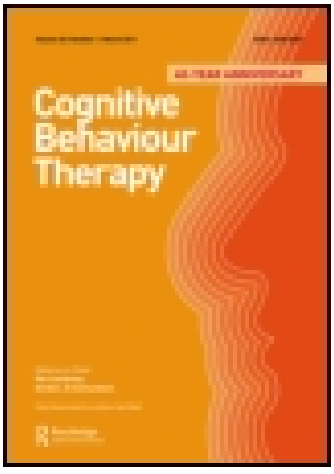

Cognitive Behaviour Therapy
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/sbeh20

Development and Psychometric Investigation of an Inventory to Assess Fight, Flight, and Freeze Tendencies: The Fight, Flight, Freeze Questionnaire

Danielle J. Maack[ab], Erin Buchanan[c] & John Young[ab]

[a] Department of Psychology, University of Mississippi, 205 Peabody Building, University, MS 38677, USA

[b] Delta Autumn Consulting, LLC, 101 Ricky D. Britt Sr. Blvd., Oxford, MS 38655, USA

[c] Department of Psychology, Missouri State University, Hill Hall 214D, 901 S National Ave, Springfield, MO 65897, USA

Published online: 03 Nov 2014.

To cite this article: Danielle J. Maack, Erin Buchanan & John Young (2014): Development and Psychometric Investigation of an Inventory to Assess Fight, Flight, and Freeze Tendencies: The Fight, Flight, Freeze Questionnaire, Cognitive Behaviour Therapy, DOI: 10.1080/16506073.2014.972443

To link to this article: http://dx.doi.org/10.1080/16506073.2014.972443

Cognitive Behaviour Therapy, 2014
http://dx.doi.org/10.1080/16506073.2014.972443

Routledge
Taylor & Francis Group

Development and Psychometric Investigation of an Inventory to Assess Fight, Flight, and Freeze Tendencies: The Fight, Flight, Freeze Questionnaire

Danielle J. Maack[1,2], Erin Buchanan[3] and John Young[1,2]

[1]*Department of Psychology, University of Mississippi, 205 Peabody Building, University, MS 38677, USA;* [2]*Delta Autumn Consulting, LLC, 101 Ricky D. Britt Sr. Blvd., Oxford, MS 38655, USA;* [3]*Department of Psychology, Missouri State University, Hill Hall 214D, 901 S National Ave, Springfield, MO 65897, USA*

Abstract. Fear is a psychological construct inherent in assessment of and reaction to threat. Its expression has been associated with individual differences in temperament, personality, and behavioral inhibition. Defining and subsequently assessing these individual differences in fear as a trait-like variable, however, have been largely neglected by researchers. Although there are well-established measures of fear, these primarily assess response to phobic stimuli rather than a reaction tendency to acute fear. As such, the goals of the present studies were to create, pilot, and revise a scale to assess the general construct of trait-like response to fear as it relates to underlying individual differences. Following guidelines for scale development, outlined by Haynes, Richard, and Kubany (1995 [Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238–247]) results of the current investigation provide strong, initial support for the factor structure, reliability, and construct validity of a new measure of trait-like fear: the Fight, Flight, Freeze questionnaire. *Key words: scale development; anxiety; fear; fight; flight; freeze.*

Received 10 July 2014; Accepted 29 September 2014

Correspondence address: Danielle J. Maack, PhD, Department of Psychology, University of Mississippi, 205 Peabody Building, University, MS 38677, USA. Tel: +1 662 915 1775. Fax: +1 662 915 5398. Email: djmaack@olemiss.edu, jnyoung1@olemiss.edu

Fear is a basic emotion which functions to promote avoidance goals by assisting in escape from threats (flight and freeze) or defensive approach (fight) in situations where avoidance is not an option. Fear arises from the threat of harm and reflects the anticipation of something aversive (see Blanchard, Hynd, Minke, Minemoto, & Blanchard, 2001 for review of defensive behaviors). From an evolutionary perspective, animals developed a defense system in order to cope with threats to longevity and successful procreation. Depending on the situation, escape, avoidance, or attack may be required to successfully respond to danger. The emotion associated with activation of any of these motivational states is referred to as fear (Blanchard et al., 2001; Ohman & Mineka, 2001).

Contemporary theories of emotion consider motivational states, such as fear, to be organized around two basic survival systems: one defensive and one appetitive in nature. More relevant to a trait conceptualization of fear, Gray and McNaughton's (2000) revised Reinforcement Sensitivity Theory proposes that fear is the underlying emotion of the fight, flight, freeze system which is part of the body's defensive motivational system. Fear manifests as flight (if escape is available), freeze (which constitutes a more passive form of avoidance, as elicited by threats that need not be approached), or fight (if escape is not an option and defensive approach is required; Corr, 2008; Heym, Ferguson, & Lawrence, 2008).

Currently, psychopathologies related to fear are taxonomically conceptualized as anxiety disorders, which comprise both cognition and somatic sequelae of anxiety as primary symptoms (e.g., worry and physical tension in social anxiety disorder, hyperventilating,

racing heart, and prospective fear of experiencing another attack in panic disorder). However, researchers have proposed that disorders be distinguished with separate fear-based and anxiety-based factors (Krueger, 1999; Watson, 2005). Specifically, using data from the National Comorbidity Study, Krueger (1999) assessed the structure of 10 common psychological disorders. The best-fit statistical model categorizing these disorders included superordinate factors of internalizing and externalizing. Subfactors of internalizing disorders included two latent constructs: anxiety/misery and fear, suggesting that the distinction between anxiety and fear is not only of theoretical interest but also important to clinical classification.

Similarly, recent studies and reviews (e.g., Cooper, Perkins, & Corr, 2007; Heym et al., 2008; Perkins, Kemp, & Corr, 2007; see Sylvers, Lilienfeld, & LaPrairie, 2011 for a comprehensive review) have demonstrated both conceptual and measurement differences between fear and anxiety. These differences point toward an important distinction thought to underlie the directional motivation of behavior. The NIMH Research Domain Criteria (RDoC) Project, which emphasizes increased understanding of the differentiations in negative valence systems, calls for research in this area, particularly "responses to acute threat (fear)" (NIMH RDoC Team, 2011, p. 2). Fear was further defined as "an activation of the brain's defensive motivational system to promote behaviors that protect the organism from perceived danger." Instrumentation to support this initiative is lacking, however, in that existing self-reports relevant to fear are not consonant with this definition.

Contemporary measurements of fear tend to either focus on one aspect of the fear construct (namely phobic fear) and/or have a conceptualization of fear that does not differentially address the fight, flight, and freeze systems. The two most commonly used assessments of fear are the Fear Survey Schedule-III (Wolpe & Lang, 1964) and the Fear Questionnaire (Marks & Mathews, 1979). Both measures are self-report assessments related to *phobic* reactions in response to specific stimuli, and as such do not capture the underlying tendency to react to acute fear. 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. The emphasis of existing measures of fear on phobic reactions rather than a propensity to experience fear more generally has been criticized as possibly being confounded as a secondary assessment of neuroticism and trait anxiety related specifically to phobias (Corr, 2008; Torrubia, Avilas, & Caseras, 2008).

Two other measures have also been used to assess trait fear, the Activities Preference Questionnaire (APQ; Lykken, Tellegen, & Katzenmeyer, 1973) and the Multidimensional Personality Questionnaire-Harm Avoidance subscale (MPQ-HA; Tellegen, 1982). The APQ was initially presented as a measure of anxiety reactivity (Blankstein, 1975) and forces respondents to choose between two activities judged to be of equal unpleasantness in different ways. One choice is considered frightening or embarrassing (referring to either actual physical danger or squeamishness or to social fears or embarrassment) and one is considered tedious or distasteful. As the respondent chooses preference of activities, there is no indication that the response chosen is due to avoidance of fear or reflective of general fear responses (i.e., fight, flight, or freeze).

Similar to the APQ, the MPQ-HA has also been used in a variety of studies to assess fearfulness (Sylvers et al., 2011). The MPQ-HA scale was based on the forced choice format used by the APQ with the addition of items asking how likely an individual would be to fear or avoid a potentially frightening stimulus. Although it has demonstrated adequate internal consistency (Tellegan & Waller, 2007), the measure is primarily focused on an individual's preference to avoid potentially scary/harmful situations. It does not assess any specific emotional reactions or tendencies toward general fight, flight, and freeze reactions to fear-inducing experiences.

Other ancillary attempts to assess the construct of a more trait-like fear (i.e., fight/flight/freeze) have been incomplete and the lack of an instrument to assess this construct has been pointed out as a gap in the literature (Cooper et al., 2007; Perkins & Corr, 2006; Sylvers et al., 2011; Torrubia et al., 2008). Thus, the current paper presents an effort to develop a comprehensive assessment of fear that is aligned with the RDoC initiative as well as contemporary evolutionary theory regarding both escape from threats and defensive approach.

General approach to measure development

Overall study

The Fight Flight Freeze Questionnaire (FFFQ) was thus developed through a series of psychometric studies to assess the construct of temperamental fear in accord with theory. Items were initially developed and selected to reflect fear reactivity in specific domains stereotypic of fight, flight, and freeze responses. Subsequent studies were conducted to explore the factor structure, reliability, and convergent/divergent validity of the instrument, thereby aiding in the development of a cogent, theoretically-driven, comprehensive, empirically-supported assessment of the different range of trait-like fight, flight, and freeze fear responses to threat.

Overall materials and method

The overall methods of this series of studies closely followed a published test construction guide (Haynes et al., 1995). The first of these steps was identification of a target construct (trait-like response to fear) and its subdivisions (fight, flight, freeze). What follows are general procedures for several iterations, examinations, and refinements of a measure designed to capture this construct. All participants described in a given phase of study participated in *only* that phase of study (i.e., each sample comprised unique individuals). All study procedures were approved through the University's IRB board and all participants provided informed consent prior to participation.

Study 1: Vignette construction and item generation

Method for vignette construction

To develop an initial pool of representative items from which to construct the measure, the authors created nine vignettes (three per proposed factor) describing individuals in situations where a specific element of fight, flight, or freeze was exemplified. These vignettes were based on theory as well as previous work in this area (e.g., Blanchard et al., 2001; Perkins & Corr, 2006). Each was subjected to lab review and revision prior to being implemented in an undergraduate sample. Subsequent to refinement all nine vignettes were given to a group of university students with instructions to provide a list of adjectives, words, or phrases that came to mind in describing the individual depicted in the vignette.[1]

Participants and development of scale items

A total of 74 undergraduates participated in this phase of study (age $M = 20.39$ years; $SD = 5.18$; 72% females; 62.16% White; 27.02% Black; 6.76% Asian; 4.05% Multiracial). All participants' responses to vignettes were entered into a database that was subsequently examined in terms of frequency. Words appearing frequently in response to a theoretically relevant stimulus (example "frozen" in response to the "freeze" vignette given above) were retained and formed the basis of item generation for this study. The authors also generated additional relevant items after reviewing participants' most frequent responses. This process produced 64 words subjected to expert review (i.e., PhD psychologists with related research areas outside of authors' universities) concerning their relevance, representativeness, specificity, and clarity (Haynes et al., 1995). This procedure resulted in nine items being removed for being redundant with other items or lack of fit via the aforementioned criteria. The resultant pool of 55 items comprised the first iteration of the FFFQ. The format of the FFFQ employs a Likert-type rating scale. Instructions on the scale were as follows: "Please read each word from the list and indicate (1–5 with 5 being almost always and 1 being almost never) how you typically react to potentially threatening situations. Do not spend too much time thinking about each word."

Study 2: Version 1 of the FFFQ

Participants

Three hundred and sixty-five participants from a large Southern University completed the 55-item initial version of the FFFQ. Cases with less than 10% missing data were retained for analyses, with missing data imputed via series mean with SPSS. Overall $< 3\%$ of the data were imputed. Five participants were excluded due to having greater than 10% missing data, and another 39 were omitted as

multivariate outliers (assessed via Mahalanobis distance; Tabachnick & Fidell, 2012). The remaining 321 participants were included in an exploratory factory analysis designed to test the factor structure of the FFFQ. The demographic characteristics of this sample were as follows: age $M = 21.92$ years ($SD = 4.84$, range 17–59); 57.0% Female; 81.9% Caucasian; 15.3% African American; 1.5% Asian; 1.2% Multiracial.

Analytic procedure

All analyses were performed using the FACTOR program (Lorenzo-Seva & Ferrando, 2006) and evaluated via Preacher and MacCallum's (2003) guidelines for exploratory factor analysis (EFA). This program was selected because it provides several fit indices for EFA including the comparative fit index (CFI; Bentler, 1990), Tucker-Lewis non-normed fit index (NNFI; Bentler & Bonnett, 1980), standardized root mean residual (SRMR; Jöreskog & Sörbom, 1981), and the root mean square error of approximation (RMSEA; Steiger, 1990). The CFI and NNFI indices are considered suitable at .90 and good at .95 (Bryant & Yarnold, 1995). The RMSEA and SRMR are considered to be suitable at values lower than .10 and good below .06 (Thompson, 2004). In all analyses, only items with loadings over .30 were considered to adequately load on a factor. Items with loadings over .30 on multiple factors were considered to "split" (Preacher & MacCallum, 2003), which would thus lead to exclusion.

Results

A parallel analysis and scree plot examination were conducted on the data and analyses indicated that a four-factor solution was warranted. Therefore, an EFA with four factors was analyzed with direct oblimin rotation (an oblique rotation that allows correlated factors) and maximum likelihood estimation. This analysis indicated 17 questions split factors and 1 question did not load on any of the four factors. Given the lack of fit, and consistent with expectations for this stage of the instrument development process (Haynes et al., 1995), these items were eliminated from further analyses (see Supplemental Table 1 for fit indices for each analysis step).

A second parallel analysis and examination of a new scree plot indicated that three factors would fit the data; thus, a three-factor EFA model was analyzed with the remaining 39 questions. This model indicated two questions did not load on any of the factors, and five questions split factors. Removal of these items with poor properties resulted in a further reduced item pool again subjected to parallel analysis and scree plot examination. Results suggested a two-factor fit. The EFA conducted within this two-factor model elucidated two items that did not load on a factor and one that split factors. The final two-factor model had 27 questions with fairly good fit indices, no split factors, and all items loading on a factor (see Supplemental Table 2).

Eleven items loaded on Factor 1 (surreal, passive, faltering, numb, dazed, detached, disengaged, blank, vacant, disconnect, empty), which, consistent with expert review of items and vignette responses, seemed to reflect a "Freeze" factor. Factor 2 consisted of the remaining 16 items (scared, fearful, frightened, afraid, threatened, terrified, distressed, startled, overwhelmed, flustered, tense, over-reactive, shaky, petrified, trapped, breathless), which indicated a separate "Flight" factor (again consistent with expert review and vignette responses). Despite adequate loadings and a cohesive factor structure, the two subscales produced during this phase of study were not consistent with theoretical understanding. As such, researchers began anew in constructing items to comprise a "Fight" scale. In effect, phase one culminated in statistical support for 11 "Freeze" items, 16 "Flight" items, and the realization that reformulation was necessary to adequately assess "Fight." Therefore, researchers constructed 13 distinct items rationally determined to more appropriately assess this construct. The combination of these items with the 27 Freeze and Flight items mentioned above comprised a second version of the instrument (40 total questions).

Study 3: Version 2 of the FFFQ

Method

Participants. A total of 244 new participants completed the second, 40-item version of the FFFQ. Data were again screened for multivariate outliers via Mahalanobis distance, and

15 participants were excluded from further analyses. Demographics for the remaining 229 participants were as follows: 55.5% Female; 66.8% Caucasian, 24.9% African American; 2.6 Asian; 1.3% Hispanic; 3.5% Multiracial; .8% Pacific Islander or Native American. Mean age was 20.00 years (SD = 3.00; range 18–45).

Results

The same EFA procedures and recommendations for evaluation used above were implemented to analyze the second version of the FFFQ. Parallel analysis and scree plot examination suggested a four-factor solution, which was used to inform the EFA that followed. All models were analyzed using direct oblimin and maximum likelihood estimation. The results of this analysis yielded four questions that split factors, which were therefore eliminated from the scale. A second parallel analysis and scree plot examination indicated a three-factor fit for the data produced by the remaining 36 items. The factor loadings on the second EFA indicated one question split factors, which was removed. A final three-factor model EFA was performed on the remaining 35 items, and demonstrated adequate loadings for each item, no split factor loadings for any items and adequate estimations of fit via conventional mechanisms (Preacher & MacCallum, 2003; see Supplemental Table 1 for fit indices and Supplemental Table 3 for individual item loadings).

This three-factor model cleanly separated distinct "Fight" and "Freeze" factors. The "Flight" factor, however, was somewhat confounded as it contained not only items that were determined by expert review and vignette responses to be "Flight" words but also five items theorized to load on the "Fight" subscale. Despite adequate loadings, these theoretically inconsistent items were eliminated (again through consensus of authors and expert reviewers). Additionally, subscales were balanced to retain the seven highest-loading and most theoretically consistent items. A final EFA was conducted on this refined, 21-item version of the FFFQ (loadings and questionnaire provided in supplemental materials). The model retained good fit with high NNFI and CFI values (>.90) and acceptable RMSEA (.08) and SRMR (.04) values. These items were retained for a final examination in an independent sample, as is consistent with recommendations for psychometric investigations (Nunnally & Bernstein, 1994).

Study 4: Version 3 of the FFFQ

Method

Participants. Another 225 unique participants were recruited. Five participants were excluded as multivariate outliers via Mahalanobis distance. Demographics of the remaining 220 participants were as follows: 58.2% Female; 76.8% Caucasian; 17.7% African American; 2.7% Asian; 2.5% Multiracial; .5% Pacific Islander or Native American. Mean age was 19.30 years (SD = 2.04; range 18–43). There were several additions to procedures for this phase of study. First, tests of internal consistency were conducted. Second, in order to test convergent and divergent validity of the FFFQ, all participants also completed the Behavioral Inhibition and Behavioral Activation System Scales (described below; Carver & White, 1994). Finally, a subsample ($n = 100$) completed the FFFQ 2 weeks later to assess test–retest reliability.

Additional measure. Behavioral Inhibition System and Behavioral Activation System Scales (BIS/BAS; Carver & White, 1994). The BIS/BAS is a widely used, 24-item self-report measure of behavioral inhibition and activation system sensitivity. Participants rate items on a four-point Likert-type scale (1 = *strongly agree*; 4 = *strongly disagree*). The BIS/BAS have demonstrated good reliability and convergent/discriminant validity (Carver & White, 1994). Scores are obtained for four subscales, including one for BIS sensitivity and three for BAS sensitivity (i.e., Reward-Responsiveness, Drive, and Fun-Seeking). With regard to the construct validity of BIS in particular, scores have been highly associated with neuroticism and negative affect (Jorm et al., 1999). Internal consistency in the present sample was adequate for both BIS ($\alpha = .74$) and BAS ($\alpha = .85$) subscales.

Results

EFA results. Parallel analysis and scree plot examination of data collected on the FFFQ again indicated a three-factor solution. The 21-item model of the FFFQ was analyzed via EFA with maximum likelihood, and direct

oblimin rotation. The final model exhibited a good fit with low RMSEA (.07) and SRMR (.04) values, high NNFI and CFI values (>.93), and robust, large factor loadings. The items loaded evenly on Fight, Flight, and Freeze factors, consistent with the EFA conducted on Version 2 and theoretical expectations for item categories (see Supplemental Tables 1 and 4 for fit statistics and individual item loadings, respectively).

Internal consistency and test–retest reliability. Reliability estimations indicated the FFFQ was sufficiently reliable ($\alpha = .90$). Likewise, individual subscales demonstrated similarly high reliability: Fight ($\alpha = .87$); Flight ($\alpha = .96$); Freeze ($\alpha = .85$). Additionally, a subsample ($n = 100$) was retested 2 weeks later. Demographics of this group were as follows: 48% Female; 82% Caucasian; 14% African American; 2% Asian; 1% Hispanic; 1% Multiracial. Their mean age was 19.46 years (SD = 2.71; range 18–43). The test–retest value of the FFFQ was high ($\alpha = .84$), and subscales displayed similarly high test–retest values: Fight ($\alpha = .80$); Flight ($\alpha = .89$); Freeze ($\alpha = .79$).

Initial convergent/divergent validity. To assess convergent and divergent validity of the FFFQ, overall and subscale scores were compared with the BIS/BAS scales. Based on published theory (cf., Gray & McNaughton, 2000), we hypothesized the following: (1) FFFQ overall score and subscales would demonstrate no relation with BAS (or subscales); (2) FFFQ overall score would exhibit a small, positive overall correlation with BIS; (3) FFFQ subscale scores of Flight and Freeze would be positively associated with BIS; and (4) FFFQ Fight subscale would be negatively associated with BIS (as individuals with higher BIS would theoretically be less inclined to fight rather than freeze or flee). Table 1 displays correlations between the FFFQ and the BIS/BAS scales.

%As hypothesized, the FFFQ total score and the Flight and Freeze subscales were significantly, positively associated with BIS. In addition, the expected significant, negative relation was seen between the Fight subscale and BIS scores. Neither FFFQ total score nor subscale scores were associated with BAS total or subscale scores with one exception. The Flight subscale of the FFFQ was negatively associated with the BAS-drive subscale. Although unexpected, a theoretical conceptualization of the drive construct provides an interpretation for this finding, as persistent pursuit of goals (drive) may be at odds with fleeing from a situation. Additionally, the drive subscale has previously been shown to be significantly, negatively associated with harm avoidance (Carver & White, 1994). As such, the negative relation between the drive subscale and Flight (arguably a mechanism of harm avoidance) seemed reasonable.

Study 5: Convergent/divergent validity and confirmation of factor structure

Method

Participants. Another group of participants ($n = 235$) was similarly recruited. Demographics were as follows: 62.1% Female; 76.8% Caucasian; 20.2% African American; 2.6% Asian; 1.7% Multiracial; 0.8% Pacific Islander or Native American. Mean age was 19.52 years (SD = 3.89; range 17–60). Again,

Table 1. *Means, standard deviations, and intercorrelations between FFFQ and BIS/BAS scales*

	Mean	SD	Fight	Flight	Freeze	BIS	BAS	RR	Drive	Fun
FFFQ total	45.31	13.28	.76***	.84***	.54***	.18**	−.02	.05	−.01	.02
Fight	17.02	6.20		.44***	.13	−.18**	.13	.08	.12	.12
Flight	15.89	7.27			.25***	.31***	−.12	−.03	−.15*	−.07
Freeze	12.31	4.48				.26***	.09	.12	.06	.01
BIS	20.32	3.71					.08	.27***	−.01	−.08
BAS	41.64	5.62						.77***	.86***	.82***
RR	17.71	2.07							.53***	.41***
Drive	11.42	2.42								.56***
Fun	12.50	2.37								–

Note. *p < *.05*, **p < *.01*, ***p < *.001*.

tests of internal consistency were conducted on the FFFQ (and subscales). Additionally, to continue testing the convergent and divergent validity of the FFFQ, all participants completed the State Trait Anxiety Inventory-trait scale, the Positive and Negative Affective Scales, and the NEO-Five Factor Inventory. A confirmatory factor analysis was also conducted on this sample as a final, stringent test of the stability of its factor structure.

Additional measures. *State Trait Anxiety Inventory-Trait Version* (STAI-T; Spielberger, Gorsuch, Lushene, Vaag, & Jacobs, 1983). The STAI-T is a 20-item scale assessing the propensity toward and stability of anxiety experiences, as well as the tendency to perceive stressful situations as threatening. The STAI-T is designed to capture individuals' typical, stable experiences along these dimensions. The widely used STAI-T has demonstrated high test–retest reliability, internal consistency, and concurrent validity with other anxiety questionnaires (e.g., Spielberger et al., 1983). Alpha internal consistency in this study was .88.

The positive and negative affect scales (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS consists of two, 10-item mood scales developed to provide brief measurement of positive and negative affect. Respondents are asked to rate the extent to which they have experienced each particular emotion "right now" or within the past week. Participants respond on a Likert-type scale from 1 = *very slightly or not at all* to 5 = *very much*. Alpha internal consistencies for the positive and negative scales were .83 and .88, respectively.

The NEO-five factor inventory (NEO-FFI; Costa & McCrae, 1992). The NEO-FFI is a 60-item scale assessing the domains of the five-factor model of personality: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Items are rated on a 5-point scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. The NEO-FFI has demonstrated good internal consistency and test–retest reliability, and has been validated against other personality inventories (Costa & McCrae, 1992). In this sample, internal consistency was as follows: Neuroticism = .81, Extraversion = .62, Openness to Experience = .68, Agreeableness = .72, and Conscientiousness = .82. Findings are relatively consistent with established norms of $\alpha = .68$ to .86 (Costa & McCrae, 1992).

Results

FFFQ internal consistency. Reliability estimations in this sample again indicated the FFFQ measure as a whole was sufficiently reliable ($\alpha = .92$). Likewise, individual subscales demonstrated similarly high reliability: Fight ($\alpha = .91$); Flight ($\alpha = .94$); Freeze ($\alpha = .86$).

Convergent/divergent validity. To assess the relation of the FFFQ compared with the other personality and anxiety instruments administered, we examined the Total Score, Fight, Flight, and Freeze factors in terms of their correlations with the STAI-T, PANAS, and NEO-FFI. Means and intercorrelations among variables of interest are provided in Table 1. Several notable, theoretically consistent results emerged when comparing scores on the FFFQ to the STAI-T and PANAS (constructs typically associated with emotional disorders). For example, moderate correlations were found between all subscales of the FFFQ and the STAI-T, with Pearson's *r*s ranging from .15 (Fight) to .46 (Freeze; all *p*s $< .05$). Similar results were evident for the correlation between the FFFQ and Negative Affect as measured by the PANAS (*r*s ranging from .27 (Fight) to .50 (Flight); all *p*s $< .05$). The only association between FFFQ and Positive Affect was a negative association with the Freeze scale ($r = -.20$).

Associations between the FFFQ and typical dimensions of personality (i.e., NEO-FFI) also produced several notable results. The overall FFFQ scale demonstrated a moderate relationship with Neuroticism ($r = .40$, $p < .001$), as did the Flight and Freeze subscales. Consistent with the theoretical function of a "fight" mechanism (i.e., defensive approach), the Fight subscale was not associated with Neuroticism, Conscientiousness, or Agreeableness. Flight scores, however, were specifically, differentially related to Neuroticism while higher Freeze scores directly related to neuroticism and inversely related to Conscientiousness, Agreeableness, and Extraversion (see Table 2).

Results of these analyses converge to suggest that FFFQ scores are associated with the tendency to experience negative reactions across domains. Additionally, although this

Table 2. *Means, standard deviations, and intercorrelations between FFFQ, STAI-T, PANAS, and NEO-FFI*

	Mean	SD	Fight	Flight	Freeze	STAI-T	PA	NA	O	C	E	A	N
FFFQ	39.12	13.43	.76***	.84***	.54***	.44**	− .04	.50**	.09	− .20**	.02	− .30**	.40**
Fight	14.14	6.35		.44***	.13	.15*	.04	.27**	− .01	− .14*	.09	− .38**	.10
Flight	12.72	6.22			.25***	.42**	.03	.48**	.07	− .08	.09	− .02	.39**
Freeze	12.65	4.71				.46**	− .20**	.42**	.13	− .24**	− .16*	− .27**	.45**
STAI-T	40.00	9.50					− .28**	.72**	.04	− .23**	− .14	− .26**	.68**
PA	28.05	6.07						− .05	.13	.42**	.33**	.23**	− .17*
NA	10.02	7.19							.00	− .16*	− .02	− .16*	.55**
O	23.79	6.00								− .07	− .01	− .06	.02
C	31.75	6.59									− .08	.28**	− .17*
E	29.61	5.86										.18**	− .18**
A	30.47	5.94											− .28**
N	21.98	8.03											

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

association is meaningful and provides initial evidence for the validity of the FFFQ as an instrument, it is not so high as to demonstrate completely overlapping constructs of measurement. In short, the pattern of results indicated the FFFQ assesses a different construct from the other instruments in this study in a way that is theoretically cogent.

Confirmatory factor analysis (CFA). As a final measure of FFFQ consistency, the three-factor, 21-question model derived from exploratory studies outlined above was investigated with CFA using maximum likelihood estimation. Four missing data points were found to be randomly distributed across participants; thus, mean replacement was used to impute these data in the CFA that follows. Data were examined for multivariate assumptions and outliers. All assumptions were met, and 17 outliers were found using Mahalanobis distance values and excluded from model tests.

Data analysis

Confirmatory factor analysis was conducted using AMOS 18.0, and results were evaluated via common metrics for CFA interpretation. Specifically, the χ^2/df ratio (Bryant & Yarnold, 1995), RMSEA, SRMR, CFI, and the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973) were all used as fit indicators. Excellent fit includes χ^2/df values below 3, RMSEA/SRMR values below .05 (good fit $< .10$), and CFI and TLI values above .90.

Results

The CFA analysis supported the factor structure suggested by exploratory analyses with low χ^2/df (2.519), RMSEA (.084; 90%CI: .074–.093), and SRMR (.064) values, as well as high CFI (.914) and TLI (.903) fit indices. Additional details concerning individual item loadings appear in Table 5 in online supplementary materials.

Overall discussion

Results of the current investigation provide strong, initial support for the reliability and validity of the FFFQ as a measure of temperamental fear. This outcome likely owes in strong part to the instrument's

development guided by rigorous suggestions for test construction (Haynes et al., 1995), as well as careful attention to the integration of theory into all aspects of research. For example, when the first iteration of the instrument failed to produce theoretically consistent scales, steps were taken to revise the item pool to more carefully reflect "fight" stimuli. Additionally, when the second version produced adequate factor loadings for items on an unexpected subscale (i.e., flight items loaded on freeze), items were eliminated from the measure.

This process ultimately yielded a final 21-item scale comprising three factors of separate (albeit related) dimensions (i.e., fight, flight, and freeze). Categorization of all items on the final measure was supported by the initial development sample's responses to vignettes, expert review, theoretical assertions of the authors, and high factor loadings. In short, the careful attention to understanding these constructs from the perspective of an average individual culminated in a strongly theoretically cogent instrument that addresses a long-identified gap in the literature. For example, widely-cited studies have called for not only a trait-like fear assessment, but also a scale to assess and differentiate a more general trait fear compared to anxiety (Cooper et al., 2007; NIMH RDoC Team, 2011; Smillie, Pickering, & Jackson, 2006; Sylvers et al., 2011). Results comparing the FFFQ with temperamental anxiety (BIS), the STAI-T, the PANAS, and the NEO-FFI offer preliminary support for a measure of trait-like response to fear that is distinguishable from underlying anxiety.

Specifically, when examining the FFFQ in comparison with the two other constructs from BIS/BAS scales, expected relationships were found. The FFFQ demonstrated only a small, significant correlation with BIS functioning ($r = .18$) with no association with BAS functioning. In addition, when individual subscales of the FFFQ were assessed, similar associations remained. This discovery adds initial support for the FFFQ as a much needed assessment of fear, based on conceptualization of trait fear as related to the fight, flight, freeze system.

Additionally, as shown in Table 1, the associations between the three subscales of the FFFQ are generally weak, but have moderate to high associations with the overall FFFQ score (r's = .54–.76). Theoretically, this finding was interesting and conceptually relevant in thinking about individuals' temperamental vulnerabilities. It is possible that individuals have an idiographic, habitual, primed mechanism of typical response (i.e., fight, flight, freeze) and specific traits may be high for only one of the sub-factors of fear. For example, the non-significant relation between the Fight and Freeze subscales is not surprising and can certainly be reconciled within the theoretical description of approach and avoidance necessitated for each factor. Specifically, individuals who were more prone to fight also demonstrated *negative* associations with BIS (which underlies anxiety). This relationship illustrates individuals with *decreased* inhibition (avoidance) are more likely to have general Fight tendencies (defensive approach) whereas individuals higher on Freeze demonstrated the opposite: *positive* associations with BIS. Increased inhibition may be associated with tendencies to avoid any defensive approach, which manifests as freeze tendencies.

Limitations and future directions

Despite preliminary findings that the FFFQ has high internal consistency, good test–retest reliability, and adequate convergent/divergent validity, the present studies have limitations to be addressed in future research. One limitation is reliance on only a self-report assessment of trait fear. The absence of behavioral data thus precluded examination of the practical predictive validity of the FFFQ, which could be an important focus of future research. Additionally, the current studies involved samples taken from a normative college population, and the degree to which the results generalize to other groups or contexts is not yet clear. It is therefore important that the FFFQ be further validated in diverse samples and settings, particularly in clinical environments. Additionally, examination of its associations with laboratory tasks designed to directly observe behavioral components of fear (including fight, flight, and freeze factors) would simultaneously contribute to further psychometric study and a more thorough understanding of the relationship between self-reported and observed aspects of fear.

Another limitation was the lack of other self-report measures of trait-like response to fear for comparison. As a result, the associations between the FFFQ and other aspects of the emotion of fear remain to be determined. To further examine the psychometric properties of the FFFQ, future studies should elaborate how this instrument compares to measures of phobic fear and other pathology, as well as relevant dimensions of personality. Studies to explore the relation of FFFQ with negative affect and other separate, but related constructs would be beneficial and a direction for future studies to replicate and extend these initial findings. In addition, further studies could be conducted to establish norms for the FFFQ in both community and clinical samples.

Conclusions

The goals of the present group of studies were to develop and initially validate the psychometric properties of a self-report measure assessing the general construct of fear in line with the NIMH RDoC initiative to begin differentiating between fear and anxiety. The FFFQ is a promising measure that provides a brief, self-report assessment of underlying fear tendencies. Related to general fear motivation as well as allowing for measurement of the three components of fear (fight, flight, and freeze), the FFFQ meets the demonstrated need of a more comprehensive assessment of responses to acute threat. Further research is needed to replicate the findings as well as to examine the discriminant validity of the FFFQ, especially with respect to other measures of general distress or psychopathology. Additionally, research exploring the differential role of the various dimensions of fear in clinical presentations (e.g., phobias, PTSD, psychopathy) will be especially important, as it may suggest specific targets for intervention.

Disclosure statement: The authors have declared that no conflict of interest exists.

Supplemental data

Supplemental data for this article can be accessed doi:10.1080/16506073.2014.972443.

Note

1. Vignettes can be obtained by emailing the authors.

References

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

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Blanchard, D.C., Hynd, A.L., Minke, K.A., Minemoto, T., & Blanchard, R.J. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of non-human mammals. *Neuroscience and Biobehavioral Reviews*, *25*, 761–770.

Blankstein, K.R. (1975). The sensation seeker and anxiety reactivity: Relationships between the sensation-seeking scales and the activity preference questionnaire. *Journal of Clinical Psychology*, *31*, 677–681.

Bryant, F.B., & Yarnold, P.R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L.G. Grimm & P.R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–135). Washington, DC: American Psychological Association.

Carver, C.S., & White, T. (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.

Cooper, A.J., Perkins, A.M., & Corr, P.J. (2007). A confirmatory factor analytic study of anxiety, fear, and behavioral inhibition system measures. *Journal of Individual Differences*, *28*, 179–187.

Corr, P.J. (2008). Reinforcement sensitivity theory (RST): Introduction. In P.J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 1–43). New York, NY: Cambridge University Press.

Costa, P.T., & McCrae, R.R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) manual*. Odessa, FL: Psychological Assessment Resources.

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

Haynes, S.N., Richard, D.C., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*, 238–247.

Heym, N., Ferguson, E., & Lawrence, C. (2008). An evaluation of the relationship between Gray's revised RST and Eysenck's PEN: Distinguishing BIS and FFS in Carver and White's BIS/BAS scales. *Personality and Individual Differences*, *45*, 709–715.

Jöreskog, K.G., & Sörbom, D. (1981). *LISREL V [Computer software]*. Chicago, IL: Scientific Software International.

Jorm, A., Christensen, H., Henderson, A.S., Jacomb, P.A., Korten, A.E., & Rodgers, B. (1999). Using the BIS/BAS scales to measure behavioral inhibition and behavioral activation: Factor structure, validity and norms in a large community sample. *Personality and Individual Differences*, *26*, 49–58.

Krueger, R.F. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, *56*, 921–926.

Lorenzo-Seva, U., & Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods, Instruments, & Computers*, *38*, 88–91.

Lykken, D.T., Tellegen, A., & Katzenmeyer, C.G. (1973). *Manual for the activity preference questionnaire* (Report No. PR-73-4, Reports from the Research Laboratories). Minneapolis, MN: Department of Psychiatry, University of Minnesota.

Marks, I.M., & Mathews, A.M. (1979). Brief standard self-rating for phobic patients. *Behavior Research and Therapy*, *17*, 263–267.

NIMH RDoC Team. (2011, March 13–15). NIMH Research domain criteria (RDoC) project. *Negative valence systems: Workshop proceedings*. Retrieved from http://www.nimh.nih.gov/research-funding/rdoc/negative-valence-systems-workshop.pdf

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw Hill.

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

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

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

Preacher, K.J., & MacCallum, R.C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, *2*, 13–32.

Smillie, L.D., Pickering, A.D., & Jackson, C.J. (2006). The new reinforcement sensitivity theory: Implications for personality measurement. *Personality and Social Psychology Review*, *10*, 320–335.

Spielberger, C.D., Gorsuch, R.L., Lushene, R.E., Vaag, P.R., & Jacobs, G.A. (1983). *Manual for the state-trait anxiety inventory (Form Y) (self-evaluation questionnaire)*. Palo Alto, CA: Consulting Psychologists Press.

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

Sylvers, P., Lilienfeld, S.O., & LaPrairie, J.L. (2011). Differences between trait fear and trait anxiety: Implications for psychopathology. *Clinical Psychology Review*, *31*, 122–137.

Tabachnick, B.G., & Fidell, L.S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Tellegen, A. (1982). *Brief manual for the multidimensional personality questionnaire*. Unpublished manuscript. University of Minnesota, Minneapolis, MN.

Tellegen, A., & Waller, N.G. (2007). *Exploring personality through test construction: Development of the multidimensional personality questionnaire*. Minneapolis, MN: University of Minnesota Press.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis*. Washington, DC: American Psychological Association.

Torrubia, R., Avila, C., & Caseras, X. (2008). Reinforcement sensitivity scales. In P. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 188–227). New York, NY: Cambridge University Press.

Tucker, L.R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, *114*, 522–536.

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*, *47*, 1063–1070.

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