Personality and Individual Differences 97 (2016) 210–219

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

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

Anne Campbell [a,*], Claire Coombes [b], Raluca David [c], Adrian Opre [d], Lois Grayson [a], Steven Muncer [e]

[a] Psychology Department, Durham University, United Kingdom
[b] Faculty of Engineering Science, University College London, United Kingdom
[c] Department of Experimental Psychology, Oxford University, United Kingdom
[d] Department of Psychology, Babes-Bolyai University, Romania
[e] School of Health and Social Care, Teesside University, United Kingdom

ARTICLE INFO

Article history:
Received 14 September 2015
Received in revised form 14 March 2016
Accepted 18 March 2016
Available online 31 March 2016

Keywords:
Fear
Anxiety
Rasch
Mokken
Sex difference
Sex invariance
Fear survey schedule
Avoidance

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.

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, 2013; 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

* Corresponding author at: Psychology Department, Durham University, South Road, Durham DH1 3LE, England, United Kingdom
E-mail address: a.c.campbell@durham.ac.uk (A. Campbell).

http://dx.doi.org/10.1016/j.paid.2016.03.049
0191-8869/

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, 1984; Arrindell, Eisemann, Richter, Oei, Caballo & van der Ende, 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