been used successfully in previous RST CFA studies (Gomez & Gomez, 2005; Leone et al., 2001).

Model fit was ascertained using the minimum fit function χ^2. As χ^2 values are potentially inflated by large sample sizes, fit was also examined using three practical fit indices. They were the root mean square error of approximation (RMSEA; Steiger, 1990), the comparative fit index (CFI; Bentler, 1990), and the standardized root mean square residual (SRMR). The RMSEA provides a measure of model fit relative to the population covariance matrix when the

[1] The first parcel comprised the odd-numbered items from the scale, while the second parcel comprised the even-numbered items from each scale. For scales with an odd number of items, the first parcel had one more item than the second parcel. A second parceling method was also tested. This involved the first and second parcels comprising the first and second half of the items from each scale, respectively. The results were very similar across both parceling methods, so only the results using the first parceling method are reported in the text.

Table 1. Descriptive information for the parcels and totals of the STAI, BIS, and FFS measures

	Mean		SD		Skewness		Kurtosis	
	Male	Female	Male	Female	Male	Female	Male	Female
STAI Parcel 1	19.20	20.42	5.08	4.55	0.67	0.57	0.63	0.02
STAI Parcel 2	19.93	20.78	5.04	5.26	0.90	0.71	0.75	0.37
STAI Total	39.13	41.20	9.79	10.40	0.78	0.69	0.68	0.30
BIS Parcel 1	11.30	12.45	2.57	2.41	–0.47	–0.42	0.17	–0.17
BIS Parcel 2	7.58	8.67	2.03	1.86	–0.19	–0.35	–0.47	0.13
BIS Total	18.89	21.13	4.29	3.92	–0.32	–0.35	–0.17	–0.02
TD Parcel 1	5.34	6.87	4.50	5.54	1.06	0.76	1.31	–0.20
TD Parcel 2	5.41	6.98	4.25	4.98	0.83	0.76	0.29	–0.04
TD Total	10.75	13.86	8.31	10.13	1.01	0.71	1.11	–0.34
SF Parcel 1	8.50	11.22	4.50	5.36	0.19	0.02	–0.71	–0.86
SF Parcel 2	6.73	8.73	4.57	5.33	0.71	0.37	0.45	–0.62
SF Total	15.23	19.95	8.72	10.30	0.44	0.19	–0.26	–0.79

Note. STAI = Spielberger Trait Anxiety Inventory, BIS = Behavioral Inhibition System scale, TD = Fear Survey Schedule Tissue Damage scale, SF = Fear Survey Schedule Social Fear scale.

complexity of the model is also taken into account. RMSEA values of < .05 are suggestive of good fit and .05 to .08 as moderate fit. The CFI provides a measure of the fit of the hypothesized model relative to the baseline or independent model, with values usually ranging from 0.00 to 1.00. For the CFI, values above .95 are suggestive of good model fit. SRMR is the standardized difference between the observed covariance and predicted covariance. A value of zero indicates perfect fit, and values less than .08 suggest good fit (Hu & Bentler, 1999).

To determine differences between models, the difference in their χ^2 values was used. This difference is itself distributed as χ^2. Given that χ^2 values are sensitive to sample size and model fit for large numbers of constraints (Cheung & Rensvold, 2002; Little, 1997; Marsh, Balla, & McDonald, 1988), the critical probability value was set at the $p = .001$ level, rather than the $p = .05$ level. Thus, the differences in the χ^2 values of the models compared had to be less than the critical values of χ^2 at the $p = .001$ level. Cheung and Rensvold (2002) examined the properties of the $\Delta\chi^2$ and various approximate fit indices in relation to CFA invariance testing. For practical differences between models, they suggested the difference for CFI values had to be .01 or more to infer differences between the models compared.

Having established a suitable CFA model for all participants, gender invariance involving the measurement and structural invariance of the fear, anxiety, and BIS measures was tested. The sequence of CFA models used to test measurement and structural invariance largely followed Brown (2006). Initially, the factor model was tested separately for males and females. This ascertained that the proposed model structure was appropriate for both groups. At the measurement level, invariance was tested for configural invariance (M1), metric invariance (M2), and scalar invariance (M3). The final measurement invariance model can then be used to test for structural invariance. Structural invariance in this study involved invariance in the latent factor variances (M4), the latent factor covariances (M5), and equality of the latent means (M6). This involved comparing M4, M5, and M6 (separately) with the final measurement invariance model (the full or partial M3 invariance model). The differences in latent means across groups can be examined by constraining the latent mean of one group (the reference group) to 0 and freely estimating the means of the other group. As the latent mean metric is arbitrary in this situation, only the relative difference in latent means can be examined (Brown, 2006; Byrne, 1998).

Results

Descriptive Statistics

Table 1 shows the means, standard deviations, skewness, and kurtosis values for each of the parcels and for the total scores for each measure, separately for males and females. According to Curran, West, and Finch (1996), for univariate normality, skewness and kurtosis values of 0 to 2, and 0 to 7, respectively, can be taken as demonstrating sufficient normality. On the basis of the values shown in Table 1, the data for both males and females appear to show sufficient normality.

Structural Organization of the Fear, Anxiety, and BIS Measures

Initially, a one-factor CFA model was tested. In this model, all parcels loaded on a single latent factor. All of the error terms were uncorrelated. As expected, this model showed poor global fit, χ^2 (20, N = 340) = 960.11, $p < .000$; RMSEA = 0.372 (90% confidence interval = .352–.392), CFI = 0.52; SRMR = 0.181. Subsequently, a three-factor CFA model

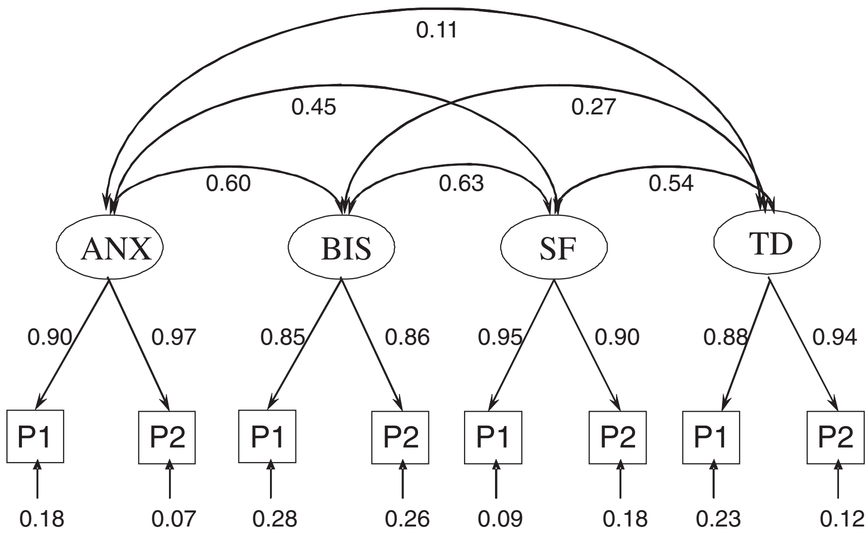

Figure 1. Parameter estimates for the four-factor model of the STAI, BIS, and fear measures. *Note*. ANX = Spielberger Trait Anxiety Inventory, BIS = Behavioral Inhibition System scale, SF = Fear Survey Schedule Social Fear scale, TD = Fear Survey Schedule Tissue Damage scale, P1 = Parcel 1, P2 = Parcel 2. Parcel 1 for each measure was fixed to unity for the purpose of identification.

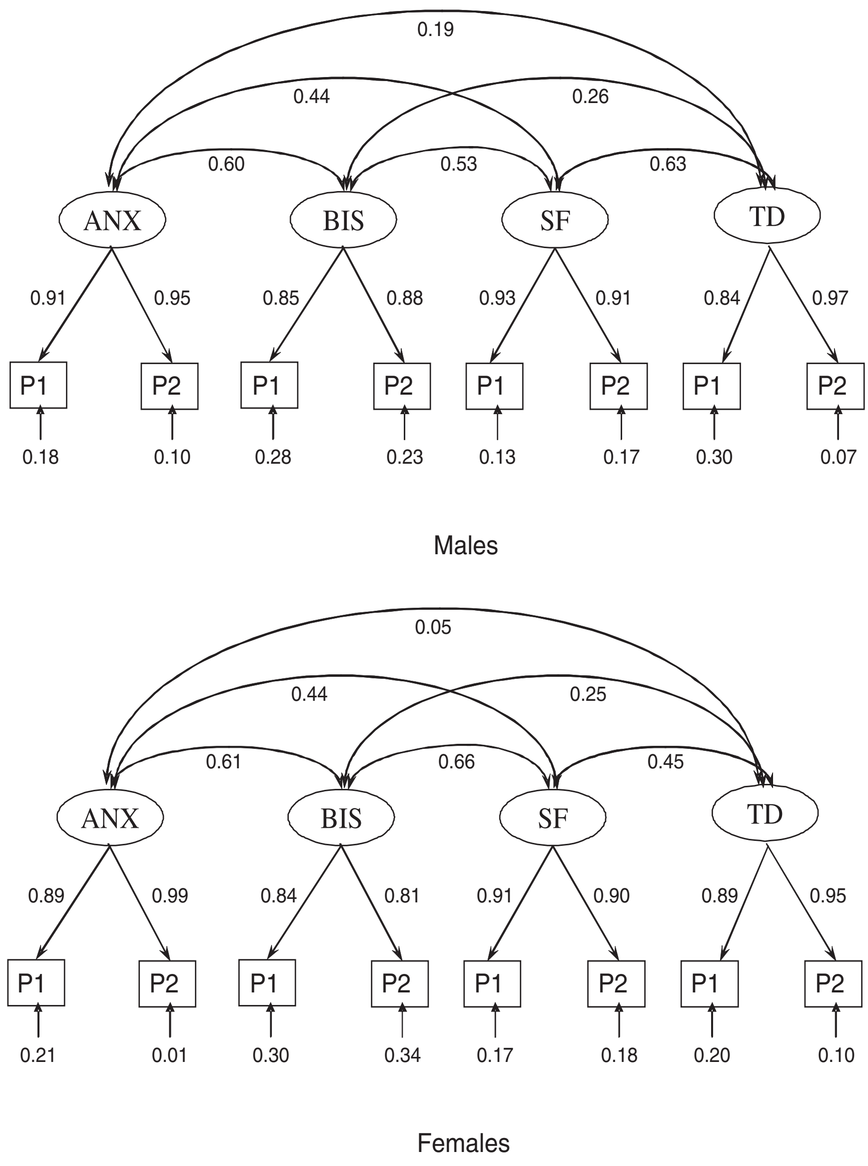

Figure 2. Parameter estimates for the four-factor model of the STAI, BIS, and fear measures for males and females. *Note*. ANX = Spielberger Trait Anxiety Inventory, BIS = Behavioral Inhibition System scale, SF = Fear Survey Schedule Social Fear scale, TD = Fear Survey Schedule Tissue Damage scale, P1 = Parcel 1, P2 = Parcel 2. Parcel 1 for each measure was fixed to unity for the purpose of identification.

Table 2. Results of tests for measurement and structural invariance for the BIS, anxiety, and fear CFA model

Model (M)	*df*	χ²	RMSEA 90% CI	CFI	SRMR	Δ Model	Δ *df*	Δχ²	δCF1
Configural invariance	28	30.80	.024 .000–.066	.99	.019	–	–	–	
M2. Full metric invariance (M1 with all factor loadings constrained equal)	32	32.32	.008 .000–.057	1.00	.023	M2–M1	4	1.52	.01
M3: Scalar invariance: (M2 with indicator intercepts constrained equal)	36	38.30	.019 .000–.059	1.00	.023	M3–M2	4	5.98	.00
M4: Factor variances invariance (M3 with all factor variances constrained equal)	40	49.98	.038 .000–.069	1.00	.084	M4–M3	4	11.68	.00
M5: Factor covariances invariance (M3 with all factor covariances constrained equal)	42	43.30	.013 .000–.054	1.00	.038	M5–M3	6	5.00	.00
M6: Test for equality of latent means (M3 with latent factor means constrained equal)	40	73.29*	.070 .044–.095	.98	.097	M6–M3	4	34.99*	–.02

Note. * $p < .001$, χ² = minimum fit function χ², RMSEA = root mean square error of approximation, CFI = comparative fit index, SMSR = standardized root mean square residual, Δ = difference. The CFI and ΔCFI values are rounded to two decimal places.

was tested. In this model, the two STAI parcels loaded on a latent Anxiety factor, the two BIS parcels loaded on a latent BIS factor, and the four tissue damage and social fear parcels loaded on a single latent Fear factor. The latent factors were allowed to covary and the error terms were uncorrelated. Although this model showed better fit than the one-factor model, $\Delta\chi^2$ ($df = 3$) = 631.72, $p < .0001$, overall it showed poor global model fit, χ^2 (17, $N = 340$) = 328.39, $p < .000$; RMSEA = 0.232 (90% confidence interval = .211–.254), CFI = 0.84; SRMR = 0.104.

Finally, a four-factor CFA model was tested. In this model, the STAI parcels loaded on a latent Anxiety factor, the BIS parcels loaded on a latent BIS factor, the tissue damage parcels loaded on a latent Tissue Damage factor, and the social fear parcels loaded on a latent Social Fear factor. The factors were allowed to covary and the error terms were uncorrelated. This model is shown in Figure 1. This model showed significantly better fit than the three-factor model, $\Delta\chi^2$ ($df = 3$) = 313.20, $p < .0001$, and showed excellent global model fit, χ^2 (14, $N = 340$) = 15.19, $p > .05$; RMSEA = 0.016 (90% confidence interval = .000–.056), CFI = 0.99; SRMR = 0.012. Figure 1 provides the standardized parameters for this model for all participants. As can be seen, all parcels had relatively high loadings on their respective latent factor. In terms of relationships between the latent factors, the STAI, BIS, and Social Fear latent factors had relatively high intercorrelations, whereas Tissue Damage Fear was not significantly related to STAI ($p > .05$). All other factor loadings and latent variable covariances were significant ($p < .01$).

Testing for Measurement Invariance Across Gender

The four-factor model was initially tested separately for males and females. The results of these analyses showed good model fit for the males, χ^2 (14, $N = 167$) = 23.81, $p < .05$; RMSEA = 0.065 (90% confidence interval = .006–.108), CFI = 0.99; SRMR = 0.024, and excellent model fit for the females, χ^2 (14, $N = 173$) = 13.90, $p > .05$; RMSEA = 0.000 (90% confidence interval = .000–.073), CFI = 1.00; SRMR = 0.013. Table 2 shows the results for the tests of measurement and structural invariance across gender. The first test for measurement invariance involved testing for configural invariance (M1). As shown in Table 2, this model showed excellent fit. Figure 2 shows the standardized parameter estimates for both males and females. All factor loadings and latent variable covariances were significant ($p < .01$), except for the covariance of Tissue Damage Fear and STAI for females ($p > .05$).

Having established configural invariance, M2 tested for full metric invariance by constraining all factor loadings to equality across the two groups. Table 2 shows that there was no statistical or practical difference in fit indices between M1 and M2, thereby indicating full metric invariance for this model across gender. M3 tested for scalar invariance by constraining indicator intercepts to equality across groups. Table 2 shows that there was no statistical or practical difference in fit indices between M2 and M3, thereby indicating full scalar invariance across gender.

Testing for Structural Invariance Across Gender

Having established full measurement invariance for the four-factor model, structural invariance was tested across gender. M4 tested for latent factor variance invariance. As shown in Table 2, M4 did not differ significantly in model fit from M3 (the test for scalar invariance), thereby indicating factor variance invariance across gender. M5 tested for latent factor covariances invariance. As shown in Table 2, M5 did not differ significantly in model fit from M3, thereby indicating factor covariances invariance across gender. Finally, M6 was compared to M3 to test the equality of latent factor means across gender. As shown in Table 2, M6 showed significantly poorer fit than M3, based on both the $\Delta\chi^2$ and ΔCFI. This result suggests there is not equality of latent means across gender. To examine differences in each latent mean across gender, the means of each latent factor were fixed to zero for the males (hence, they were the reference group). On this basis, females had significantly higher latent means for BIS ($p < .001$), Social Fear ($p < .001$), and Tissue Damage Fear ($p < .001$), but not STAI ($p > .05$).

Discussion

To date, very little research has considered the implications of Gray and McNaughton's (2000) revised model of RST for psychometric self-report measurement. Based on the distinction between fear and anxiety in the revised RST, it is clearly important that research addresses how self-report measures of these constructs relate to each other. Examining relationships between existing measures of negative emotionality currently used in RST research may help facilitate the development of new measures suitable for the revised RST.

The current study aimed to examine the convergent and discriminant validity of self-report measures of the BIS, Trait Anxiety, Social Fear, and Tissue Damage Fear using CFA, and to examine measurement and structural invariance across gender. As expected, a four-factor CFA model of the STAI, BIS, Social Fear, and Tissue Damage Fear showed the best model fit. The superior fit of the four-factor model when compared with nested one- and three-factor models, supports the notion that these scales are measuring distinct constructs.

In terms of relationships between the latent factors in the four-factor model, the STAI, BIS, and Social Fear factors tended to intercorrelate relatively highly, suggesting they are strongly related constructs. Social Fear also had a strong positive correlation with Tissue Damage Fear. BIS, however, only had a moderate positive correlation with Tissue Damage Fear, while STAI was not significantly related to Tissue Damage Fear. This suggests the BIS and Anxiety constructs are not strongly related to Tissue Damage Fear.

The second aim of the study was to examine the measurement and structural invariance of the four-factor model across gender using a sequence of nested CFA models. Initially, measurement invariance was tested. The results supported configural invariance, full metric invariance, and scalar invariance across gender. Having established measurement invariance, of interest for the current study was the examination of structural invariance across gender. Factor variance invariance and factor covariance invariance across gender was supported. The latter finding suggests that the interrelationships between the fear, anxiety, and BIS measures are relatively similar across gender. The test for equality of latent means across gender, however, showed significantly poorer fit when compared to the final measurement invariance model. An analysis of the relative difference between latent means showed that females had significantly higher scores on BIS, Social Fear, and Tissue Damage Fear but not for the STAI. By controlling for measurement error and demonstrating measurement invariance, these results provide solid support for previous RST studies that have shown gender differences on measures of negative emotionality (e.g., Carver & White, 1994).

Perkins et al. (2007) have suggested that in the context of RST there may be a crucial distinction between Tissue Damage Fear and measures such as Social Fear, STAI and the BIS scale. The current findings do tend to support the notion that Tissue Damage Fear is different to the constructs measured by the STAI and BIS scale. This may be considered surprising in the case of the BIS scale, as some of the item content of this scale would appear to reflect fear (e.g., "Even if something bad is about to happen to me, I rarely experience fear or nervousness"). Nonetheless, the BIS scale tends to contain content related to the affective experience of negative emotion, rather than sensitivity to specific aversive stimuli, as the FSS contains (Carver & White, 1994; Torrubia et al., 2001).

Overall, it is suggested that future research focus on the development of new RST measures that incorporate the Gray and McNaughton (2000) revisions. Current self-report measurement in RST research appears somewhat ad hoc on the basis that existing measures, such as the Carver and White (1994) BIS/BAS Scales, were developed in the context of the unrevised RST. While the BIS scale may be adequate for testing some form of higher-level punishment sensitivity, it may not be able to address more specific research questions related to the separation of anxiety and fear in the revised RST. Indeed, Carver (2004) suggested that the BIS scale may need to be revised to incorporate a more diverse range of processes related to threat sensitivity and avoidance. The use of the FSS as a measure of fear may also be somewhat problematic as, although it appears to be moderately effective in this role (e.g., Perkins et al., 2007), it was not designed to be administered to the general public (Arrindell, 1980; Wolpe & Lange, 1964) and so contains some potentially embarrassing items that are inappropriate for a nonclinical population. In addition, it would be preferable to have a specifically developed scale that incorporates the recent changes to RST, and is capable of distinguishing approach and avoidance of threat.

The results of the current study should be considered in the light of several limitations. First, the study only considered self-report questionnaire data. It is important that future studies examine the differential relationships these self-report measures may have with laboratory and real-world tasks designed to operationalize key facets of Gray and McNaughton's (2000) revised RST. For example, it may be possible to examine these differential relationships using one-way avoidance tasks (FFFS mediated) and twoway avoidance tasks (BIS mediated). Second, the CFA models tested in this study used item parcels as indicators rather than individual items. While using individual items can be more informative when seeking to examine aspects of an individual scale, in our case we were primarily interested in the structural relations between the latent traits. Further, it could be argued that the sample size in the current study would not have been sufficiently large to examine a model that contained 52 observed indicators and four latent traits, particularly when the sample was split by gender. Third, it should be noted that there has been some debate about the proper sequence of steps involved in testing for invariance (Byrne, 1998; Cheung & Rensvold, 2000; Vandenberg & Lance, 2000; Steenkamp & Baumgartner, 1998). Nonetheless, there is general support for the basic procedure for testing invariance that the current study followed (Brown, 2006).

References

Arrindell, W.A. (1980). Dimensional structure and psychopathology correlates of the fear survey schedule (FSS-III) in a phobic population: A factorial definition of agoraphobia. *Behavior Research and Therapy, 18*, 229–242.

Arrindell, W.A., Eisemann, M., Richter, J., Oei, T.P.S., Caballo, V.E., van der Ende, J. et al. (2003). Phobic anxiety in 11 nations. Part 1: Dimensional constancy of the five-factor model. *Behavior Research and Therapy, 41,* 461–479.

Arrindell, W.A., Pickersgill, M.J., Bridges, K.R., Kartsounis, L.D., Mervyn-Smith, J., van der Ende, J. et al. (1987). Self-reported fears of American, British, and Dutch university students: A cross-national comparative study. *Advances in Behavior Research and Therapy, 9*, 207–245.

Bandalos, D.L., & Finney, S.J. (2001). Item parcelling issues in structural equation modeling. In G.A. Marcoulides & R.E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 269–296). Mahwah, NJ: Erlbaum.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research.* New York: Guilford.

Byrne, B.M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.* Mahwah, NJ: Erlbaum.

Carver, C.S. (2004). Negative affects deriving from the behavioral approach system. *Emotion, 4*(1), 3–22.

Carver, C.S., & White, T.L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*, 319–333.

Caseras, X., Ávila, C., & Torrubia, R. (2003). The measurement of individual differences in behavioral inhibition and behavioral activation systems: A comparison of personality scales. *Personality and Individual Differences, 34*, 999–1013.

Cheung, G.W., & Rensvold, R.B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31*, 187–212.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Corr, P.J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience and Biobehavioral Reviews, 28*, 317–332.

Corr, P.J.J., & McNaughton, N. (2007). Reinforcement sensitivity theory and personality. In P. Corr (Ed.), *The reinforcement sensitivity theory of personality*. Cambridge, UK: Cambridge University Press.

Curran, P.J., West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16–29.

Eysenck, H.J. (1967). *The biological basis of personality.* Springfield, IL: CC Thomas.

Geer, J.H. (1965). The development of a scale to measure fear. *Behavior Research and Therapy, 3,* 45–53.

Gomez, R., Cooper, A., & Gomez, A. (2005). An item response theory analysis of the Carver and White (1994) BIS/BAS scales. *Personality and Individual Differences, 39*, 1093–1103.

Gomez, R. &, Gomez, A. (2005). Convergent, discriminant, and concurrent validities of measures of the behavioral approach and behavioral inhibition systems: A confirmatory factor analytic approach. *Personality and Individual Differences, 38,* 87–102.

Gray, J.A. (1970). The psychophysiological basis of introversion-extraversion. *Behavior Research and Therapy, 8*, 249–266.

Gray, J.A., & McNaughton, N. (2000). *The neuropsychology of anxiety* (2nd ed.). Oxford: Oxford University Press.

Heubeck, B.G., Wilkinson, R.B., & Cologon, J. (1998). A second look at Carver and White's (1994) BIS/BAS scales. *Personality and Individual Differences, 25*, 785–800.

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

Leone, L., Perugini, M., Bagozzi, R.P., Pierro, A., & Mannetti, L. (2001). Construct validity and generalizability of the Carver-White Behavioral Inhibition System/Behavioral Activation System Scales. *European Journal of Personality, 15*, 373–390.

Little, T.D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*(1), 53–76.

Little, T.D., Cunningham, W.A., Shahar, G., & Widaman, K.F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structur al Equation Modeling, 9*, 151–173.

Marsh, H.W., Balla, J.R., & McDonald, R.P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391–410.

McNaughton, N., & Corr, P.J. (2004). A two-dimensional neuropsychology of defence: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews, 28*, 285–305.

Muthen, L.K., & Muthen, B.O. (2006). *MPlus 4.1*. Los Angeles: Muthen & Muthen.

Oei, T.P.S., Cavallo, G., & Evans, L. (1987). Utility of the fear survey schedule with Australian samples of anxiety disorder patients. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 329–336.

Perkins, A.M., & Corr, P.J. (2006). Reactions to threat and personality: Psychometric differentiation of intensity and direction dimensions of human defensive behavior. *Behavioral Brain Research, 169,* 21–28.

Perkins, A.M., Kemp, S.E., & Corr, P.J. (2007). Fear and anxiety as separable emotions: An investigation of the revised reinforcement sensitivity theory of personality. *Emotion, 7,* 252–261.

Pickering, A.D., Corr, P.J., Powell, J.H., Kumari, V., Thornton, J.C., & Gray, J.A. (1997). Individual differences in reactions to reinforcing stimuli are neither black nor white: To what extent are they gray? In H. Nyborg (Ed.), *Scientific study of human nature: Tribute to Hans J. Eysenck at eighty* (pp. 36–67). Amsterdam: Pergamon/Elsevier Science Inc.

Spielberger, C.D., Gorsuch, R., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Steenkamp, J.E.M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78–90.

Steiger, J.H. (1990). Structural model evaluation and modification. *Multivariate Behavioral Research, 25*, 173–180.

Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences, 31*, 837–862.

Vandenberg, R.J., & Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–69.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*, 820–838.

Wolpe, J., & Lang, P.J. (1964). A fear survey schedule for use in behavior therapy. *Behavior Research and Therapy, 2,* 27–30.

Wolpe, J., & Lang, P.J. (1977). *Manual for the fear survey schedule.* San Diego, CA: Educational and Industrial Testing Service.

Zelenski, J.M., & Larsen, R.J. (1999). Susceptibility to affect: A comparison of three personality taxonomies. *Journal of Personality, 67*, 761–791.

Andrew Cooper

Department of Psychology
University of Wales, Swansea
Singleton Park, SA2 8PP
Swansea, UK
E-mail a.cooper@swansea.ac.uk