The Anxiety and Illness Behaviour Laboratory, Department of Psychology, University of Regina, Regina, SK; 2 Department of Psychiatry, McMaster University and Anxiety and Illness Behaviour Laboratory, Department of Psychology, University of Regina, Regina, SK; 3 Department of Psychiatry, Ryerson University

**Method**

- Undergraduate Participants, University of Regina
  - 61 men, ages 18-34 (M = 20.6; SD = 3.2)
  - 212 women, ages 18-32 (M = 20.2; SD = 3)

- Clinical Participants, Anxiety Treatment Centre at St. Joseph’s Healthcare, Hamilton, ON; 3 Department of Psychiatry, Ryerson University
  - 162 men, ages 16-69 (M = 34.8; SD = 11.7)
  - 193 women, ages 16-64 (M = 32.9; SD = 11.3)

- Demographics were supplemented with:
  - Social Phobia Inventory (SPIN; Conner, et al., 2000)
  - A 17-item 5-point Likert scale, 0 (not at all) to 4 (extremely), self-report measure to assess symptoms specific to SAD.

- Clinical sample internal consistency, α = .91, average inter-item correlation, r = .27.
- Undergraduate sample internal consistency, α = .94, average inter-item correlation, r = .48.

- Both samples were randomly divided in half using SPSS, allowing for EFA and CFA.

- The EFAs were performed using the recommendations based on Osborne (2008).

- CFA fit indices (Hu & Bentler, 1999): χ²/df ratio (χ²/df, should be <2.0); Comparative Fit Index (CFI; should be close to .95); Root Mean Square Error of Approximation (RMSEA; should be close to .06); Expected Cross Validation Index (ECVI; lower values, better fit).

- There were no significant differences (all p > .10) on any of the items between the two randomized groups, or significant interactions based on sample or sex.

- In the clinical sample, there were significant differences on several SPIN items between men and women (Table 1); there were no such differences in the undergraduate sample. Given the absence of a priori differences related to sex, those items were removed prior to EFA.

- Results of EFAs using Osborne’s (2008) recommendations with a half-interactive component of the undergraduate samples suggested a 2-factor 9-item solution accounting for 51% and 54% of the variance respectively; the first factor was comprised of fear and avoidance items, the second factor comprised of four physiological items.

- Results of the CFA with the other half of the clinical sample suggested the 2-factor model did not have an acceptable fit to the data (Table 2). Results of the CFA with the other half of the clinical sample suggested the 2-factor model had an acceptable fit to the data; a unitary 5-item solution using only items from factor one produced the best solution.

- The 5- and 9-item versions correlated highly (r = .91; r = .97) with the full measure in both samples. Results of a receiver operator curve analyses (ROC) supported the utility of 5- and 9-item versions in distinguishing clinical and undergraduate participants (AUC = .87; AUC = .86) at a rate comparable to the full measure (AUC = .88) and the 3-item mini spin (AUC = .87).

**Results**

- Items removed prior to the EFA were significantly different across men and women, and tend to focus on fears of authority and criticism; this apparent conceptual difference between men and women warrants subsequent research.

- The CFA results suggest that the items in the SPIN do not form three independent factors but, instead, form two factors that can be conceptualized as fear/avoidance and physiological arousal (Figure 1).

- The results of the CFA support a 5-item unitary and a 9-item 2-factor solution for the undergraduate sample; however, those solutions are not necessarily appropriate for the clinical sample. Despite content differences they are also comparable features of five mini spin items.

- In line with previous research (Connor et al., 2001), it appears there are several items that do not load robustly on to any factor and may be unnecessary for discriminating measurements of SAD. Nevertheless, it may be that the removed items provided additional specificity not accounted for in the current analyses. Future research should evaluate these possibilities before making definitive exclusions.

- For nonclinical samples, the first factor is not as distinct as the second. It may also represent unexplored heterogeneity between and within clinical and non-clinical presentations of SA based on symptoms of physiological arousal. Given the variety of items producing comparable ROC results, further research is needed into how we should measure SAD.

**Discussion**

- **Items**: The 5- and 9-item versions correlated highly (r = .91; r = .97) with the full measure in both samples. Results of a receiver operator curve analyses (ROC) supported the utility of 5- and 9-item versions in distinguishing clinical and undergraduate participants (AUC = .87; AUC = .86) at a rate comparable to the full measure (AUC = .88) and the 3-item mini spin (AUC = .87).

- **Items**: The CFA results suggest that the items in the SPIN do not form three independent factors but, instead, form two factors that can be conceptualized as fear/avoidance and physiological arousal (Figure 1).

- The results of the CFA support a 5-item unitary and a 9-item 2-factor solution for the undergraduate sample; however, those solutions are not necessarily appropriate for the clinical sample. Despite content differences they are also comparable features of five mini spin items.

- In line with previous research (Connor et al., 2001), it appears there are several items that do not load robustly on to any factor and may be unnecessary for discriminating measurements of SAD. Nevertheless, it may be that the removed items provided additional specificity not accounted for in the current analyses. Future research should evaluate these possibilities before making definitive exclusions.

- For nonclinical samples, the first factor is not as distinct as the second. It may also represent unexplored heterogeneity between and within clinical and non-clinical presentations of SA based on symptoms of physiological arousal. Given the variety of items producing comparable ROC results, further research is needed into how we should measure SAD.