Sex differences are not attenuated by a sex-invariant measure of fear:

The Situated Fear Questionnaire

ABSTRACT

Widely-used fear questionnaires may exaggerate sex differences because they do not ensure sex invariance of items and conflate anxiety with fear. Beginning with 50 descriptions of fear-eliciting situations, we used Rasch analysis to identify sex-invariant items and Mokken analysis to establish unidimensional scalability. The resulting 27-item Situated Fear Questionnaire (SFQ) correlated highly with the widely-used Fear Survey Schedule, while demonstrating better discrimination between anxiety and fear. Sex differences in three samples were all in excess of *d* = 1.00 and were not explained by gender role adherence or anxiety levels. The hedonic tone associated with fear situations (ranging from distressing/alarming to thrilling/exhilarating) was rated as more positive by men and this was only partially explained by their lower reported fear.

Key words: fear, anxiety, Rasch, Mokken, sex difference, sex invariance, Fear Survey Schedule, avoidance

1. Introduction

In the field of emotion research, the complex multi-systemic nature of fear has proved challenging for measurement (Bradley & Lang, 2000; Schaefer, Larson, Davidson, & Coan, 2014). In experimental studies, fear has been assessed using fMRI indices of regional brain activation, EEG event-related potentials, and peripheral physiological indices that reflect activity of the autonomic nervous system and hypothalamic pituitary axis. Even with such 'objective' measures of fear reactivity, laboratory researchers need self-report measures to establish baseline fear levels and to examine candidate mediators of differential responding. For researchers interested in assessing fear outside the constraints of the laboratory, reliable and valid self-report methods are also essential. There has been considerable interest in understanding the evolutionary and developmental origins of sex differences in fear as well as their physiological and neural instantiation (Bangasser & Valentino, 2014; Campbell 2014; Derryberry & Rothbart, 1997). Sex differences in fearfulness are central to models of emotional regulation (e.g. aggression, behavioural inhibition), cognition (e.g. selective attention and recall, decision making) and personality (impulsivity, sensation seeking). The aim of the studies presented here was to develop and validate a sex-invariant self-reported measure of situational fear in a non-clinical sample.

1.1. The need for a new fear inventory

Given the extensive research interest in fear, it is surprising that the three major psychometric measures dedicated to its assessment were developed over thirty years ago. The Fear Questionnaire (FQ: Marks & Mathews, 1979) was designed to measure phobic fears in clinical populations. Sixteen questions ask the respondent to rate their tendency to approach or avoid stimuli which range from the sight of blood to travelling on public transport. A further six questions ask the participant to rate the degree to which anxiety symptoms are problematic in their daily lives. In a representative sample of the US population, no sex differences were found for the FQ total score or its subscales (Gillis, Haaga, & Ford, 1995). The Fear Survey Schedule (FSS) has two commonly-used versions. Like the FQ, the FSS-III (Wolpe & Lang, 1964) was developed for clinical

evaluation of phobic patients: It asks respondents to rate how much they are disturbed by 'fear or other unpleasant feelings' in response to 72 stimuli. Among undergraduates sampled from eleven nations, fear was higher among women although the magnitude of the sex difference varied over the subscales between an average of *r* = .16 (*d* = .32) for Agoraphobia and *r* = .33 (d = .70) for Harmless Animals (Arrindell et al., 2003). The FSS-II (Geer, 1965), developed for research purposes, presents 51 stimuli and respondents rate their fear on a 7-point scale from 'None' to 'Terror'. (Many items overlap with those in the FSS-III.) Effect sizes for the sex difference in total scores have been reported as *d* = 0.76 (Bernstein & Allen 1969) and d = 0.70 (Geer, 1965). There are a number of concerns about these scales, many of which have important implications for the accurate measurement of sex differences.

1.1.1 Over-representation of phobic items

The first is the heavy reliance on phobic items. Phobias are object-specific fears that are excessive, unreasonable, or out of proportion to the actual risk. Phobias are more prevalent among women than men (Park et al., 2013: Xu et al., 2012). Twice as many women (21.2%) as men (10.9%) meet the criterion for a single specific phobia and women are more likely to report multiple phobias (Fredrikson, Annas, Fischer & Wik, 1996). The inclusion of a large number of phobic items may tend to artificially increase the magnitude of the sex difference in fear. They may also distort it since phobic objects, by their nature, are unlikely to capture the full range of everyday fears. Although some objects or events which evoke phobic reactions in patients also provoke a lesser degree of fear in the general population (e.g. *the sight of blood, dentist appointments*), many phobic objects do not provoke fear in the majority of people and may even be regarded as pleasurable (e.g. *being a passenger in a car, being with a member of the opposite sex*). This limits the usefulness of such inventories for assessing fear and sex differences in fear in non-clinical populations.

1.1.2. Fear is not anxiety

A second and related issue is the conflation of fear and anxiety items in many inventories (Sylvers, Lilienfeld & LaPrairie, 2011). While some writers have treated the two concepts as

interchangeable (Izard & Ackerman, 2000; Wolpe, 1987) or as complementary facets of the same concept (Beck & Emery, 2005), many more have argued for their independence (McNaughton, & Corr, 2004; Öhman, 2008; Perkins, Kemp, & Corr, 2007; Sylvers et al., 2011). Fear is a short-lived, acute state that motivates avoidance and escape, after which the emotion quickly dissipates (Epstein 1972; Gray & McNaughton, 2000). The source of danger is specific, present and immediate (Adolphs 2013; Lang, Davis, & Ohman, 2000). Anxiety is a response to a stimulus which is ambiguous or uncertain in terms of threat. In contrast to fear, the temporal orientation is to the future rather than the present. Anxiety is a response to a possible threat that must be faced (Gray & McNaughton, 2000), where avoidance is impossible (Tellegen, 1982). It is associated with a protracted state of generalised hyper-vigilance and arousal that can persist even when the individual is removed from the triggering situation. The conflation of anxiety with fear is reflected in item wording. The FSS contains items assessing anxiety, while the Manifest Anxiety Scale contains items that ask explicitly about fear. The average correlation between anxiety measures and phobic fear measures (such as the FSS) is r = .46 (Sylvers et al., 2011). This has implications for the measurement of sex differences. Cross-culturally women score higher than men on Big Five neuroticism, especially the anxiety facet (Costa, Terracciano, & McCrae, 2001). They are twice as likely to suffer from anxiety disorders (Altemus, Sarvaiya, & Epperson, 2014; Bangasser & Valentino, 2014). To the extent that fear inventories simultaneously tap anxiety, women's scores may be artificially inflated relative to men's. Noting that correlations between anxiety and fear became smaller as the content overlap between the measures decreased, Sylvers et al. (2011, p. 133) identified a need for "refinement of self-report measures of fear and anxiety to reduce cross-contamination of constructs and construct irrelevant variance". We aim to develop a measure that addresses fear as an immediate emotional reaction rather than brooding anticipation.

1.1.3. Differential item responding by sex?

Another issue relevant to accurately estimating sex differences is the need to ensure that items perform invariantly over sex. In other words, an item should be as likely to be endorsed by a

high-scoring (or low-scoring) man as by a high-scoring (or low-scoring) woman. Without this equivalence, cross-sex comparisons of total scores can be misleading, analogous to comparing apples and oranges. No test of differential item functioning has been conducted on extant fear measures. The development of a sex-invariant fear questionnaire was a key aim of this study. This was established using Rasch analysis to measure differential item functioning which assesses whether an item is more 'difficult' for one sex than the other. Classical test theory which has been used to analyse inventories such as the FSS (e.g. Arrindell, 1980) is based on covariance between items. Factor analysis assumes that scores can be summed to the extent that they load on a common factor. This means that two individuals (or two sexes) could receive the same trait score even though they have endorsed non-overlapping sets of items. For example, in a test of arithmetic ability, a person who correctly answered *2 + 2* would receive the same score as someone who correctly answered *234 – 153 / 9.* By contrast, Rasch analysis examines the structure of the items based on ordering them in terms of difficulty. It is predicated on the premise that an individual who achieves a high overall score would be more likely to get the second question correct than someone who gets a lower overall score. This can be applied to personality traits also to reveal the dimensionality of the items. A 'difficult' item corresponds to one which is endorsed only by those with a high level of the latent trait. In the present study, we used Rasch analysis to compare the 'difficulty' structure of items for the two sexes. Differential item functioning was used to reduce the initial item number, retaining only those which were sex invariant. This was followed by Mokken analysis on a new sample to confirm the hierarchical scalability of the items.

1.1.4. Gender, fear and self-presentation

An additional concern about sex differences in fear is whether they are artefacts of gendered self-presentation, resulting from men's reluctance to admit fear on self-report instruments (Jansz, 2000). Fear scores are negatively associated with masculinity (Arrindell, 2000) and positively with femininity (Tucker & Bond, 1997). Nevertheless, studies which control for gender role adherence still find a significant effect of biological sex (Arindell, Kolk, Pickersgill & Hageman, 1993;

Dillon, Wolf & Katz, 1985) and informing participants that the honesty of their responses are verifiable by physiological measures does not eliminate the sex difference (McLean & Hope, 2010; Pierce & Kirkpatrick, 1992). We examined this issue further in the present study.

1.1.5. Dimensionality

We also addressed three further concerns about existing fear measures. The first was questionnaire dimensionality. Both the FQ and FSS are multi-dimensional. The four factors of the FQ (Agoraphobic; Blood-Injury; Social; Anxiety-Depression) and the five factors of the FSS-III (Social; Agoraphobic; Bodily Injury, Death and Illness; Sex and Aggression; Harmless Animals) show evidence of stability across sex and culture (Arrindell, Emmelkamp, & van der Ende, 1985; Arrindell et al., 2003). However, the orthogonal nature of the factors (as well as the specificity of phobic disorders) means that subscale scores should not be summed to create total scores (Arrindell et al., 1984). This is problematic for researchers who want a global fear measure for research proposes. We therefore sought to create a unidimensional scale.

1.1.6. Number and range of items

A further concern was ensuring a sufficient number and range of items. Large-scale fear surveys use single items measuring the frequency or intensity of experiencing fear (Brebner, 2003; Fischer, Mosquera, van Vianen, & Manstead, 2004; Simon & Nath, 2004). Fear has also been assessed as one component of more general psychometric mood inventories. On the Positive and Negative Affectivity Scale (PANAS: Watson, Clark & Tellegen, 1988), two of the 10 negative mood items ('scared', 'afraid') assess fear and respondents are asked to rate 'to what extent' (the wording does not clearly distinguish between frequency and intensity) they have experienced these emotions in a given time frame. Similarly, three of the 12 emotions on the Differential Emotions Scale (Izard, Libero, Putnam, & Haynes, 1993) tap the frequency of experiencing fear (e.g. 'Feel scared, uneasy, like something might harm you') over a specified time period. The use of a single item (or a small number of items) limits reliability and prevents examination of the range of situations which elicit fear. To ensure coverage of a full range of fear-inducing situations and to compare their relative

'difficulty' for men and women, it was important to include items ranging from mildly to extremely frightening. This means that some of the situations described may not have been experienced by respondents. There has been debate about the accuracy with which respondents are able to rate their reaction to hypothetical situations (Jackson & Mauran, 1996; Zuckerman, 1996). Although we recognise these concerns, we believe that they are less applicable in the case of fear measurement. Fear is an emotion adaptively designed to keep the organism safe by causing avoidance or withdrawal. The capacity to anticipate fear is what allows us to make informed choices about future behaviour (Loewenstein, Weber, Hsee & Welch 2001). A person who refuses an invitation to take part in a parachute jump does so based on the degree of fear they would experience were they to jump. Without this ability, a person would have to experience potentially lethal risks before they could establish whether or not they were frightening. We therefore believe that respondents are in a good position to provide reasonable estimates of their fear in situations they may not have directly experienced.

1.1.7. Need for context

A final concern is with item wording. In the FSS, many items are presented as single words (*death, heights*) or short phrases without elaboration (*meeting someone for the first time*), so that the questions lack detail and context. For example, for many people the degree of fear associated with *snake* may crucially depend on situational factors (e.g. proximity to target, availability of escape routes, whether the snake is uncaged, etc.). Fear associated with *airplane travel* may depend on whether or not there is turbulence. Thus fear-ratings may vary widely on account of the lack of specificity of the items (Lick, Sushinsky, & Malow, 1977).

1.2 Aims of the present study

To overcome these issues (over-representation of phobic items, conflation of anxiety and fear, differential item functioning, gendered self-presentation, questionnaire dimensionality, item number and wording), we developed a measure of fear intensity as a situated and short-lived emotional response to risky or potentially dangerous situations that might be encountered in the

real world. The items were designed to represent a unidimensional scale with a hierarchical structure that was invariant over sex: This was assessed by comparing Rasch scores calculated for each item between males and females. The sex-invariant items retained following Rasch analysis were subsequently submitted to a nonparametric item response method, another Rasch analysis and also confirmatory factor analysis. We then examined the convergent validity of the scale in relation to a range of related measures and the effect of sex on fear scores after controlling for variation in gender role adherence.

In addition, we considered whether sex differences in fear might explain the sex difference in enjoyment of risky activities. It is paradoxical that while fear can motivate avoidance, it can also be pleasurable. The popularity of theme parks and scary movies attest to this. Men's greater average taste for risk is visible in their greater involvement in extreme sports (Robinson, 2008), and their faster and more reckless driving (Beattie, 2008). Men score significantly higher than women on Zuckermans' Sensation Seeking Scale, a finding attributable to the degree of risk involved since sex differences are not found on the subscale of Experience Seeking which measures preferences for new experiences that are not marked by risk e.g. eating exotic food (Zuckerman, 1994). Furthermore, sex differences in reported risk-taking are completely mediated by differences in sensation seeking scores (Zuckerman & Kuhlman, 2000). Men make riskier decisions than women (Croson & Gneezy, 2009) and the magnitude of the sex difference increases as the risks involved become more life-threatening, leading Byrnes, Miller and Schaffer (1999, p. 378) to conclude "fear responses may explain gender differences in risk taking more adequately than the cognitive processes involved in the reflective evaluation of options". Women are more sensitive to the dangers associated with risky actions (Harris & Miller, 2000; Harris, Jenkins & Glaser, 2006). Experimentally framing a risky choice in such a way as to draw attention to potential negative consequences reduces risk taking in women but not men (Gabriel & Williamson, 2010). In the present study, in addition to completing fear ratings, we asked respondents to rate the hedonic valence, ranging from negative (distress) to positive (exhilaration), associated with the

questionnaire's various fear-provoking situations. Given findings of elevated female fear and danger sensitivity, we hypothesised that a sex difference found in the hedonic valence of these risky situations would be mediated by women's higher level of fear.

2. Method

2.1 Situated Fear Questionnaire: Initial item selection

Fifty items were created based on personal experience, peer discussion and past inventories (see Table 1). The items represented fear across a broad range of situations including: fear of physical trauma, illness or death; situations of heightened personal vulnerability; fear of public humiliation; fear of harm to loved ones. Items also varied in terms of whether the situations had been freely entered into (e.g. roller coaster rides, scuba diving) or had occurred unexpectedly (e.g. aircraft turbulence, skidding on ice while driving). For each item participants rated their fear on a 5-point Likert scale *(Not at all afraid; A little afraid; Moderately afraid; Very afraid; Extremely afraid).*

2.2 Participants and materials

To determine the required sample sizes, we conducted a priori power analyses. For Rasch analysis, a sample of 245 participants gives 99% confidence that no item calibration is more than 0.5 logit away from its stable value (Linacre, 1995). For the other analyses, a sample of N=245 provides 99% power to detect a correlation of .30 and an effect size of d = .50. Therefore N=248 was the minimum sample size (Sample 3) with the other two samples being larger (Sample 1 N=301; Sample 2 N=320). The study was approved by the Ethics Subcommittee of the Durham University Psychology Department and informed consent was given by all participants.

2.2.1 Sample 1

Participants. An opportunity sample of 301 British participants (209 women and 92 men) completed the questionnaires online. Ninety seven per cent (N=292) participants were in the 18-25 age group. The participants were recruited for the study via a university participant pool and social networking sites.

Measures. Participants completed the 50-item Situated Fear Questionnaire (SFQ). In addition, they rated the hedonic tone (subjective experience) of the same 50 situations. This was measured on a 5-point Likert scale *(Extremely distressing/alarming; Quite distressing/alarming, Neutral, Quite exhilarating/thrilling; Extremely thrilling/exhilarating*. This scale was scored from 1 (Extremely alarming) to 5 (Extremely thrilling/ exhilarating).

2.2.2 Sample 2

Participants. Three hundred and twenty Romanian participants took part (236 women and 83 men). Of those reporting their age, ninety three per cent (N=287) were in the 18-25 age group (68 men, 219 women). All participants were volunteer first and second year students at Babes-Bolyai University (Cluj, Romania) who completed the questionnaires in class time.

Measures. Participants completed the 50-item SFQ as described above. In addition, they completed the Zuckerman Kuhlman Personality Inventory (ZKPQ: Zuckerman d & Kuhlman, n.d.); Approach Avoidance Temperament Questionnaire (ATQ: Elliott & Thrash, 2010); and the Harm Avoidance scale from the Multidimensional Personality Questionnaire (MPQ: Tellegen, 1982). The ZKPQ had been already been translated into Romanian for a previous study (Opre & Albu, 2010). The remaining questionnaires were translated into Romanian by two of the authors (RD and AO) and then independently back translated by an academic in the Faculty of Letters (Languages and Literature) at Babes-Bolyai University.

2.2.3 Sample 3

Participants. Two hundred and forty eight British participants took part (116 men and 132 women). All but one participant were in the 18-25 age group. Participants were invited to take part by undergraduate psychology students as part of a practical assignment. They completed the questionnaires online.

Measures. Participants completed a shortened 36-item version of the SFQ (see Table 1). As in Sample 1, they also rated the hedonic tone of these situations (*Extremely distressing/alarming* to *Extremely thrilling/exhilarating*). They completed the Fear Survey Schedule (FSS: Wolpe & Lang,

1964); Stress Reaction and Harm Avoidance scales of the MPQ (Tellegen, 1982); and the Instrumental (masculine) and Expressive (feminine) scales of the Personal Attributes Questionnaire (Spence & Helmreich, 1978).

3. Results

3.1 Situated Fear Questionnaire: Psychometric analysis

The psychometric analyses included item response theory methods and factor analysis. We used both Rasch and Mokken analysis to investigate unidimensionality and the hierarchical ordering of items. We first used Rasch analysis to investigate differential item functioning over sex as the procedures are more well-established for this purpose. We then used Mokken analysis to check the scale on a new sample of participants as another method of investigating unidimensionality. The items that were considered to form a scale were then examined from a Rasch perspective. Lastly we tested the factor structure with ordinal confirmatory factor analysis.

3.1.1 Unidimensionality and invariance (Rasch analysis).

We began by using Rasch analysis to reduce item redundancy and create a set of items that were invariant over sex. Rasch analysis takes a set of item responses and uses their likelihood of endorsement ('item difficulty') as a means of testing whether they fit a latent unidimensional scale. Respondents are then assigned a score on this scale. The analysis was performed on RUMM2030

Using data from Sample 1, we first checked the person fit statistics for the 50 item scale. The person fit statistics indicate how well each individual's responses follow the Rasch model, by identifying any participants who are following idiosyncratic response sets rather than responding differentially to the items. Although six participants had relatively poor infit mean square statistics of >2, only one of these had clearly responded inappropriately and answered with an alternating pattern of 1 and 5. This participant was not considered in the remaining analysis. The item statistics were encouraging (see Table 1). The fit measures indicate the ratio of observed variance to that expected by the Rasch unidimensional model. The infit statistic is an adjustment to the outfit value which reduces the effect of outliers. A mean square fit statistic of 1 indicates ideal fit and values of

over 1.5 indicate poor fit. All of the items had fit statistics in the acceptable range. All of the items also had positive point-measure correlations (between the item and Rasch total score) with a median correlation of .49.

Next we considered the issue of differential item functioning by sex. This is examined by calculating Rasch measure scores for each item separately for males and females, and then comparing the scores. Rasch analysis takes account of the 'ability' of the participants (in this case their fearfulness) when computing item difficulty so any sex differences between scores on the items are not related to differences in fear. Rather, a significant difference between the scores indicates that the item is not being treated in the same way by the two sexes. This would mean that the instrument as a whole was not functioning the same way for male and female participants. The statistical significance was determined by using Tristan's (2006) DIF analysis with an adjustment for sample size (see Table 1). (We did not correct for multiple tests because our aim was to be conservative in ensuring that only gender-neutral items were retained.) There were fourteen items that showed significant differences (4, 6, 10, 15, 19, 20, 21, 25, 38, 41, 43, 45, 48, 49). These were split fairly evenly in terms of which sex showed the bias. We removed these items from the remainder of the analyses.

We then checked the unidimensionality of the 36 item fear measure by conducting a PCA of the residuals left after the Rasch component had been removed (Mavranezouli, Brazier, Young, & Barkham, 2011). Monte Carlo studies have shown that an eigenvalue of over 2 should be investigated as a source of a potential factor (Linacre, 2012). The Rasch unidimensional component had an eigenvalue of 35.7 and accounted for 49.8 per cent of the variance. The second potential component had an eigenvalue of 2.5 and accounted for 3.5 per cent of the variance. The factor sensitivity ratio (Wright & Stone, 2004) which compares the residual variance to that explained by Rasch dimension was .07. This was acceptably low and suggested that the second potential component did not affect measurement accuracy. The impact of this factor on person scores (Smith, 2002) was also negligible. The item reliability of the 36-item questionnaire (the extent to which the

hierarchical ordering of items would be replicated in new samples) was high at .99. It also had excellent person reliability of .90 (person reliability indicates whether the ordering of persons would be replicated if the sample was given a parallel set of items). Values of above .7 are considered acceptable and those above .8 are considered good (Bond & Fox, 2007). The Cronbach's alpha of the 36-item scale was α = .91.

3.1.2 Scalability and factor structure (Mokken and CFA analysis).

We then examined the hierarchical scalability of the 36 item measure on a new sample (Sample 3). To do this, we used Mokken analysis, a non-parametric form of item response theory (Mokken, 1971; Molenaar, 1982). Although based on Guttman scaling, Mokken does not assume error-free data. Nor does it include assumptions about the sigmoid shape of item characteristic curves that can result in the rejection of many items and so decrease the reliability of the resultant measure. Invariant item ordering (IIO) means that items can be ranked according to their 'difficulty' (or frequency of endorsement), allowing for hierarchical ordering of scale items. This requires the calculation of three coefficients. Coefficient *H* for each item provides a measure of the item's scalability (and unidimensionality). From these values, an *H* coefficient for the full scale can be calculated which indexes the extent to which scale items accurately order respondents. H^T reverses the role of persons and items, and thus indexes the extent to which a sample of individuals agree on the ordering of the items (Sijtsma, Meijer & van der Ark, 2011). Taken together, *H* and H^T are indicative of the strength and structure of a scale. Values greater than .3 indicate a hierarchical scale (van Schur, 2003).

Data were analysed using 'mokken', a programme in the package of freely-available statistical software 'R' (van der Ark 2007). This programme has an automated item selection procedure which places items onto scales directly, making it easy to see if the items fall onto one or more scales. We entered all 36 items and specified a solution that corresponded to a minimum acceptable scale *H* = 0.3. Two items (34, 44) were rejected as not fitting the first extracted scale. Although this scale had good internal consistency (*rho* = .93) and H^T = .50, the scalability was only

marginally acceptable at H = .31. Examination of the item H values indicated that seven items were unacceptably low at H < .28 (2, 5, 8, 9, 16, 30, 33). These 9 items were removed and the analysis was run again on the remaining 27 items. A single scale was identified with stronger scalability indices for person ordering (H = .36) and item ordering (H^T = .53). All items had H values > .31. Backward selection for item removal (based on the Crit criterion for monotonicity) indicated no items that required removal. Internal consistency remained high (*rho* = .93) for the final 27 item scale. Rasch analysis was then conducted on the 27 item scale. The fit statistics were all in the acceptable range with the worst item having an outfit statistic of 1.2, and the median point-measure correlation was 0.57. The item reliability of .99 and person reliability of .95 were also good. The PCA of residuals revealed that the Rasch component had an eigenvalue of 38.1and explained 58.6% of the variance. The sensitivity ratio was also acceptable at 0.06. Overall these measures are better than those from the 36 item scale, even though those were acceptable from a Rasch standpoint. The correlation between male and female mean scores across the 27 items was r = .97, p<.001 (Sample 1), r = .94, p<.001 (Sample 2) and r = .99, p<.001 (Sample 3).

Finally, to confirm the single factor structure of the 27 items with classical test theory, we used confirmatory factor analysis. This was run the R 'lavaan' program for ordinal data which uses diagonally weighted least squares to estimate model parameters. The fit statistics indicated a unitary construct was a good fit to the data, CFI = .98, RMSEA = .06 (90% CI [.05, .07]), although the chi square value remained significant, $X^2(324) = 610.49$, p<. 001.

3.2 Convergent and construct validity

Mean scores were computed for each participant on the 27 SFQ items. We first examined the correlation between the SFQ and the most widely used measure of fear, the Wolpe and Lang's (1964) Fear Survey Schedule (See Table 2). This correlation was strong and significant, r = .69, p<.001 (Sample 3).

In line with our interest in the relationship between fear and anxiety, we computed correlations between the SFQ and MPQ Stress Reaction Scale, r = .29, p<.001 (Sample 3) and the

ZKPQ Neuroticism-Anxiety scale, r = .31, p<.001 (Sample 2). These moderate correlations suggest that, although SFQ fear is conceptually distinct from anxiety, more anxious people report higher levels of situated fear. However the correlation of the MPQ Stress Reaction scale with FSS, r = .41, p<.001, was significantly higher than with SFQ, t(245) = 2.61, p<.01. This suggests that the FSS with its large complement of phobic items is more strongly associated with anxiety than the SFQ.

Fear is an emotion associated with avoidant motivation. The MPQ Harm Avoidance scale explicitly assesses avoidance of danger and excitement, and this scale as expected showed a positive correlation with SFQ (Sample 3: r = .39, p<.001; Sample 2: r = .32, p<.001) and with the FSS measure, r = .37, p<.001. We also examined correlations with measures of more general avoidance and approach temperament (ATQ; Eliott & Thrash, 2010) in Sample 2's data. Avoidance temperament is defined as a general neurobiological sensitivity to negative stimuli. The SFQ was positively correlated with avoidance, r = .33, p<.001 although it was unrelated to approach temperament, r = .09, p=.10.

3.3 Sex differences

Sex differences across the three samples are presented in Table 3. Effect sizes were in excess of d = 1.00 indicating a large effect which was consistent across all three samples. In a 2 (sex) by 3 (sample) analysis of variance, the main effect of sex was significant F(1,847) = 251.78, p<.001, d = 1.09. Although the Romanian sample scored significantly lower than both British samples, F(2,847) = 14.69, p<.001, there was no significant interaction between sample and respondent sex. In Sample 3, respondents completed both the SFQ and the Fear Survey Schedule. The effect size for the sex difference on the FSS was of medium size, t(246) = 4.12, p<.001, d = .53, while it was considerably larger for the SFQ, t(243) = 8.16, p<.001, d = 1.04.

The sex difference was only slightly reduced when measures of anxiety were controlled. In Sample 2, when ZKPQ Neuroticism-Anxiety was entered as a covariate, the effect size dropped from d = -1.23 to d = -1.12, F(1,303) = 82.68, p<.001. In Sample 3, the introduction of MPQ Stress Reaction reduced the effect size from d = -1.04 to d = -1.00, F(1,242) = 59.12, p<.001.

It has been suggested that sex differences in self-reported fear result from (or are exaggerated by) a reluctance to acknowledge fear associated with the male gender role. In Sample 3, participants also completed the PAQ which provides measures of instrumentality (associated with a masculine gender role) and expressivity (associated with a feminine gender role). There was a significant sex difference with women scoring higher in expressivity (femininity), $t(246) = 2.06$, $p = .04$, although men's score on instrumentality (masculinity) was not significantly higher than women's, $t(246) = -1.75$, $p = .08$. Both scales were significantly but modestly correlated with SFQ scores in the expected direction. To examine whether the association between biological sex and fear was mediated by gender roles, we used 10,000 bootstrapped samples to estimate confidence intervals. The total effect for the model, $B = .56$ [*CI* -.69, -.42] and the direct effect, $B = -.52$ [*CI*-.65, -.38] were both significant. The indirect effect of sex on fear through masculinity and femininity was not significant, $B = -.04$ [*CI* -.09, -.01]. Biological sex remained a highly significant predictor of fear when gender role was controlled.

3.4 Fear and enjoyment of danger

The SSImp scale (ZKPQ) is composed of two factors (Zuckerman & Kuhlman, n.d.). It has been argued that an aggregate score conflates two distinct constructs. Sensation seeking is associated with willingness to engage in risky and dangerous activities such as parachute jumping. However such activities typically involve a considered decision and careful preparation. By contrast, impulsivity assesses a tendency to act without thought or planning. A meta-analysis by Cross, Copping and Campbell (2011) found a sex difference on the former but not the latter. This was also true in the present data (Sample 2: Sensation seeking $t(313) = -2.92$, $p = .004$, $d = 0.45$; Impulsivity $t(312) = -1.08$, $p = .28$, $d = 0.15$). As expected, there was no correlation between SFQ and Impulsivity, $r = -.09$, $p = .11$, but there was a significant although modest correlation with Sensation Seeking, $r = -.15$, $p = .008$ (see Table 2).

The ZKPQ Sensation Seeking subscale includes statements that represent a carefree approach to life (e.g. "I'll try anything once"; "I Prefer friends who are excitingly unpredictable") but

only a few items refer explicitly to frightening or risky behaviours. In order to examine associations between fear and enjoyment of risk while holding situation constant, we computed the correlation between participants' ratings of fear and their ratings of hedonic tone (from 'alarming' to 'exhilarating') across the 27 situations (Samples 1 & 3 excluding multivariate outliers; N= 668). The correlation was highly significant, r = -.51, p<.001. The correlation was of similar magnitude for males, r = -.43, p<.001 and females, r = -.42, p<.001. There was a significant sex difference, $t(668)$= 9.27, p<.001, d = 0.74 with women rating the hedonic tone associated with fear more negatively than men, $Mean_{women}$=2.27, SD=.40; $Mean_{men}$ = 2.58, SD=.44). In Sample 3, for whom PAQ scores were available, we examined the association between gender role and reported hedonic tone. These correlations were not significant for either Masculinity, r = .09, p=.13 or Femininity, r = .12, p=.07.

We used mediation analysis to investigate whether the sex difference in hedonic tone was explained by the sex difference in fear. We used 10,000 bootstrapped samples to estimate confidence intervals. The total effect for the model was B = .30 [CI .24, .36]. The indirect effect of fear was significant, B = .18 [CI .14, .23] although the direct effect of sex also remained significant B = .12 [CI.05, .18]. Thus approximately 60 per cent of the variance in the sex difference in enjoyment of risky situations was explained by men's lower levels of fear in these same situations.

4. Discussion

We used three samples to refine the Situated Fear Questionnaire and to examine its convergent validity. Our first sample completed the initial 50-item version of the questionnaire. Rasch scores were computed for men and women and the item-level results compared. Because Rasch analysis takes account of the 'ability' of the participants (in this case their fearfulness) when computing item difficulty, a discrepancy indicates that the item is not of equivalent difficulty (ease of endorsement) for men and women. This analysis led to the exclusion of fourteen items.

On a number of these excluded items, men scored higher than would be expected given their overall lower fear level. These included coping practically and emotionally with another

person's health emergency (items 20, 25) and failing to live up to obligations to a workmate, employer or friend (items 10, 21, 45). Unexpectedly, given men's riskier driving, they also scored higher than expected on fear when being 'cut up' by another motorway driver (item 6). Similarly, despite their greater involvement in risky recreational activities, their scores for a roller coaster ride (item 4), a cliff top walk (item 38), and when swimming during a shark warning (item 49) were unexpectedly high. It may be that the conscious experience of fear is not equally distressing to the two sexes and this is supported by the significant sex difference in hedonic tone with men experiencing fear situations as less alarming and distressing than women. For men, the fear inherent in these recreational activities may have associated elements of thrill or exhilaration. Relative to their overall greater level of fear, women scored higher than expected on items reflecting personal vulnerability including being isolated and alone (items 15, 19, 41) and receiving threatening anonymous calls (item 43).

With the above items excluded and the questionnaire reduced to 36 items, the Rasch component accounted for fifty per cent of the variance and the questionnaire had excellent item and person reliability. On a new sample, we then used Mokken analysis to investigate the scalability of the items. A further nine items were removed resulting in a 27-item scale. Confirmatory factor analysis indicated a single factor model. Internal consistency was excellent and the correlation between male and female mean scores across the 27 items was near unity. The most fear-inducing situations were traumatic injury (loss of sensation in the legs following a car accident; risk of blindness following retinal detachment). These were followed by inability to escape a fire, loss of contact with a family member following a local terrorist attack, and losing control of a car on an icy road. The least fear-provoking situations were an electrical power outage when alone in a house, being caught outside in a thunder storm, being called aside by customs officials, being stranded abroad without money and finding oneself drunk and alone at a party.

Sex differences on the SFQ fear measure were strong with effect sizes exceeding d = 1.00 in each of the three samples. The magnitude of this sex difference exceeds those found in a number of

domains where sex differences are well established, including aggression (Archer, 2004), sensation seeking (Cross, Copping & Campbell, 2011) and sexual behaviour (Petersen & Hyde, 2010). Furthermore, the present sex difference was found on a questionnaire which was explicitly designed to be sex invariant in content. In recent years, there have been only a few attempts to apply item response theory to the measurement of domains where sex differences are of interest. Invariance as a function of sex was examined (and confirmed) in a recent analysis of the Empathy Quotient scale (Allison, Baron-Cohen, Wheelwright, Stone, & Muncer, 2011). Sexual symmetry in admissions of intimate partner violence has been a contentious issue and consequently one in which item invariance over sex is of key importance. Although Jose, Olino and O'Leary (2012) reported measures of item difficulty for men and women on the Conflict Tactics Scale, these were not explicitly used to modify the measurement instrument. In developing psychometric measures, items showing differential functioning should be revised or discarded and this is especially critical in domains where sex differences are controversial or disputed. Differential item functioning means that there is a significant difference between the proportion of men and women at the same fear level who positively endorse the item. This can make comparison between men and women in overall scores misleading.

The substantial sex difference in fear was not attributable to personality traits associated with male and female gender roles. It has been suggested that instrumental traits associated with masculinity (such as 'stands up well under pressure') are antithetical to the public acknowledgement of traits like fear that can indicate vulnerability or weakness. Reciprocally, expressive traits associated with femininity may enhance readiness to recognise emotions such as fear. Between 1970 and 1995, women's scores on the PAQ masculinity scale rose significantly, although men's femininity scores did not alter (Twenge 1997; 2009). In line with this, we found a significant sex difference for femininity only. Mediation analysis indicated that gender role adherence did not explain the sex difference in fear. This supports an expanding body of neuroimaging and

electrophysiological research suggesting a biological basis for the greater reactivity of women than men when exposed to the same fear-inducing stimulus.

The effect size for the sex difference in fear was only slightly reduced when anxiety was statistically controlled. There is increasing recognition that fear and anxiety are distinct (although related) concepts and there have been calls for more sensitive measures that are capable of distinguishing between them. This is especially relevant for the estimation of sex differences. Sylvers et al. (2011) in their meta-analysis calculated the weighted mean correlation between phobic fear measures and self-reported anxiety at r = .46 and, in the present data, the FSS correlated with the MPQ Stress Reaction scale at r = .41. However the SFQ showed a lower correlation with both Stress Reaction and Neuroticism-Anxiety suggesting that the SFQ is less contaminated by variance associated with anxiety. Attempts to distinguish between fear and anxiety have been made at conceptual and definitional levels (Ohman, 2008; Tellegen, 1982), as well as in terms of evolutionary function (Boyer & Bergstrom, 2011), proposed psychological systems (Gray & McNaughton, 2000), differential pharmacological sensitivity (McNaughton & Corr, 2004) and neuroanatomy (Davis, Walker, Miles, & Grillon, 2010). Nonetheless, if two biologically distinct systems exist, it would be surprising if there was not a positive association between them. The strength with which an individual experiences anxiety (a heightened anticipation of potentially challenging situations and concern about coping with them) is likely to have implications for fear (an intense phasic avoidant response to a short-lived threat)---as well as vice versa. Moderate correlations may reflect a genuine association rather than a measurement issue of item overlap (Perkins et al., 2007).

As expected, a lower threshold for experiencing fear was associated with a more avoidant temperament. Elliott and Thrash (2010) aligned their two measures of temperament with Gray's original Behavioural Activation System (BAS) and Behavioural Inhibition System (BIS). However, in an important revision of the theory, Gray and McNaughton (2000) made a key alteration to the proposed function of the BIS system: Rather than sensitivity to conditioned punishing stimuli, the new role of BIS was to resolve approach-avoidance and avoidance-avoidance conflicts. BIS was seen

as central to anxiety in terms of orchestrating behaviour where the organism cannot avoid engagement with the situation. By contrast, behavioural responses to imminent danger were proposed to be managed by the Fight/Flight/Freeze system (FFFS). This new formulation underlined the distinction between anxiety (associated with approach decisions) and fear (associated with avoidance reactions).

There have been efforts to develop psychometric measures of the reconceptualised tri-partite fear system (FFFS) using classical test theory. Jackson (2009) and Corr and Cooper (in preparation) developed items assessing respondents' rated likelihood of engaging in the three different behaviours in response to fear. It would be of interest to examine these behaviourally-oriented scales in relation to the SFQ scale which focuses on the intensity of affective reactions, and to examine associations between different SFQ fear situations and behavioural response tendencies. However there are psychometric issues with the Fight scales. Although Jackson's (2009) Flight and Freeze scales showed a significant inter-correlation (and significantly higher scores in women), the Fight scale was uncorrelated with either (and showed a significantly higher score in men). Similarly, Corr and Cooper's (in preparation) FFFS scale included items referring only to Flight and Freeze; Fight items were captured by a separate scale called 'defensive fight'. More recently, a lexical approach was taken by Maack, Buchanan and Young (2015) who identified synonyms for the emotional states corresponding to the three behavioural responses e.g. 'disengaged' (Freeze), 'argumentative' (Fight), 'terrified' (Flight). This study also reported difficulties identifying a clean Fight scale. Nevertheless, it would be of interest to examine associations between this scale (which assesses experiential states associated with the various response styles) and the current measure of situated fear intensity.

Finally we examined whether fear is experienced as more distressing by women than by men and whether this may explain women's greater avoidance of risky and dangerous activities. Affective responses ('feelings') play an important role in appraising the rewards and costs of risky activities (Loewenstein et al., 2001), with fear intensity acting as an emotional cue for stronger avoidance. Our data indicate a marked sex difference in fear and this is congruent with studies that

have attributed women's desistance from risky activities to their greater emotional sensitivity to negative outcomes (e.g. Harris et al., 2006). However it is also possible that the same intensity of fear is experienced differently by women and men. It is paradoxical that while fear can motivate avoidance, it can also be pleasurable, even appetitive. Extreme sports enthusiasts report fear but also find the experience thrilling. Women rated the various fear situations in the SFQ as more distressing and alarming than men. Mediation analysis indicated that this sex difference was chiefly attributable to the sex difference in fear intensity: Women experience greater fear than men and greater fear is experienced as more distressing. Nonetheless, after controlling for women's greater fear, there remained a significant sex difference in the rated enjoyment of the fear experience.

In summary, we found sex differences in fear using questionnaire items that were selected to show no gender bias, to generate a meaningful aggregate score, to assess fear rather than anxiety and to focus on situated fears rather than reactions to phobic stimuli. Despite near perfect agreement between the sexes about the relative ranking of the situations in terms of fear elicitation, the sex difference in reported fear was large.

Adolphs (2013) has called for closer integration between the ecology of fear (its triggers and phenomenology as they are experienced 'in the wild') and studies of the neural circuitry that underlie its wide-ranging effects. We hope the SFQ can help to bridge this gap. Experimentally, the questionnaire can be of use in establishing sex (and individual) differences in fearfulness that may mediate differential neural, physiological and behavioural reactivity to manipulations. Beyond the laboratory, sex invariant measurement allows hypotheses about the evolutionary, developmental and cultural origins of sex differences in fear to be pursued more rigorously. We hope that it will encourage researchers in the development of models of fear in relation to the wide range of sex-linked phenomena in which it has been implicated including psychopathy, PTSD, inhibitory control, decision-making, attentional bias, risky impulsivity, sensation seeking, disgust, and aggression.

References

Adolphs, R. (2013). The biology of fear. *Current Biology 23* (2), R79–R93. doi:10.1016/j.cub.2012.11.055

Allison, C., Baron-Cohen, S., Wheelwright, S.J., Stone, M.H. & Muncer, S. (2011). Psychometric analysis of the Empathy Quotient (EQ). *Personality and Individual Differences, 51,* 829–835. doi:10.1016/j.paid.2011.07.005

Altemus, M., Sarvaiya, N., & Epperson, N. (2014). Sex differences in anxiety and depression clinical perspectives. *Frontiers in Neuroendocrinology, 35,* 320–330. doi:10.1016/j.yfrne.2014.05.004

Archer, J. (2004) Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology, 8,* 291–322. doi:10.1037/1089-2680.8.4.291

Arrindell, W.A. (1980). Dimensional structure and psychopathology correlates of the fear survey schedule in a phobic population. *Behavior Research & Therapy, 18,* 229−242.

Arrindell, W. A. (2000). Phobic dimensions. IV. The structure of animal fears. *Behaviour Research and Therapy*, *38*(5), 509–530. doi:10.1016/S0005-7967(99)00097-2

Arrindell, W.A., Eisemann, M., Richter, J., Oei, T.P.S., Caballo, V.E., van der Ende, J. … Cultural Clinical Psychology Study Group (2003). Phobic anxiety in 11 nations Part 1: Dimensional consistency of the five factor model. *Behaviour Research and Therapy, 41,* 461-479. doi:10.1016/S0005-7967(02)00047-5

Arrindell, W.A., Emmelkamp, P.M.G., & ven der Ende, J. (1984). Phobic dimensions: I. Reliability and generalizability across samples, gender and nations. *Advances in Behaviour Research and Therapy, 6,* 207-254.

Arrindell, W. A., Kolk, A. M., Pickersgill, M. J., & Hageman, W. J. (1993). Biological sex, sex role orientation, masculine sex role stress, dissimulation and self reported fears. *Advances in Behaviour Research and Therapy, 15,* 103–146. doi:10.1016/0146-6402(93)90018-W

Bangasser, D.A. & Valentino, R.J. (2014). Sex differences in stress-related psychiatric disorders: Neurobiological perspectives. *Frontiers in Neuroendocrinology, 35,* 303–319. doi.org/10.1016/j.yfrne.2014.03.008

Beattie, G. (2008). *Sex differences in driving and insurance risk: Understanding the neurobiological and evolutionary foundations of the differences.* Manchester, England: University of Manchester.

Beck, A. T., & Emery, G. (2005). *Anxiety disorders and phobias: A cognitive perspective.* New York: Basic Books.

Bernstein, D.A., & Allen, G.J. (1969). Fear Survey Schedule (II): Normative data and factor analyses based on a large college sample. *Behaviour Research and Therapy, 7,* 403-407.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Boyer, P. & Bergstrom, B. (2011). Threat-detection in child development: An evolutionary perspective. *Neuroscience and Biobehavioral Reviews, 35*(4), 1034-1041. doi:10.1016/j.neubiorev.2010.08.010

Bradley, M. M., & Lang, O. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). New York: Oxford University Press.

Brebner, J. (2003). Gender and emotions. *Personality and Individual Differences, 34,* 387-394. doi:10.1016/S0191-8869(02)00059-4

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125,* 367–383. doi:10.1037/0033-2909.125.3.367

Campbell., A. (2013). *A mind of her own: The evolutionary psychology of women* (Second Edition). Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199609543.001.0001

Corr, P. J., & Cooper, A. (in preparation). *The Corr-Cooper Reinforcement Sensitivity Theory Personality Questionnaire (RST-PQ): Development and validation.* Retrieved from http://www.philipcorr.net/uploads/downloads/153.pdf

Costa, P.T., Terracciano, A., & McCrae, R.R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*(2), 322-331. doi:10.1037/0022-3514.81.2.322

Craig, A.D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10,* 59-70. doi:10.1038/nrn2555

Croson, R. & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47* (2), 448-474. doi:10.1257/jel.47.2.448

Cross, C.P., Copping, L.T. & Campbell, A. (2011). Sex differences in impulsivity: A meta-analysis. *Psychological Bulletin, 137* (1), 97–130. doi:10.1037/a0021591

Davis, M., Walker, D.L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology, 35,* 105–135. doi:10.1038/npp.2009.109

Derryberry, D. & Rothbart, M.K. (1997). Reactive and effortful processes in the organisation of temperament. *Development and Psychopathology, 18* (4), 633-652.

Dillon, K. M., Wolf, E., & Katz, H. (1985). Sex roles, gender, and fear. *Journal of Psychology, 119,* 355–359. doi:10.1080/00223980.1985.10542907

Elliot, A.J., & Thrash, T.M. (2010). Approach and avoidance temperament as basic dimensions of personality. *Journal of Personality, 78,* 865-906. doi:10.1111/j.1467-6494.2010.00636.x

Epstein, S. (1972). The nature of anxiety with emphasis upon its relationship to expectancy. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research,* Vol. 2. (pp. 291−337). New York: Academic Press.

Fischer, A. H., Mosquera, P. M.R., van Vianen, A., & Manstead, A. S. R. (2004). Gender and culture differences in emotion. *Emotion, 4*, 87-94. doi:10.1037/1528-3542.4.1.87

Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996). Gender and age differences in the prevalence of specific fears and phobias. *Behaviour Research and Therapy 34,* 33–39. doi:10.1016/0005-7967(95)00048-3

Gabriel, K.I., & Williamson, A. (2010). Framing alters risk-taking behaviour on a modified balloon analogue risk task (BART) in a sex-specific manner. *Psychological Reports, 107,* 699-712. doi:10.2466/02.09.20.PR0.107.6.699-712

Geer, J. H. (1965). The development of a scale to measure fear. *Behaviour Research and Therapy, 3,* 45−53. doi:10.1016/0005-7967(65)90040-9

Gillis, M.M., Haaga, D.A.F., & Ford, G.T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire and Social Phobia and Anxiety Inventory. *Psychological Assessment, 7*(4), 450-455. doi:10.1037/1040-3590.7.4.450

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

Harris, C. R., Jenkins, M., & Glaser, D. (2006). Gender differences in risk assessment: Why do women take fewer risks than men? *Judgment and Decision Making Journal, 1,* 48-63.

Harris, M.B. & Miller, K.C. (2000). Gender and perceptions of danger. *Sex Roles* **43,** 843-863. doi:10.1023/A:1011036905770

Izard, C. E., & Ackerman, B. P. (2000). Motivational, organizational, and regulatory functions of discrete emotions. In M. Lewis, & J.M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 253−264). 2nd Ed. New York: Guilford Press.

Izard, C.E., Libero, D.Z., Putnam, P. & Haynes, O.M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology, 64* (5), 847-860. doi:10.1037//0022-3514.64.5.847

Jackson, C.J. (2009) Jackson-5 scales of revised Reinforcement Sensitivity Theory (r-RST) and their application to dysfunctional real world outcomes. *Journal of Research in Personality, 43,* 556–569. doi:10.1016/j.jrp.2009.02.007

Jackson, J.S.H. & Maraun, M. (1996). The conceptual validity of empirical scale construction: The case of the Sensation Seeking Scale. *Personality and Individual Differences, 21*(1), 103-110.

Jansz, J. (2000). Masculine identity and restrictive emotionality. In A. H. Fischer (Eds.), *Gender and Emotion: Social Psychological Perspectives,* (166-186). Cambridge: University Press.

Jose, A., Olino, T.M., & O'Leary, K.D. (2012). Item Response Theory analysis of intimate-partner violence in a community sample. *Journal of Family Psychology, 26* (2), 198–205. doi:10.1037/a0027100

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders, 61*, 137−159. doi:10.1016/S0165-0327(00)00343-8

Lick, J.R., Sushinsky, L.W., & Malow, R. (1977). Specificity of fear survey schedule items and the prediction of avoidance behaviour. *Behavior Modification, 1,* 195-203. doi:10.1177/014544557712003

Linacre, J.M. (1995). Sample size and item calibration stability. *Rasch Measurement Transactions, 7:4, p.328.*

Linacre, J.M. (2012). *A user's guide to Winsteps, Ministep Rasch-Model computer programmes.* http://www.winsteps.com.winman/index.htm

Loewenstein, G.F., Weber, E.U., Hsee, C.K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127* (2), 267-286. doi: 10.1037//0033-2909.127.2.267

Maack, D.J., Buchanan, E. & Young, J. (2015). Development and psychometric investigation of an inventory to assess fight, flight, and freeze tendencies: The Fight, Flight, Freeze Questionnaire. *Cognitive Behavior Therapy, 44* (2), 116-127. doi:10.1080/16506073.2014.972443

Marks, I.M., & Mathews, A.M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy, 17,* 263-267. doi:10.1016/0005-7967(79)90041-X

Mavranezouli, I., Brazier, J.E., Young, T.A., & Barkham, M. (2011). Using Rasch analysis to form plausible health states amenable to valuation: The development of CORE-6D from a measure of common health problems (CORE-OM). *Quality of Life Research, 20,* 321-333. doi:10.1007/s11136-010-9768-4

McLean, C.P., & Hope, D.A. (2010). Subjective anxiety and behavioural avoidance: Gender, gender role, and perceived confirmability of self-report. *Journal of Anxiety Disorders, 24,* 494-502. doi:10.1016/j.janxdis.2010.03.006

McNaughton, N., & Corr, P.J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews, 28,* 285–305. doi:10.1016/j.neubiorev.2004.03.005

Mokken, R. J. (1971). *A theory and procedure of scale analysis: With applications in political research* (Vol. 1). Berlin: Walter de Gruyter. doi:10.1515/9783110813203

Molenaar, I.W. (1982). Mokken scaling revisited. *Kwantitatieve Methoden, 3*(8)*,* 145–164.

Öhman, A. (2008). Fear and anxiety: Overlaps and dissociations. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 709−729)., 3rd Ed. New York, NY: Guilford Press.

Opre, A., & Albu, M. (2010). Proprietăţi psihometrice ale versiunii române a Chestionarului de personalitate Zuckerman-Kuhlman (ZKPQ): Date preliminare. Anuarul Institului de Istorie "George Bariţiu" din Cluj Napoca, Series Humanistica - VIII.

Park, S. Sohn, J.H., Hong, J.P., Chang, S.M., Lee, Y.M., Jeon, H.J., Cho S-J., Bae, J.N., Lee, J.Y., Son J-W., & Cho, M.J. (2013). Prevalence, correlates, and comorbidities of four DSM-IV specific phobia subtypes: Results from the Korean Epidemiological Catchment Area study. *Psychiatry Research, 209,* 596–603. doi: 10.1016/j.psychres.2012.12.025

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. doi:10.1037/1528-3542.7.2.252

Petersen, J.L. & Hyde, J.S. (2010). A meta analytic review of research on gender differences in sexuality, 1993-2007. *Psychological Bulletin, 136* (1), 21–38. doi:10.1037/a0017504

Pierce, K. A., & Kirkpatrick, D. R. (1992). Do men lie on fear surveys? *Behavior Research and Therapy*, *30*, 415–418. doi:10.1016/0005-7967(92)90055-L

Robinson, V. (2008). *Everyday masculinities and extreme sports: Male identity and rock climbing.* Oxford, England: Berg.

Schaefer, H.S., Larson, C.L., Davidson, R.J., & Coan, J.A. (2014). Brain, body, and cognition: Neural, physiological and self-report correlates of phobic and normative fear. *Biological Psychology, 98,* 56-69. doi:10.1016/j.biopsycho.2013.12.011

Sijtsma, K., Meijer, R. R., & van der Ark, L. A. (2011). Mokken scale analysis as time goes by: An update for scaling practitioners. *Personality and Individual Differences*, *50* (1), 31-37. doi:10.1016/j.paid.2010.08.016

Simon, R.W., & Nath, L.E. (2004). Gender and emotion in the United States: Do men and women differ in self-reports of feelings and expressive behavior? *American Journal of Sociology, 109,* 1137-1176. doi:10.1086/382111

Smith, E. V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, *10*, 205-231.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates and antecedents.* Austin: University of Texas Press.

Sylvers, P., Lilienfeld, J.L., & LaPrairie, J.L. (2011). Differences between trait fear and trait anxiety: Implications for psychopathology. *Clinical Psychology Review, 31,* 122–137. doi:10.1016/j.cpr.2010.08.004

Tellegen, A. (1982). *Brief manual for the Multidimensional Personality Questionnaire*. Unpublished manuscript, University of Minnesota, Minneapolis.

Tristan. A. (2006). An adjustment for sample size in DIF analysis. *Rasch Measurement Transactions*, *20*, 2-3.

Tucker, M., & Bond, N. W. (1997). The roles of gender, sex role, and disgust in fear of animals. *Personality and Individual Differences, 22,* 135–138. doi:10.1016/S0191-8869(96)00168-7

Twenge, J.M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles, 36* (5/6), 305-325. doi:10.1007/BF02766650

Twenge, J.M.(2009). Status and gender: The paradox of progress in an age of narcissism. *Sex Roles, 61* (5/6) 338–340. doi:10.1007/s11199-009-9617-5

van der Ark, L. A. (2007). Mokken scale analysis in R. *Journal of Statistical Software, 20*(11), 1–19.

van Schur, W. H. (2003). Mokken scale analysis: Between the Guttman scale and parametric item response theory. *Political Analysis, 11*(2)*,* 139–163. doi:10.1093/pan/mpg002

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54* (6), 1063-1070. doi:10.1037/0022-3514.54.6.1063

Wolpe, J. (1987). The promotion of scientific psychotherapy: A long voyage. In J. K. Zeig (Ed.), *The evolution of psychotherapy* (pp. 133−142). New York: Routledge.

Wolpe, J., & Lang, P. J. (1964). A fear survey schedule for use in behaviour therapy. *Behavior Research and Therapy, 2,* 27−30. doi:10.1016/0005-7967(64)90051-8

Wright, B.D., & Stone, M.H. (2004). *Making measures*. Chicago: The Phaneron Press.

Xu, Y., Schneiera, F., Heimberg, R.G., Princisvalle, K., Liebowitz, M.R., Wang, S. & Blanco, C. (2012). Gender differences in social anxiety disorder: Results from the national epidemiologic sample on alcohol and related conditions. *Journal of Anxiety Disorders, 26,* 12– 19. doi:10.1016/j.janxdis.2011.08.006

Zuckerman, M. (1994). *Behavioural expressions and biosocial bases of sensation seeking.* Cambridge: Cambridge University Press.

Zuckerman, M., & Kuhlman, D. M. (2000). Personality and risk-taking: Common biosocial factors. *Journal of Personality, 68,* 999–1029. doi:10.1111/1467-6494.00124

Zuckerman, M. (1996). “Conceptual clarification” or confusion in “The study of sensation seeking.” *Personality and Individual Differences, 21*(1), 111-114.

Zuckerman, M. & Kuhlman, D.M. (n.d.) *Norms for the Zuckerman Kuhlman Personality Questionnaire.* Department of Psychology, University of Delaware.

Table 1

Means (and standard deviations) by sex, Rasch item fit statistics and Mokken indices of differential item functioning

	Fear score		Rasch item indices			Mokken item indices	
Item	Men	Women	Infit Outfit (mean square)	Partial r	Tristan *t* Differential Item Function	Item H values (36 items)	Item H values (27 items)
1. You are walking down a poorly lit alleyway at night. You can hear footsteps behind you but you can't see anybody there	2.73 (0.96)	3.80 (0.92)	0.80 0.80	0.64	-0.37[a]	0.29	0.31
2. You are on a transatlantic flight and suddenly you feel the plane bumping about as it loses and gains altitude	1.98 (1.13)	2.78 (1.24)	1.06 1.07	0.5	-0.93 [a]	0.20	
3. You are just about to do a bungee jump off the side of a bridge; the initial drop is over 200m	3.20 (1.32)	3.72 (1.16)	1.08 1.07	0.48	0.45 [a]	0.30	0.31
4. You are on the world's largest roller coaster, it is juddering as it ascents to the first peak	2.50[a] (1.22)	2.74 (1.19)	1.12 1.12	0.44	2.15		
5. When you are in the middle of a presentation in front of some influential people. Your powerpoint stops working so you have to speak ad lib	2.22 (1.03)	2.60 (1.05)	1.03 1.05	0.47	1.85 [a]	0.18	

6. You are driving along the motor way at 70mph. Somebody cuts in front of you very suddenly and you have to break hard	2.51 (0.90)	3.05 (1.02)	0.96 0.96	0.51	-2.30		
7. You are alone in your house during the night when you hear what sounds like someone breaking in	3.28 (0.98)	4.11 (0.83)	0.87 0.87	0.57	-1.21 [a]	0.34	0.36
8. You are scuba diving and suddenly your oxygen tank starts to malfunction	3.78 (1.07)	4.22 (0.91)	1.01 0.97	0.47	-1.60 [a]	0.27	
9. You are in a large crowd at a music concert and have been separated from your friends. They are nowhere to be seen	1.21 (0.46)	2.06 (0.94)	0.90 0.82	0.55	-1.79 [a]	0.27	
10. You have overslept and missed an important meeting with your boss	3.03[a] (0.95)	3.09 (0.95)	1.17 1.18	0.31	3.01		
11. You are on the top floor of a building when a fire alarm goes off. You reach the staircase and see the fire blocking the stairs down to the ground floor	3.90 (0.97)	4.50 (0.77)	0.94 0.93	0.49	-0.49 [a]	0.35	0.39
12. You are swimming in the sea with a few friends. You can feel yourself being rapidly pulled away from the shore by a rip tide	3.23 (1.05)	3.71 (0.93)	1.01 1.03	0.45	-1.09 [a]	0.29	0.31
13. You lose your train ticket and	2.23	2.73	0.96	0.5	0.12 [a]	0.31	0.33

realise that the guard will be checking soon. You have no more cash to pay a fine or pay for a new one	(0.87)	(1.04)	0.96				
14. You are drunk at a party out of town and your friends have gone home without you	1.71 (0.82)	2.42 (1.06)	0.95 0.92	0.53	-0.54 [a]	0.36	0.39
15. On holiday, you leave a party late and realise you do not know how to get back to your holiday flat	2.03 (0.91)	3.02 (0.95)	0.79 0.79	0.65	-3.77		
16. You can feel a lump beneath your breast / prostate	3.18 (0.95)	3.47 (1.06)	1.15 1.14	0.36	-0.85 [a]	0.23	
17. You realise you have left the iron/gas on when you left for work this morning	2.94 (0.83)	3.27 (0.95)	1.08 1.08	0.37	-0.28 [a]	0.33	0.35
18. Sitting on top of the bus, a group of teenagers is in front of you are making threatening comments	2.30 (0.93)	2.64 (1.01)	1.02 1.00	0.46	1.04 [a]	0.32	0.35
19. You are the only person standing on the train platform late at night	1.37 (0.61)	2.56 (1.02)	0.89 0.87	0.59	-3.85		
20. You are looking after a child with asthma – she complains she feels breathless	2.50 (0.85)	2.76 (0.96)	1.00 1.01	0.45	4.30		
21. You have borrowed your friend's	2.19	2.58	1.06	0.4	3.24		

laptop without asking. She now claims it is not working properly	(0.87)	(0.92)	1.06				
22. You check the petrol gauge – your car is running at empty and there is no garage for miles	2.16 (0.92)	2.53 (0.88)	0.93 0.93	0.51	1.44 [a]	0.34	0.36
23. All the electrics fuse in your house, and you are alone and unsure of where the fuse box is	1.52 (0.87)	2.44 (1.12)	1.00 1.04	0.51	-0.62 [a]	0.35	0.37
24. You are involved in a car accident on a busy motorway. You have been taken to hospital and you cannot feel any sensation in your legs	4.42[a] (0.67)	4.57 (0.63)	1.14 1.33	0.19	1.22 [a]	0.34	0.36
25. You have just been told a close relative has had to have emergency surgery after a traffic accident	3.86 (0.92)	4.15 (0.85)	1.10 1.16	0.34	3.22		
26. Your flight has been cancelled due to bad weather conditions. You think you may be stranded abroad with little money	1.83 (0.84)	2.33 (0.89)	0.90 0.93	0.53	1.72 [a]	0.35	0.38
27. A member of your family is on holiday abroad, and you hear that there have been terrorist attacks in the city where they are staying. You cannot get hold of them	3.40 (0.90)	4.10 (0.87)	0.94 0.95	0.51	1.44 [a]	0.32	0.34
28. You hit an icy patch on a busy main road, and lose complete control of the car	3.64 (0.98)	4.07 (0.85)	0.90 0.90	0.53	-0.63 [a]	0.38	0.42

29. You stall your car on a busy roundabout. You cannot seem to start it again	2.04 (0.98)	2.60 (1.05)	0.89 0.89	0.57	1.57 [a]	0.33	0.36
30. You are crossing a field with a large bull in it	2.11 (0.81)	2.71 (1.05)	0.98 0.99	0.49	1.19 [a]	0.28	
31. Your neighbour has hurt herself very badly in the garden – she appears to be unconscious	2.61 (0.93)	3.06 (0.99)	1.01 1.00	0.47	-1.47 [a]	0.29	0.30
32. A member of your family should have been home many hours ago – you cannot contact them	2.70 (1.05)	3.32 (1.03)	0.96 0.97	0.52	-1.92 [a]	0.29	0.32
33. You are told your sexual partner is HIV positive	4.47[a] (0.66)	4.58 0.69)	1.17 1.39	0.2	0.22 [a]	0.22	
34. You have just bet a large amount of money in a casino	2.71[a] (1.03)	2.86 (1.02)	1.18 1.18	0.34	1.47 [a]	0.25	
35. You are called to one side as you go through 'Customs'	1.54 (0.74)	1.92 (0.89)	0.96 0.97	0.46	-0.72 [a]	0.31	0.33
36. You are asked to provide a breathalyser test – you think you should be below the limit, but you have had a bit to drink	2.70 (0.87)	3.01 (0.93)	1.05 1.05	0.4	0.53 [a]	0.35	0.38
37. You are abroad and the taxi driver is driving over 70mph. There are no	2.33 (0.92)	2.96 (0.89)	0.90 0.89	0.54	1.50 [a]	0.33	0.36

safety belts in the vehicle							
38. You go for a walk on the cliffs with friends – there is no barrier and the path is very narrow with a steep drop to the sea	2.27[a] (1.15)	2.47 (1.08)	1.04 1.09	0.46	2.66		
39. You are caught outside in a heavy thunder storm with lightening very close by	1.53 (0.69)	2.19 (0.99)	0.94 0.92	0.52	-0.83 [a]	0.34	0.36
40. You are told you risk permanent blindness as your retina is detaching, you must get to hospital quickly	4.18[a] (0.82)	4.33 (0.85)	1.10 1.12	0.34	0.70 [a]	0.33	0.35
41. Your lift gets stuck between floors – you are alone	2.14 (1.01)	3.21 (1.19)	0.99 0.99	0.55	-3.42		
42. The police are moving people out of the airport area you are in because of a suspect package	1.89 (0.81)	2.57 (1.06)	0.91 0.90	0.55	-1.31 [a]	0.34	0.37
43. You have received several anonymous calls to your mobile, and this time the person is making threatening comments	2.41 (0.97)	3.26 (0.98)	0.90 0.89	0.57	-2.58		
44. Your computer is stolen and you have not backed up your work	3.18 (1.17)	3.52 (1.10)	1.12 1.10	0.43	-0.22 [a]	0.26	
45. You have just been asked at short notice to give a 'vote of thanks' speech at a friend's wedding	2.20[a] (1.14)	2.33 (1.08)	1.28 1.35	0.3	2.80		

46. You are white water-rafting in Australia, and the instructor is finding it difficult to keep control of the raft	2.52 (0.81)	3.17 (0.91)	0.84 0.84	0.59	1.07 [a]	0.39	0.42
47. You are on your gap year travels and have to cross a make-shift jungle bridge high above a chasm	2.91 (1.07)	3.23 (1.17)	1.02 1.03	0.5	1.92 [a]	0.36	0.39
48. You lie down on your bed and see a large house spider on your pillow	1.81 (1.15)	3.07 (1.36)	1.23 1.36	0.45	-2.86		
49. Whilst you are swimming off the coast of South Africa; the shark warning is sounded	3.51 (0.91)	3.85 (1.01)	1.01 1.01	0.46	-2.95		
50. You are in the middle of taking an important exam. You have just read the questions and realise that you have not revised the relevant topics	3.47 (1.12)	3.93 (0.96)	1.02 1.03	0.47	-1.20 [a]	0.32	0.33
Final 27-item scale	2.70 (0.42)	3.23 (0.44)				Scale H=.31	Scale H=.36

Items in **bold** were those retained on the 36-itemscale.
Items **bold and underlined** were retained on the 27-item scale

Table 2

Convergent and divergent validity of SFQ with personality and temperament measures

	FSS	PAQ Masculinity	PAQ Femininity	ATQ Approach	ATQ avoidance	ZKPQ Impulsivity	ZKPQ Sensation Seeking	ZKPQ Neuroticism-Anxiety	MPQ Stress Reaction	MPQ Harm Avoidance
SFQ	.69*** Sample 3 (N=245)	-.20** Sample 3 (N=245)	.20** Sample 3 (N=245)	.09 Sample 2 (N=309)	.33*** Sample 2 (N=309)	-.09 Sample 2 (N=305)	-.15** Sample 2 (N=306)	.31*** Sample 2 (N=306)	.29*** Sample 3 (N=245)	.39*** Sample 3 (N=245) .32*** Sample 2 (N=297)

** p<.01

*** p<.001

Table 3

Sex differences in mean Situated Fear Questionnaire scores in three samples

Sample	Male	Female	*t*	*d*
1 (UK)	2.70 (0.42)	3.23 (0.45)	9.75 (df=299)	-1.22
2 (Romanian)	2.38 (0.59)	3.08 (0.53)	9.83 (df=308)	-1.23
3 (UK)	2.66 (0.57)	3.22 (0.50)	8.16 (df=243)	-1.04