Short Communication

Gender-based measurement invariance of the Substance Use Risk Profile Scale

Jasmina Memetovic [a,*], Pamela A. Ratner [b,c,1], Christopher G. Richardson [a,2]

[a] *School of Population and Public Health, Faculty of Medicine, University of British Columbia, 2206 East Mall, Vancouver, BC, V6T 1Z3| Canada*
[b] *School of Population and Public Health, Faculty of Medicine, University of British Columbia, 302-6190 Agronomy Road Vancouver, BC, V6T 1Z3 Canada*
[c] *School of Nursing, Faculty of Applied Science, University of British Columbia, Vancouver, BC, 302-6190 Agronomy Road Vancouver, BC, V6T 1Z3 Canada*

HIGHLIGHTS

- We demonstrated gender-based measurement invariance of the SURPS
- A confirmatory factor analysis demonstrated good fit of a 4-dimensional structure
- Findings contribute to validation of the SURPS in non-clinical adolescent samples

ARTICLE INFO

ABSTRACT

Gender may influence the psychometric properties of psychological and affect-related measures. The American Psychological Association has made recommendations to conduct tests of group-level measurement invariance (MI) before comparing scores between groups. Gender-based measurement invariance has been examined in many well-known psycho-social scales such as the CES-D and the Big Five Inventory. The Substance Use Risk Profile is a 23-item scale measuring affective- and personality-related traits known to increase risk for substance use, with 4 dimensions: anxiety sensitivity, hopelessness, sensation seeking and impulsivity. Despite similarities in the constructs assessed by the SURPS, the CES-D and the Big Five Inventory, gender-based measurement invariance of the SURPS has not yet been published. Multi-group confirmatory factor analysis was used to assess the measurement invariance of the four dimensions of the SURPS across gender. MI was conducted with M-Plus 6.2 using a 2-step analysis for ordinal variables suggested by Muthén and Muthén, and model fit was assessed using the comparative fit index (CFI) criteria recommended by Cheung and Rensvold. A single group confirmatory factor analysis (CFA) was also conducted. The sample was composed of 1352 adolescents (56% female, mean age of 14 years) participating in the BC Adolescent Substance Use Survey, an online survey capturing substance use and psychosocial trends in secondary students across British Columbia, Canada. Measurement invariance across gender was demonstrated for the SURPS (ΔCFI = 0.003), and the single group CFA supported a four-dimensional structure for the SURPS items (CFI = 0.92, RMSEA = 061, 95% CI = 0.058–0.065).

1. Introduction

Research of possible gender differences in psychosocial and behavioral traits frequently involves the use of psychometric instruments to measure the constructs under investigation. Before considering the interactive effects of gender on traits or comparing the total scores of psychological measures, it is important to first establish whether the underlying psychometric properties of the scales used to measure the relevant latent constructs are measurement invariant across the genders (Meredith, 1993; Millsap & Yun-Tein, 2004; Steenkamp & Baumgartner, 1998). Scales with gender-sensitive items could lead to biased total scores for a particular group, which could subsequently distort or bias any relationships observed among the latent constructs and the outcomes of interest (Meredith, 1993; Millsap & Yun-Tein, 2004; Steenkamp & Baumgartner, 1998). An instrument that is found to be structurally equivalent, in terms of its measurement properties, in all groups of interest is said to be measurement invariant (Breithaupt & Zumbo, 2002; Horn, Jack, & Mason, 1983; Marsh, Nagengast, & Morin, 2012; Steenkamp & Baumgartner, 1998; Zumbo, 1999).

Methodological guidelines set forth by the American Psychological Association (1999) recommend that researchers perform validity tests of psychological assessment tools, such as assessments of measurement invariance between groups of interest, before using them to explore differences among those groups. An instrument that is found to be

* Corresponding author. Tel.: +1 778 855 9377.
  E-mail addresses: jmem@shaw.ca (J. Memetovic), pam.ratner@ubc.ca (P.A. Ratner),
chris.richardson@ubc.ca (C.G. Richardson).
  [1] Tel.: +1 604 822 7427.
  [2] Tel.: +1 604 827 4023.

structurally equivalent, in terms of its measurement properties, in all groups of interest is said to be measurement invariant (Breithaupt & Zumbo, 2002; Horn et al., 1983; Marsh et al., 2012; Steenkamp & Baumgartner, 1998; Zumbo, 1999). Despite the aforementioned guidelines, surprisingly few studies involving the use of psychological measures have established their gender-based measurement invariance. A few noteworthy exceptions include a gender-based measurement invariance of the Centre for Epidemiological Studies Depression Scale (CES-D) (Breithaupt & Zumbo, 2002; Zumbo & Kho, 2005), the Perceived Stress Scale (PSS) (Lavoie & Douglas, 2012) and, more recently, the NEO-Five Inventory (Marsh et al., 2012). An examination of the CES-D by Zumbo and Koh found that, given equivalent levels of depressive symptomatology, women were more likely to endorse the "crying spells" item, and men were more likely to endorse "talking less than usual" (Breithaupt & Zumbo, 2002; Zumbo & Kho, 2005). Since the SURPS construct of *hopelessness* is very similar to the concept of depression assessed by the CES-D, it too may contain items that are more (or less) likely to be endorsed by one gender over the other. Most recently. Marsh demonstrated gender-based differences in latent mean scores in all the Big Five Inventory dimensions, except for openness (Marsh et al., 2012). Although there is considerable overlap between the NEO-Five Factor Inventory dimensions and the SURPS (Woicik, Stewart, Pihl, & Conrod, 2009, the wording of the SURPS and the NEO-Five Inventory items are quite different. Establishing gender-based measurement invariance of the SURPS represents an important step towards establishing the validity of this recently developed measure and specifically ensuring that total scores derived from it are not gender biased (Steenkamp & Baumgartner, 1998; Zumbo, 1999; Zumbo & Kho, 2005).

1.1. The Substance Use Risk Profile Scale

The development of the Substance Use Risk Profile Scale (SURPS), which measures personality-related risk of substance abuse, was informed by research related to motivation and the personality profiles of alcohol users (Cloninger, 1987a; Conrod, Petersen, & Pihl, 1997; Conrod, Pihl, Stewart, & Dongier, 2000; Cooper, Russell, Skinner, & Windle, 1992). The items were derived from well-known personality-related inventories (Conrod et al., 2000; Krank et al., 2011; Woicik et al., 2009), such as the Tri-dimensional Personality Questionnaire (Cloninger (1987b, 1987c)) and the NEO-Five Factor Personality Inventory (Costa & McCrae, 1992). The SURPS is based on four personality- and affect-related factors or styles that have previously been shown to increase the risk of engaging in maladaptive behavior related to substance misuse and abuse (Caspi, Moffitt, Newman, & Silva, 1996; Conrod, Pihl, Stewart, & Dongier, 2000; Krank et al., 2011; Woicik et al., 2009). The most recently validated 23-item version of the SURPS captures four constructs that profile a risk for substance use: anxiety sensitivity, hopelessness, sensation seeking and impulsivity. The dimensions have been assessed for current and predictive validity with substance use in clinical and non-clinical samples, and the dimensions *impulsivity* and *sensation seeking* have been shown to be associated with substance use (Conrod, Pihl, Stewart, & Dongier, 2000; Krank et al., 2011; Woicik et al., 2009).

1.2. Measurement invariance

When an instrument is said to be measurement invariant, each construct is measured with the same metrics across the study groups (Horn et al., 1983). As described by (Meredith, 1993; Steenkamp & Baumgartner, 1998), instrument bias can result from differences in the construct configuration, the magnitude of the factor loadings, the intercepts or the residual variances (Meredith, 1993; Steenkamp & Baumgartner, 1998). When conducting a measurement invariance assessment of a latent variable with ordinal indicators, factor loading, factor variance and residual (error) variance invariance can be simultaneously tested in two steps to establish *strict invariance* (Marsh et al.,

2012; Muthén & Christoffersson, 1981). If strict invariance of the instrument is shown, both construct correlations and means can be compared across groups.

1.3. Study aims

This study aimed to assess the gender-based measurement invariance of the Substance Use Risk Profile Scale (SURPS) using a multi-group confirmatory factor analysis (CFA) approach recommended by Muthén and Muthén (2010). The results of an un-grouped CFA are also presented to provide further support for the four-factor model of the SURPS; these results are compared with previously published analyses (Krank et al., 2011; Woicik et al., 2009).

2. Materials and methods

2.1. Data source and study design

The BC Adolescent Substance Use Survey (BASUS) was a prospective cohort study designed to capture changes in psychosocial functioning and substance use trends in adolescents attending publicly funded English-speaking secondary schools (grades 8 through 12) in British Columbia, Canada. The cross-sectional data analyzed here were obtained from a sample of adolescents ($N = 1,352$) enrolled in grades 8 and 9 in 62 BC secondary schools (i.e., from BASUS Wave 3 in the fall of 2010). The study was approved by the University of British Columbia Behavioural Research Ethics Board.

2.2. Measures

2.2.1. The Substance Use Risk Profile Scale
The SURPS contains 23 items that measure four dimensions of personality-related risk of substance abuse: anxiety sensitivity (5 items), hopelessness (7 items), sensation seeking (6 items) and impulsivity (5 items). Responses to each item were provided on a 4-point Likert-type scale: 1 ("strongly disagree") to 4 ("strongly agree").

2.2.2. Gender
Participants were asked to identify themselves as being either male or female.

3. Analyses

3.1. Gender-based measurement invariance

The goal of the analyses presented here was to examine the gender-based measurement invariance of the SURPS using multi-group CFA, with the M-plus 6.12 software, according to the methods described by Muthén and Muthén (2010) and Millsap and Yun-Tein (2004). A robust weighted least squares estimator (WLSMV) was used to accommodate the categorical nature of the data. With ordered, categorical observed variables, it is threshold invariance that is tested rather than intercept or latent mean invariance (Muthén & Muthén, 2010). Thresholds, or the number of item categories less one, can be used to estimate the probability of a latent response (for a more detailed description, see Feldman, Masyn, & Conger, 2009). To test for measurement invariance, two nested models of the invariance parameters were statistically compared (as per Muthén & Muthén, 2010) to determine whether the difference in the comparative fit index (CFI) was ≤ -0.01 (Cheung & Rensvold, 2002; Hu & Bentler, 1999).

3.2. Confirmatory factor analysis and reliability estimates

Once the measurement invariance of the four dimensions of the SURPS was demonstrated, a single-group CFA was conducted to examine the overall fit of the four-factor model. The comparative fit index