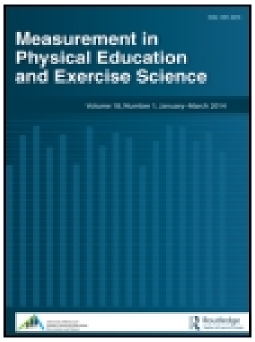

Measurement in Physical Education and Exercise Science

ISSN: 1091-367X (Print) 1532-7841 (Online) Journal homepage: https://www.tandfonline.com/loi/hmpe20

Bifactor Approach to Modeling Multidimensionality of Physical Self-Perception Profile

ChihMing (Ryan) Chung, Xiaolan Liao, Hairong Song & Taehun Lee

MEASUREMENT IN PHYSICAL EDUCATION AND EXERCISE SCIENCE
2016, VOL. 20, NO. 1, 1–15
http://dx.doi.org/10.1080/1091367X.2015.1081594

Routledge
Taylor & Francis Group

Bifactor Approach to Modeling Multidimensionality of Physical Self-Perception Profile

ChihMing (Ryan) Chung[a], Xiaolan Liao[a], Hairong Song[b], and Taehun Lee[b]

[a]Office of Academic Assessment, University of Oklahoma, Norman, Oklahoma; [b]Department of Psychology, University of Oklahoma, Norman, Oklahoma

ABSTRACT
The multi-dimensionality of Physical Self-Perception Profile (PSPP) has been acknowledged by the use of correlated-factor model and second-order model. In this study, the authors critically endorse the bifactor model, as a substitute to address the multi-dimensionality of PSPP. To cross-validate the models, analyses are conducted first in exploratory structural equation modeling (ESEM) framework with a randomly selected subsample and then in confirmatory factor analysis (CFA) framework with a second subsample. Results from both ESEM and CFA analyses suggest that bifactor model is the best-fitted model. A general physical self-esteem factor and four domain specific factors are identified with the bifactor model. Coefficient omega hierarchical of the general factor is .86. The ω_s of the specific factors are .19, .20, .59, and .29, respectively. Gender difference at both general factor level and domain specific factor level is examined within the bifactor model. Discussions of the use and limitations of bifactor model as well as ESEM are provided.

KEYWORDS
bifactor model; exploratory structural equation modeling; multi-dimensionality; Physical Self-Perception Profile

This article presents a new approach to analyzing the multi-dimensional structure of Physical Self-Perception Profile (PSPP) (Fox & Corbin, 1989) using bifactor analysis. By applying the bifactor model to explore underlying factors of physical self-perception and comparing it to other traditional approaches (i.e., common one-factor model, correlated multiple-factor model, and second factor model), this study advances the understanding not only of the PSPP scale and its relationship with external variables (e.g., gender) but also of the utility of the bifactor model in multi-dimensional analysis. Cross-validation of four models are conducted using exploratory structural equation modeling (ESEM) and confirmatory factor analysis (CFA), the results of which both suggest that of the correlated four-factor model, second-order factor model, and bifactor model, the bifactor model represents the best fit.

Physical self-perception is defined as a sense of competence in the physical domain involving physical appearance and physical body movement (Fox & Corbin, 1989; Harter, 1999). Specific perceptions within the physical domain, such as sport competence, physical strength, physical condition, and body attractiveness, are recognized as influential in the development of the sense of physical self-perception. Fox and Corbin (1989) developed the PSPP, a well-known psychometric tool used to examine self-perception in the physical domain and the most widely used measurement in the fields of physical education, sport psychology, and social psychology research to assess levels of physical self-perception.

The PSPP is based on a hierarchical, multi-dimensional, and theoretical model using theory of self-perception profiles created by Harter (Fox & Corbin, 1989; Harter, 1985). A major advancement for measuring the concept of self-perception was the widespread acceptance of multi-dimensionality in the late 1980s (Fox & Corbin, 1989). Fox and Corbin (1989) included four subdomains in the PSPP to provide a more precise measurement for investigating the concept of physical self-perception among college-aged students. The four sub-domains in the 30-item self-report inventory consist of the following: (a) Sports Competence (Sport), (b) Physical Strength (Strength), (c) Physical Condition (Condition), and (d) Body Attractiveness (Body). Each subscale consists of six items in which participants are presented with two contrasting descriptions. A fifth subscale was initially included in the PSPP to measure a global Physical Self-Worth construct.

Fox and Corbin (1989) used four data collection phases which included the sub-domain identification (phase one), instrument construction (phase two),

CONTACT ChihMing (Ryan) Chung rchung@ou.edu; Xiaolan Liao xiaolanliao@ou.edu Office of Academic Assessment, University of Oklahoma, 860 Van Vleet Oval, Copeland Hall 101, Norman, OK 73019.

instrument reliability and factorial validity (phase three), and factor confirmation and preliminary construct validity (phase four). Each phase had a specific function in the construction of the profile.

In phase one, an open-ended questionnaire was designed to allow college students to provide more direct insight on the content for physical self-perceptions of college-aged students. The purpose of the subdomain identification was to indicate possible subscales of the physical self as the foundation for constructing a multidimensional physical self-perception instrument. Four subdomains were chosen from the initial profile, including body attractiveness, sports competence, physical strength, and fitness and exercise. In phase two, the instrument construction process, six questions were developed under each of the four sub-domains identified in phase one (Fox & Corbin, 1989). In the third phase, three data collection processes were included to assess instrument reliability and factorial validity (Fox & Corbin, 1989). Exploratory factor analyses (EFA) were conducted and a four-factor structure emerged from the items found in phase one. Items and subscale stability were assessed based on test–retest estimates. The PSPP subscales demonstrated preliminary internal reliability, as well as stability over samples. In phase four, factor confirmation and preliminary construct validity, was designed to confirm the psychometric properties and factor structure of the PSPP (Fox & Corbin, 1989). Results of the fourth phase yielded the finished product of the PSPP. The reliability of the PSPP established coefficient alphas ranging from .81 to .92.

Karteroliotis (2008) examined the validity of the PSPP using college students attending two midwestern universities in the United States. The purpose of the study was to examine the four-subscale structure of the PSPP, and to assess any potential differences in the factor structure across genders. Exploratory factor analysis (EFA) and CFA were applied in the study. The researcher indicated that 67.9% of the variance was explained in the male data and 70% of the variance was explained by the factor structure in the female data through the EFA. However, Karteroliotis did not find evidence to fully support the proposed four-factor structure for both males and females in the study based on the CFA. The Chi-square goodness-of-fit test was statistically significant in both groups (Karteroliotis, 2008).

Many researchers have confirmed the PSPP as a useful psychometric inventory to assess the concept of physical self-perception, especially among college students (Asçi, Asçi, & Zorba, 1999; Fox & Corbin, 1989; Karteroliotis, 2008; Xu & Yao, 2001). As a result of past studies, many researchers have used the PSPP as a helpful measurement in the fields of physical education, sport psychology, and social psychology research to assess levels of physical self-perception (Fox & Corbin, 1989; Hayes, Crocker, & Kowalski, 1999; Lindwall & Hassmen, 2004).

As PSPP becomes a widely used tool to assess the concept of physical self-perception and the development in the methodological world, understanding of PSPP has been advanced. Initially, Fox and Corbin (1989) used a correlated four-factor model to handle the multi-dimensionality of the PSPP. In a correlated four-factor model, each of the four latent factors loaded on the intended items and the four latent factors are allowed to be correlated (Figure 1(b))(Fox & Corbin, 1989). (Presented in Figure 1(a) is a common one-factor model, which will be introduced later when the authors describe the bifactor model.) However, the correlated four-factor model identified by Fox and Corbin (1989) was not able to capture the general factor that was hypothesized in their theoretically hierarchical model.

With the advancement in statistical modeling, second-order factor model was adopted to handle the multi-dimensionality of PSPP (Hagger, Asci, & Lindwall, 2004; Hagger, Biddle, Chow, Stambulova, & Kavussanu, 2003). In a second-order factor model, the entire first-order factors load on the intended observed variables and a second-order factor is introduced to the model to explain the covariation among the first-order factors (Figure 1(c)). A second-order factor representing general physical self-esteem factor (as the authors term it) and four first-order factors representing the four subdomain physical areas were identified in their work (Hagger et al., 2003, 2004).

Recently, a bifactor model has been proposed to handle multi-dimensionality of concepts in a variety of fields for its superiority to second-order factor model (Chen, West, & Sousa, 2006; Reise, Morizot, & Hays, 2007). Bifactor model was first introduced by Holzinger to demonstrate the multi-dimensionality of preliminary reports on the Spearman–Hotzinger Unitary Trait (Holzinger & Swineford, 1937). Although bifactor model was developed a few decades ago, its superiority over second-order factor model in handling multidimensional concept has only recently been recognized. Bifactor model has been used primarily in studying intelligence (Acton & Schroeder, 2001; Gault, 1954; Gignac & Watkins, 2013; Golay, 2011; Hammer, 1950; Jensen & Weng, 1994; Watkins, 2010; Watkins & Beaujean, 2014) and personality (Armon & Shirom, 2011; Martel, Roberts, Gremillion, von Eye, & Nigg, 2011; McAbee, Oswald, & Connelly, 2014; McCrae & John, 1992; Revelle & Wilt, 2013; Rushton & Irwing, 2009a, 2009b, 2009c). Recently it has been applied in various areas including depression and anxiety, health

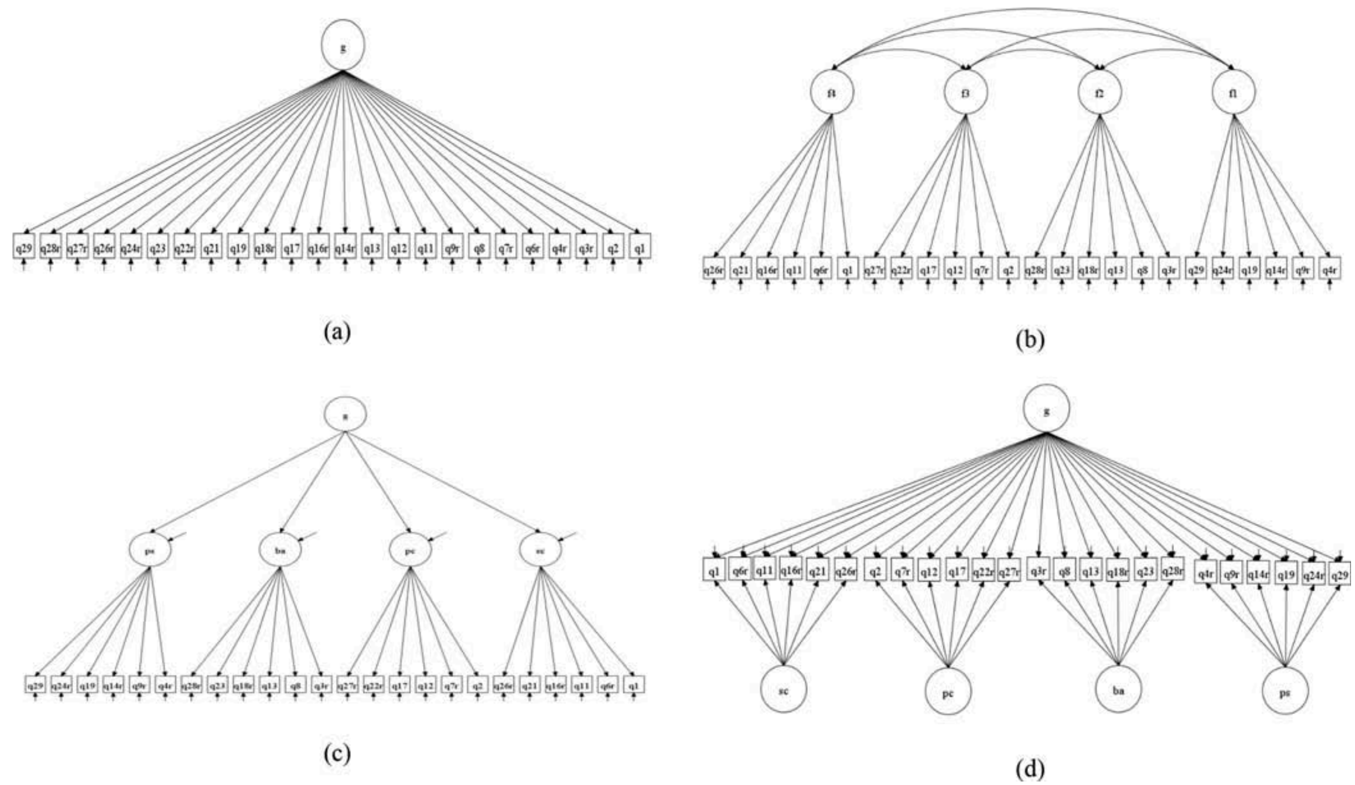

Figure 1. (a) Common one-factor model, (b) correlated four-factor model, (c) second-order factor model, and (d) bifactor model.

outcomes, sport performance, quality of life, Attention-Deficit/Hyperactivity Disorder (ADHD), and the Positive and Negative Affect Schedule (PANAS) (Chen et al., 2006; Leue & Beauducel, 2011; Martel et al., 2011; Myers, Martin, Ntoumanis, Celimli, & Bartholomew, 2014; Reise et al., 2007; Simms, Grös, Watson, & O'Hara, 2008; Xie et al., 2012).

Bifactor model can be viewed as an extension to common one-factor model. A common one-factor model (Figure 1(a)) hypothesizes that a single factor accounts for the covariation among a set of observed variables, in which the entire observed variables load on the common factor. Bifactor model extends its precursor by including domain specific factors aiming to explain the covariation among the residuals that are resulting from the general factor. A bifactor model hypothesizes that (a) there is a general factor that accounts for the communality shared by the observed variables and (b) there are multiple domain specific factors, each of which accounts for the residual covariation among a cluster of items (Figure 1(d)) (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; Reise, 2012; Reise, Moore, & Haviland, 2010). In a bifactor model, all items load on the general factor and one domain specific factor (Reise, 2012).

Bifactor model is an alternative to handle multi-dimensionality of a scale. Traditional approaches to modeling multi-dimensionality include multi-factor model and hierarchical model (e.g., second-order factor model). Chen and colleagues compared the strengths of bifactor model and a few other approaches in modeling multi-dimensionality (Chen et al., 2006, 2012). The bifactor model is endorsed for its advantages, namely its ability to (a) separate domain specific factors from the general factor, (b) study the relation between items and general factors, and between items and domain specific factors, (c) identify whether a domain specific factor still exists after partialling out the general factor, (d) test whether a subset of the domain specific factors predict external variables, over and above the general factor, (e) test mean difference on both the general and specific factor levels, and (f) test measurement invariance at both the general and specific factor levels (Chen et al., 2012, 2006; Reise, 2012; Reise et al., 2010).

Methodology

In the current study, the intention is to propose the use of bifactor model to handle the multi-dimensionality of PSPP. By comparing bifactor model to the three traditional latent factor models (i.e., common one-factor model, correlated multiple-factor model, and second factor model) in terms of model fit and interpretation of the results, the study shows how the bifactor model can help advance the standing of PSPP and its relationship with external variables (e.g., gender). To cross-validate the tested models, the authors first conduct exploratory factor analysis with a randomly selected sample with n = 250 from the entire data and then do CFA with the remaining sample with n = 150. The exploratory analysis of the three models correlated multiple-factor model, second factor model, and bifactor model are conducted in the framework of ESEM under target rotation. There are a number of methods to determine the sample size for exploratory factor analysis (Henson & Roberts, 2006; Schmitt, 2011). In the authors' practice, they adopt the rule of thumb suggested by Stevens (1996) that five to 20 participants per variable are appropriate. In the current study, the ratio of participants to variables is approximately 10:1 (i.e., 250:24), suggesting that the sample size is sufficient. This decision leaves a subsample with n = 150, which is sufficient to test the confirmatory factor models. All the factor analyses are conducted with Mplus 7.

Participants

In the study, the participants were 400 full-time male (n = 200) and female (n = 200) undergraduate students, who were at least 18 years old and enrolled in three medium-sized colleges and universities in northeastern United States. In the data selection process, the researcher checked each questionnaire after every classroom visit. A total of 400 students were selected. Only questionnaires answered completely were included. All participants were asked to read and sign an Informed Consent form prior to participation. The authors assumed no sample can be 100% representative of the entire population, however they did consider the most relevant factors regarding a reasonable sample size and representation of the population for this study before they collected their data (i.e., total population, margin of error, confidence level). The authors have identified that this study requires at least 386 respondents. They believe that this sample (N = 400) is sufficient and represents the relevant characteristics for the population.

Data collection procedures

With IRB approval, college and university administrators from a set of randomly selected 4-year schools on a predetermined list were contacted for approval to participate in the study. General education course instructors at each institution were then consented and the

PSPP instrument was administered in person to a sample of students in spring 2010 courses. The researcher was in attendance during the 15–20-minute completion timeframe to guide through the question format using an example and answer questions.

Analysis

Exploratory factor analysis

The robust continuous maximum likelihood estimation (MLR) as a method to extract factors in the exploratory factor analyses[1]. MLR is one of the most widely used factor-extraction method for its capacity in providing a wide range of model fit indexes and significance test of factor loadings (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Schmitt, 2011). The MLR method provides maximum likelihood parameter estimates with standard errors and a Chi-square test statistic that are robust to non-normality. In evaluating and comparing the models, the authors rely on several goodness-of-fit indexes along with the Chi-square test: the comparative fit index (CFI; Bentler, 1990), the Tucker–Lewis index (TLI; Tucker & Lewis, 1973), the root mean square error of approximation (RMSEA; Steiger, 1990) with its confidence interval, the standardized root mean square residual (SRMR; Hu & Bentler, 1999), the Akaike information criteria (AIC; Akaike, 1987), the Bayesian information criteria (BIC; Schwartz, 1978), and the sample size adjusted BIC (ABIC; Sclove, 1987). Chi-square test is affected by sample size (i.e., small discrepancy is easier to detect in larger sample samples) and tends to overfactoring (i.e., extract too many factors; Mulaik et al., 1989; Schmitt, 2011). But Chi-square test is the only method that is based on distributional theory and allows for a formal significance test of overall model fit and difference for model comparison (e.g., Chi-square difference test between two nested models; Schermelleh-Engel, Moosbrugger, & Müller, 2003; Yuan & Bentler, 2008)[2]. It is recommended that the Chi-square test be used together with the SRMR when evaluating improvement in fit resulting from inclusion of additional factors (Schmitt, 2011).

In assessing model fit, the authors follow a set of cutoff criteria researchers have recommended as described as follows: values smaller than .08/.06 for RMSEA indicate acceptable/good model fit, values greater than .90/.95 for CFI and TLI indicates acceptable/good model fit (Mulaik et al., 1989; Sharma, Mukherjee, Kumar, & Dillon, 2005), and values smaller than .10/.05 for SRMR indicates acceptable/good model fit (Schermelleh-Engel et al., 2003). In comparing nested models, along with the Chi-square difference test, the change in RMSEA, CFI, AIC, BIC, and ABIC will also be examined as supplemental support. Lower values for RMSEA, AIC, BIC, ABIC reflect better fit.

Common one-factor model EFA. By requesting the type of analysis to be EFA and specifying the range of the number of factors, Mplus produces output for each model with a given number of factors in the range. In this analysis, the authors specify the range of the number of factors to be from 1 to 5. The oblique GEOMIN (Yates, 1987) rotation is used to obtain factor loadings and factor correlations. The residuals are not correlated. Modification indices are requested.

Correlated-factor model ESEM. The correlated-factor model EFA is conducted within the ESEM framework. The ESEM is recently developed by Asparouhov and Muthén (2009) and (2010) implemented in Mplus program. In estimating ESEM, a priori hypothesis regarding the expected factor structure and a target rotation is used (Asparouhov & Muthén, 2009; Marsh, Morin, Parker, & Kaur, 2014). In this analysis, four factors are specified to load on all the items. The four factors are allowed to be correlated. The residuals are not correlated.

Second-order factor model ESEM. The second-order factor model EFA is conducted within the ESEM framework. In this analysis, a target rotation is used. The second-order factor is specified to load on all the first-order factors, the first-order factors are specified to load on all the observed variables with factors loadings estimated from the multiple-factor model established from preceded common factor analysis as starting values. Residual correlations between first-order factors are not allowed. Four factor loadings for each first-order factor are fixed at given values for model identification purpose. The residuals are not correlated.

Bifactor model ESEM. Bifactor exploratory factor analysis is also conducted in the framework of ESEM. The bi-GEOMIN (ORTHOGONAL) rotation is used to obtain factor loadings. By specifying that a general factor and the m group factors (i.e., four in this study) are indicated by the observed variables, Mplus will produce results of a bifactor model with one general factor and m group factors. The residuals are not correlated.

CFA

In specifying the CFA models, the structures identified from exploratory analysis are used as a priori hypotheses. For model identification purpose, one item is selected for each latent factor and its factor loading is fixed at 1 (Vandenberg & Lance, 2000). The MLR method is used to obtain parameter estimates and model fit indexes that are robust to non-normality. The same criteria established for evaluating model fit and comparing nested models for EFA is used for CFA.

Common one-factor model CFA. In this model, one single factor is specified to load on all the observed variables. The residuals are uncorrelated and freely estimated as by default in Mplus. One-factor loading is fixed at 1 for model identification purpose.

Correlated four-factor model CFA. In this model, four factors are specified to load on the intended observed variables. The factors are specified to be correlated. The residuals are uncorrelated and freely estimated. One-factor loading for each latent factor is fixed at 1 for model identification purpose.

Second-order factor model CFA. In this mode, four first-order factors are specified to load on the intended observed variables, and a second-order factor is specified to load on the first-order factors. The residuals of the observed variables are uncorrelated and freely estimated. The residuals of the first-order factors are uncorrelated and freely estimated as by default in Mplus. One-factor loading for each first-order factor is fixed at one and one of the factor loadings for second-order factor is fixed at one for model identification purpose.

Bifactor model CFA. In this model, a general factor is specified to load on all the observed variables, and four domain-specific factors are specified to load on the intended observed variables. The general factor and the domain-specific factors are specified to be uncorrelated. The domain-specific factors are specified to be uncorrelated. The residuals of the observed variables are uncorrelated and freely estimated. One-factor loading for the general factor is fixed at one and onefactor loading for each domain-specific factor is fixed at one for model identification purpose.

CFA with a method factor. Half of the items in the PSPP are positively worded and half of the items are negatively worded. The positively worded items were reverse coded before the analysis. A *method* factor is then introduced to each of the four CFA models to account for the communities shared by the negatively worded items. In all the four models, the method factor is specified to be uncorrelated with other latent factors. One-factor loading for the method factor is fixed at one for model identification purpose.

Bifactor model with gender as a covariate. To investigate gender difference in both the general factor and domain specific factors, the authors then fit a bifactor model with gender included as a predictor on both the general factor and the four domain specific factors to the sample. The base model is specified in the same way as in the bifactor model.

Results

Exploratory factor analysis

The mean, standard deviation, and correlations of the observed variables are presented in Table 1. The goodness-of-fit indices associated with the ESEM models are presented in Table 2 (top section). The one-factor EFA model (CFI = .600; TLI = .562; RMSEA = .133; SRMR = .116) provides an unacceptable degree of fit to the data, suggesting that the one-factor model is not a good representation of the data. The Kaiser eigenvalues criterion suggests a four-factor model, with four eigenvalues are greater than 1 (i.e., 9.546, 2.795, 1.865, and 1.635). The analysis reveals that the four factors explain 66% of the total variance among the items, with the first factor explaining 39.8% of the total variance. The scree test also indicates a four-factor model, with a first elbow happens after the first factor and a second elbow happens after the fourth factor. The simple structure criterion also suggests a four-factor model; the standard factor loadings from the correlated four-factor model are presented in Table 3. As indicted by the pattern of loadings, a simple structure with four factors loaded on each of the six targeted items is supported. The four-factor EFA model (CFI = .956; TLI = .935; RMSEA = .051; SRMR = .028) provides good fit to the data. A correlated four-factor model is then selected among the common factor models as the best model.

The correlated four-factor ESEM model (CFI = .956; TLI = .935; RMSEA = .051; SRMR = .028) provides acceptable model fit. The second-order ESEM model (CFI = .956; TLI = .936; RMSEA = .051; SRMR = .030) provides acceptable model fit, the bifactor ESEM model (CFI = .968; TLI = .948; RMSEA = .046; SRMR = .022) provides good model fit, suggesting that all three models can be viewed as good representations of the data. The bifactor model is the best one, however supplemented by the fact that the smallest values for information criteria (AIC = 12024; BIC = 12581; ABIC = 12080) are associated with the bifactor model. The Chi-square difference test ($\chi^2 = 73.0774, df = 20, p < 0.001$) indicates that the bifactor model is the best-fitted model as well.

Standardized factor loadings of second-order ESEM model and bifactor ESEM model are presented

Table 1. Mean, *SD*, and correlations of the observed variables (*N* = 400).

	Q1	Q2	Q3	Q4	Q6	Q7	Q8	Q9	Q11	Q12	Q13	Q14	Q16	Q17	Q18	Q19	Q21	Q22	Q23	Q24	Q26	Q27	Q28	Q29
Q1	1.00																							
Q2	.42	1.00																						
Q3	.27	.44	1.00																					
Q4	.35	.38	.25	1.00																				
Q6	.53	.46	.42	.59	1.00																			
Q7	.36	.50	.37	.38	.50	1.00																		
Q8	.17	.47	.56	.09	.29	.27	1.00																	
Q9	.32	.35	.26	.79	.54	.37	.19	1.00																
Q11	.51	.53	.30	.32	.49	.34	.37	.31	1.00															
Q12	.41	.57	.35	.35	.58	.56	.40	.30	.49	1.00														
Q13	.26	.44	.55	.20	.32	.32	.56	.20	.37	.33	1.00													
Q14	.38	.39	.29	.62	.58	.30	.19	.58	.44	.43	.30	1.00												
Q16	.48	.40	.37	.42	.69	.39	.31	.43	.69	.60	.34	.53	1.00											
Q17	.32	.58	.35	.39	.42	.55	.39	.34	.45	.56	.49	.50	.50	1.00										
Q18	.21	.44	.68	.34	.45	.39	.54	.35	.33	.39	.50	.32	.36	.34	1.00									
Q19	.40	.52	.37	.54	.43	.33	.36	.57	.46	.52	.35	.58	.40	.48	.39	1.00								
Q21	.46	.41	.26	.30	.43	.31	.27	.26	.51	.50	.26	.33	.59	.37	.23	.35	1.00							
Q22	.33	.57	.39	.42	.55	.78	.38	.40	.41	.60	.38	.42	.49	.65	.41	.40	.32	1.00						
Q23	.35	.57	.68	.28	.51	.46	.68	.39	.44	.53	.55	.30	.51	.50	.66	.45	.35	.55	1.00					
Q24	.35	.49	.35	.63	.56	.40	.34	.72	.43	.51	.26	.61	.54	.51	.36	.68	.41	.50	.52	1.00				
Q26	.42	.44	.35	.36	.48	.38	.41	.30	.66	.53	.32	.47	.64	.53	.31	.40	.48	.46	.47	.39	1.00			
Q27	.45	.64	.49	.48	.65	.67	.40	.41	.52	.71	.43	.47	.61	.60	.49	.52	.46	.71	.60	.51	.60	1.00		
Q28	.27	.46	.69	.19	.45	.31	.60	.25	.39	.34	.71	.34	.39	.39	.58	.38	.24	.39	.64	.36	.29	.47	1.00	
Q29	.34	.46	.26	.52	.48	.37	.34	.51	.52	.46	.28	.65	.47	.51	.29	.56	.40	.44	.39	.61	.46	.45	.29	1.00
Mean	2.71	2.76	2.66	2.75	2.73	3.31	2.54	2.65	3.10	2.98	2.71	2.68	2.82	3.10	2.52	2.71	3.08	3.15	2.54	2.68	3.06	2.95	2.54	2.72
SD	.86	.96	.84	.82	.84	.88	.89	.87	.93	.88	1.04	.82	.86	.80	.83	.82	.85	.84	.89	.81	.93	.89	.93	.80

Table 2. Model fit of ESEM and CFA models.

Models	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90%CI	SRMR	AIC	BIC	ABIC
ESEM (*n* = 250)										
One-factor[a]	1,373	252	.600	.562	.133	.127–.140	.116	13,239	13,493	13,264
Correlated four-factor	308	186	.956	.935	.051	.041–.061	.028	12,076	12,562	12,124
Second-order	309	188	.957	.936	.051	.040–.061	.028	12,072	12,551	12,120
Bi-factor	254	166	.968	.948	.046	.034–.057	.022	12,024	12,581	12,080
CFA (*n* = 150)										
One-factor	921	252	.678	.647	.133	.124–.142	.097	7,626	7,843	7,615
Correlated four-factor	474	246	.890	.877	.079	.068–.089	.061	7,112	7,347	7,100
Second-order	477	248	.890	.877	.078	.068–.089	.063	7,112	7,341	7,100
Bi-factor	392	228	.921	.904	.069	.058–.081	.048	7,045	7,334	7,030
CFA with method factor (*n* = 150)										
One-factor	787	240	.737	.697	.123	.114–.133	.090	7,483	7,736	7,470
Correlated four-factor	392	234	.924	.910	.067	.055–.079	.062	7,034	7,305	7,020
Second-order	396	236	.923	.910	.067	.056–.079	.064	7,035	7,300	7,021
Bi-factor	324	216	.948	.934	.058	.044–.070	.042	6,989	7,314	6,973

[a]One-factor exploratory analysis is conducted in traditional EFA framework.

in Table 3. The factor loadings for the second-order factor ESEM model were obtained with that the factor loadings from the correlated four-factor model being given as starting values. The loadings of the second-order factor on first-order factors were freely estimated and are .54, .82, .60, and .46 as presented in Table 3. The pattern of factor loadings for the bifactor model shows a clear simple structure of four-factor model.

CFA

The goodness-of-fit indices associated with the CFA models are presented in Table 2 (middle section). The one-factor CFA model (CFI = .678; TLI = .674; RMSEA = .133; SRMR = .116) provides an unacceptable degree of fit to the data, suggesting that the one-factor model is not a good representation of the data. The correlated four-factor CFA model (CFI = .890; TLI = .877; RMSEA = .079; SRMR = .061) and the second-order CFA factor model (CFI = .890; TLI = .877; RMSEA = .078; SRMR = .061) provide unacceptable model fit, whereas the bifactor CFA model (CFI = .921; TLI = .904; RMSEA = .069; SRMR = .048) provides acceptable model fit.

Model fit of the models with a *method* factor is presented in Table 3 (bottom section). The one-factor CFA model (CFI = .737; TLI = .697; RMSEA = .123; SRMR = .090) provides an unacceptable degree of fit to the data, suggesting that the one-factor model is not a good representation of the data. The correlated four-factor CFA model (CFI = .924; TLI = .910; RMSEA = .067; SRMR = .062), the second-order factor model (CFI = .923; TLI = .910; RMSEA = .067; SRMR = .064), and the bifactor model (CFI = .948;

Table 3. Standardized factor loadings of ESEM models.

	Correlated four-factor ESEM					Second-order ESEM					Bifactor ESEM				
N = 250	F1	F2	F3	F4		F1	F2	F3	F4		G	F1	F2	F3	F4
Q1	.67	.08	−.04	.01	Q1	.67	.08	−.04	.01	Q1	.45	.53	.03	−.04	.02
Q6	.59	.13	−.01	.21	Q6	.63	.13	−.01	.20	Q6	.58	.50	.09	.01	.17
Q11	.83	−.03	.07	−.03	Q11	.82	−.03	.07	−.03	Q11	.55	.60	−.12	−.03	−.09
Q16	.69	−.07	.08	.20	Q16	.74	−.06	.07	.19	Q16	.57	.55	−.07	.02	.14
Q21	.75	−.01	−.02	0	Q21	.75	−.01	−.02	.00	Q21	.45	.58	−.05	−.05	−.01
Q26	.77	.11	−.07	.01	Q26	.77	.10	−.07	.01	Q26	.52	.61	.03	−.07	.01
Q2	.09	.54	.15	.06	Q2	.10	.54	.14	.08	Q2	.64	.02	.29	.09	−.08
Q7	−.04	.84	−.23	.09	Q7	−.05	.81	−.22	.13	Q7	.53	−.04	.53	−.12	.03
Q12	.15	.71	.03	−.17	Q12	.10	.71	.03	−.13	Q12	.63	.00	.32	−.06	−.32
Q17	.13	.54	.09	.05	Q17	.13	.54	.09	.07	Q17	.67	.00	.22	−.01	−.14
Q22	−.10	.86	.05	−.06	Q22	−.12	.86	.05	−.02	Q22	.57	−.06	.58	.11	−.11
Q27	.06	.74	.00	.11	Q27	.06	.74	.00	.14	Q27	.65	.08	.54	.08	.06
Q3	−.09	−.05	.81	.07	Q3	−.05	.03	.76	.05	Q3	.41	.02	.11	.71	.03
Q8	.02	.05	.74	−.14	Q8	.03	.12	.70	−.14	Q8	.50	−.05	−.02	.49	−.30
Q13	.06	.10	.71	−.03	Q13	.06	.17	.67	−.04	Q13	.63	−.06	−.04	.43	−.26
Q18	−.15	.06	.73	.11	Q18	−.11	.13	.69	.09	Q18	.44	−.05	.17	.64	.04
Q23	−.01	.10	.73	−.06	Q23	−.01	.17	.69	−.06	Q23	.57	−.08	−.01	.48	−.25
Q28	.08	−.06	.81	.09	Q28	.11	.02	.76	.07	Q28	.58	.07	.00	.59	−.05
Q4	−.06	−.02	−.04	.87	Q4	.11	−.02	−.04	.81	Q4	.53	−.05	.00	−.02	.64
Q9	−.03	.06	−.04	.83	Q9	.14	.06	−.04	.78	Q9	.64	−.09	−.04	−.10	.52
Q14	.14	−.06	.01	.62	Q14	.27	−.06	.01	.58	Q14	.55	.03	−.15	−.09	.36
Q19	.21	.04	.18	.57	Q19	.33	.06	.17	.53	Q19	.74	.09	−.07	.05	.29
Q24	.10	.17	.12	.61	Q24	.22	.19	.12	.57	Q24	.73	.05	.08	.06	.36
Q29	.28	.09	.07	.40	Q29	.36	.10	.07	.37	Q29	.70	.07	−.15	−.10	.10
F2	.43				G	.54	.82	.60	.46						
F3	.37	.57													
F4	.43	.47	.34												

G = General Factor.

TLI = .934; RMSEA = .058; SRMR = .042) provide acceptable model fit, indicating that all three models fit the data well, and the bifactor model is the best-fitted model.

The standardized factor loadings of CFA models with *method* factor included are presented in Table 4. As indicated in the table, the factor loadings from the correlated-four factor model and second-order factor model are almost the same. The factor loadings of the observed variables on the four latent factors range from .61 to .86, suggesting that the items are good indicators of the intended factors. In the correlated-factor model, the correlation coefficients range from .34 to .57, suggesting that the four factors are moderately correlated. In the second-order factor model, the loadings of first-order factors on the second-order factor range from .71 to .90, indicating that the second-order factor accounts for the covariation among the four-first factors.

As presented in Table 4, in the bifactor model, the factor loadings of the observed variables on the general factor range from .46 to .80, indicating that all the items are good indicators of the general factor; the factor loadings of the observed variables on the domain specific factors range from .25 to .68, except that three items have loadings smaller than .20 (i.e., .13 for Q6, .19 for Q2, and .17 for Q12), indicating that the domain specific factors are viable after the general factor has been partialled out. Most of the loadings on the general factor are greater than the loadings on the domain specific factors. The items (e.g., Q1, Q6, Q11) with greater loadings on the general factor are referred to as marker items of general factor and those items (e.g., Q3, Q7, Q8) with loadings on the domain specific factors are referred to as marker items of domain specific factors.

Coefficient omega hierarchical (ω_h)

The coefficient Omega hierarchical (ω_h) is computed for both the general factor and specific factors[1]. ω_h of the general factor for the total score is .86. ω_s of the specific factor for the sports competence (SC) subscale score is .19, for the physical condition (PC) subscale is .20, for the body attractiveness (BA) subscale score is .59, and for the physical strength (PS) subscale score is .29. The Omega reliability (ω) for the total score is .96, for the SC subscale score is .90, for the PC subscale is .95, for the BA subscale score is .95, and for the PS subscale score is .91.

Gender as a covariate

To examine the role of gender in both the general factor and domain specific factors, the bifactor CFA model with method factor is used as a baseline model. Results from the bifactor model with gender included as a predictor suggest that male students scored higher on the general *physical self-esteem* factor than female students ($\beta = -.742$, $p < .05$), female students scored higher on the domain

Table 4. Standardized factor loadings of CFA with method factor models.

	Four-factor CFA model					Second-order CFA model					Bifactor CFA model				
n = 150	SC	PC	BA	PS		SC	PC	BA	PS		GPSE	SC	PC	BA	PS
Q1	.61				Q1	.61				Q1	.54	.25			
Q6	.76				Q6	.76				Q6	.70	.13			
Q11	.79				Q11	.79				Q11	.67	.48			
Q16	.86				Q16	.86				Q16	.72	.45			
Q21	.65				Q21	.65				Q21	.57	.33			
Q26	.75				Q26	.75				Q26	.68	.36			
Q2		.74			Q2		.74			Q2	.72		.19		
Q7		.74			Q7		.74			Q7	.54		.65		
Q12		.79			Q12		.80			Q12	.78		.17		
Q17		.76			Q17		.76			Q17	.70		.31		
Q22		.81			Q22		.81			Q22	.66		.56		
Q27		.86			Q27		.86			Q27	.80		.26		
Q3			.81		Q3			.81		Q3	.51			.64	
Q8			.76		Q8			.76		Q8	.53			.55	
Q13			.73		Q13			.73		Q13	.50			.54	
Q18			.76		Q18			.76		Q18	.50			.54	
Q23			.84		Q23			.84		Q23	.68			.49	
Q28			.81		Q28			.81		Q28	.53			.63	
Q4				.73	Q4				.72	Q4	.49				.56
Q9				.75	Q9				.74	Q9	.46				.68
Q14				.77	Q14				.77	Q14	.62				.41
Q19				.79	Q19				.79	Q19	.66				.44
Q24				.83	Q24				.83	Q24	.67				.50
Q29				.76	Q29				.76	Q29	.64				.37
PC	.80				GPSE	.90	.90	.71	.78						
BA	.61	.69													
PS	.74	.68	.52												

GPSE = Global Physical Self-Esteem.

specific factor *physical condition* than male students (β = .715, p < .05), and no gender difference was found on the other three domain specific factors sport competence (β = − .374, p > .05), *body attractiveness* (β = − .163, p > .05), and *physical strength* (β = .214, p > .05).

Results Summary

Exploratory factor analyses are conducted with a randomly selected subsample and CFA are conducted with the remaining sample. Exploratory common factor analysis are conducted in a traditional way, whereas exploratory correlated-factor model, second-order model, and bifactor model analyses are conducted in the framework of ESEM (Asparouhov & Muthén, 2009, 2010; Marsh et al., 2014). All the analyses are conducted with Mplus 7.0 program, in which target rotation is used as the embedded rotation method for ESEM. As suggested by the results, overall, the ESEM models have better model fit than their CFA models. Bifactor model is the best-fitted model from both EFA and CFA analyses in comparison to the three alternatives: common one-factor model, correlated-four factor model, and second-order factor model. Beyond the fact that it best represents the sample data, bifactor model is also superior to its competitors by allowing separation of the general factor from the domain specific factors, and allowing investigation of the relationship of external factor (i.e., gender as in this study) with general factor and domain specific factors separately and simultaneously. Gender difference in PSPP is found to exist at both the general factor level and one of the domain specific factors level.

Discussion

Physical self-perception is a multi-dimensional concept defined as a sense of competence in multiple physical domains (Fox & Corbin, 1989; Harter, 1999). Fox and Corbin (1989) developed the PSPP to examine self-perception in the physical domains. The PSPP was designed to represent the four physical domains sport competence, physical strength, physical condition, and body attractiveness to indicate a general physical self-esteem construct. To clarify, initially, PSPP included six items to measure a general physical self-worth concept. This general physical self-worth concept was originally used as a criterion variable to test the convergent validity of the PSPP in Fox and Corbin's (1989) work. By examining the relationships between the general physical self-worth and the four specific domain factors, the authors claimed that the PSPP was measuring a general physical self-esteem concept, which is related to, but distinct from, physical self-worth. The inclusion of the general physical self-worth factor in the initial PSPP seems to have misled followers who treated it as a fifth physical domain and took it as an essential component of PSPP (Hagger et al., 2004). In the current study, only the 24 items that are designed to measure physical self-perception are used in the analyses for the purpose of the study.

To acknowledge the dimensionality of PSPP, Fox and Corbin (1989) initially adopted the correlated four-factor model in their analysis. In the model, the four factors were loaded on their intended items and allowed to be correlated with each other. However, this model could not recover the general physical self-esteem factor that was designed to be measured by the PSPP. Recognizing the deficiency of the correlated four-factor model in reflecting the general physical self-esteem factor, Hagger and colleagues (2003) used a second-order factor model to analyze PSPP data (Hagger et al., 2003). In the second-order factor model, the general physical self-esteem factor was included as a higher-order factor to explain the covariance shared by the four specific domain factors (i.e., called first-order factors in the second-order factor model). Both the correlated factor analysis and the second-order factor analysis have been widely applied in handling multi-dimensionality of scales in other areas (Digman, 1997, 1990; Judge & Bono, 2000; Marsh & Hocevar, 1985; McCrae & Costa, 1987; McCrae & John, 1992; Saklofske, Austin, & Minski, 2003; Tellegen, Watson, & Clark, 1999). Until recently, the bifactor model, as an alternative, is re-proposed to be a superior model to analyze multi-dimensionality of scales (Holzinger & Swineford, 1937, 1939; Reise, 2012; Reise et al., 2007, 2010). Both second-order factor model and bifactor model assume that there exists a general factor that is over and above domain specific factors, and this assumption has recently been called into question. But recently, Revelle and Wilt (2013) challenged Musek's (2007) approach as problematic in identifying the general big one-factor using principal component analysis to extract factors and in viewing the first factor as the general factor (please see Revelle & Wilt, 2013 for detailed information).

Instead, Revelle and Wilt (2013) suggest using McDonald's (1999) approach coefficient Omega hierarchial (ω_h) to evaluate the importance of a general factor. Coefficient Omega hierarchial (ω_h) is the ratio of squared sum of factor loadings on the general factor to the modeled variance of scale scores (Reise, 2012; Revelle & Wilt, 2013). Conceptually, ω_h can be viewed as an index to indicate the proportion of variance of scale that is explained by the general factor. Reise (2012) also proposed that omega subscale (ω_s) can be used as a model-based reliability index of a subscale to evaluate the importance (or viability) of a subscale after the general factor has been controlled. ω_s can be computed by dividing sum of the group factor loadings squared by the the sum of variance of subscale scores (Reise, 2012). In addition, two other approaches have been used in evaluting the importance of a general factor but are not without criticism. The first approach relies on model fit comparion and the other is the index common variance explained (ECV) (Reise, 2012; Revelle & Wilt, 2013). As a matter of fact, a model including a general factor would fit better than a model excluding it, because the latter is nested within the former one (Chen et al., 2006). Thus improvement in model fit does not necessarily mean there is a need for a general factor. The ECV is the ratio of variance explained by the general factor to the variance explained by the general plus the group factors. However, the ECV is valued as "an index of unidimensionalty but not an index of importance of a general factor, because ECV only speaks of how strong the general factor is relative to group factors." A high ECV can be obtained when the general factor is weak and group factors are even weaker (Reise, 2012).

In this study, in proposing the use of bifactor model to study PSPP, the authors heavily rely on the theoretical background of how the instrument PSPP was created. Besides theory, the authors provide statitistical evidence to support their proposal. Theoretically, the PSPP derived from a hierachical model of self-perception profiles with a general factor at the higher level of the hierarchy (Fox & Corbin, 1989). Statistically, the use of bifactor model to analyze PSPP data is justified in multiple ways: (a) the large ratio of first eigenvalue value to second eigenvalue value (i.e., 9.546 to 2.795), which is indicative of a general factor, (b) better model fit compared to its competitors from both ESEM analysis and confirmatory facto analysis, and (c) a ω_h value of .86 of the general factor for the total score from the bifactor CFA model, which indicates the existence of a general factor.

In this study, the authors did the exploratory factor analysis in the ESEM framework. ESEM has recently been implemented in Mplus program (Asparouhov & Muthén, 2009, 2010; Marsh et al., 2014). Many researchers have already taken advantage of this latest development and applied ESEM in their research (Booth & Hughes, 2014; Dombrowski, 2014; Ebesutani et al., 2012; Jennrich & Bentler, 2011, 2012; Marsh et al., 2009, 2010; Reise et al., 2007).

ESEM is expected to provide more accurate parameter estimates than the over-restricted CFA model while being more parsimonious than the EFA model (Asparouhov & Muthén, 2009). However, ESEM has a few limitations, as pointed out by Marsh and colleagues (2014) and Asparouhov and Muthén (2009), in that it is vulnerable to "rotational indeterminacy" (i.e., "different rotation strategies result in different solutions that all fit the data equally well") as is the traditional EFA and it loses parsimony when there is a large number of indicators. As suggested, ESEM should be used as a baseline model to conduct subsquent analysis if "it fits the data better than a correponding CFA model" (Marsh et al., 2014).

In this study, the authors conducted exploratory second-order analysis and exploratory bifactor analysis in the framework of ESEM under target rotation. Target rotation is conceptually viewed as an approach fitting in between "EFA rotation and the CFA specification" (Marsh et al., 2014), in which the target loading values are zeros or near zeros to indicate a priori model. The exploratory bifactor analysis is well supported by Mplus program whereas the exploratory second-order factor analysis is not fully established with Mplus program. In conducting the second-order analysis in the ESEM framework, the authors follow the approach used in the latest work by Morin, Morin, Arens, and Marsh (2015; referred to therein as hierarchical-ESEM). One thing worth articulating that was not mentioned in their work is that a number of m^2 (m is the number of first-order factors; $m = 4$ in the authors' sample) constraints has to be forced to achieve model identification in fitting second-order ESEM under target rotation (Marsh et al., 2014). Asparouhov and Muthén (2009) suggested that $m \times (m - 1)$ constraints are sufficient to achieve model identification. However, with the authors' sample, to identify the model, at least 16 (i.e., m^2) zero targets have to be fixed at given values.

In the analysis, the authors add a method factor to the models to account for the fact that half of the items in PSPP are negatively worded. Including the method factor has substantially improved the model fit, and again this should not be surprising. The wording effect has been reported in many studies (Hazlett-Stevens, Ullman, & Craske, 2004; Horan, DiStefano, & Motl, 2003; Motl & DiStefano, 2002), but has not yet received enough attention. Ignoring *method* factor can lead to inaccurate interpretation of factor analysis results and hinder the understanding of a construct, especially when the concept is one-dimensional (Schriesheim & Eisenbach, 1995; Tomas & Oliver, 1999). In studies of self-esteem, there has been inconsistency in the reports on the dimensionality of the self-esteem concept. As reported, factor analysis yields a two-factor solution if the method factor was not considered. However, a solution with a general factor on all items and a method factor on the negatively worded items is preferred when the method factor is taken into account (Horan et al., 2003; Tomas & Oliver, 1999). The method factor should not be misunderstood as a conceptual factor and not be associated with external related variables (Hazlett-Stevens et al., 2004).

The use of bifactor analysis to study PSPP helps to advance the understanding of the physical self-perception concept. In the bifactor analysis framework, the general physical self-esteem factor is separated from the domain-specific factors. Its relationship with external variables can be examined, and at the same time, the relationships between the domain-specific factors and external variables can be examined while controlling each other. As indicated by this study, on average, male students have higher general physical self-perception scores than female students. In addition, female students scored higher on the domain specific factor physical condition than male students whereas no gender difference was found on the other three domain factors, which are interesting findings. In previous studies, gender difference on PSPP was examined either by total score approach (i.e., a composite score of the 24 items) or individual score approach (i.e., a composite score of each subscale) (Chen et al., 2012), both of which provide less accurate estimates than the latent score approach (i.e., the bifactor approach in the current study). Furthermore, the general factor and specific factors could not be studied separately in previous studies.

In comparing the three methods, the bifactor model emerges as the most useful for examining complex inter-relations. The bifactor model allows applied researchers to directly examine relationships which the second-order factor model does not support: (1) the relationship between the general factor and observed items and (2) between the specific factors and observed items. The majority of current PSPP-related research addresses findings of the second-order model. The bifactor model allows separation of the domain specific factors from the general factor. This is significant because the bifactor model allows simultaneous examination of the relationship of an external variable (i.e., gender in this case) to a general factor and specific factors. Though findings from the second-order approach are valuable, using bifactor model allows PSPP researchers to begin to re-examine the concept, which could lead to a deeper understanding about hierarchical concepts.

Building on previous models, this approach transforms the field's abilities to understand connections between factors. This research verified that domain factors still exist after partialling out the general factor. This result reveals a more in-depth connection between items and domain factors which may not be apparent when using either the correlated factor model or the second-order factor model. These two models are useful; however, this model illuminates a new approach to examining the PSPP, previously only addressed using the second-order factor model. Further, though this model was tested using participants from the United States, it should extend the understanding of the PSPP when applied in other cultural settings, leading to a better understanding of participant characteristics. Additionally, this approach can aid in determining whether gender is a predictor of the PSPP domains within those different cultural contexts.

Limitations and future directions

A few limitations are worth mentioning regarding the use of bifactor modeling. First, bifactor model hypothesizes that there is a general factor along with a number of domain specific factors in measuring a concept. When there is no general factor presented, it is not appropriate to use bifactor modeling. This study provides both theoretical and statistical criteria on evaluating the importance of a general factor. Additionally, the study offered evidence to justify the use of bifactor model in studying the general factor measured by PSPP. Furthermore, statistically, bifactor models seem to always provide better model fit than the alternative models (Yung, Thissen, & McLeod, 1999). Hence, bifactor modeling should be used cautiously. However, the statistical criteria this study has relied on have been criticized for limitations in evaluating the importance of a general factor. For example, there is no guidance on the cutoff point of coefficient omega hierarchical to indicate whether the general factor is important enough to be included in the model. The eigenvalue criterion and the model fit comparison criterion have received criticism as well. Simulation studies are needed to evaluate the performance of different criteria on evaluating the importance of a general factor.

The authors have a critical decision to make in choosing between ESEM and CFA model as an appropriate baseline model to conduct subsequent analysis. The bifactor ESEM and CFA model both provide acceptable model fit, with the ESEM providing a slightly better fit but less parsimony. The authors believe the improvement in fit is worth the loss in parsimony. Recent work on ESEM claimed that ESEM provides more accurate estimates, but yet no convincing conclusion is available to make an appropriate selection between ESEM and CFA. Simulation studies are needed to evaluate the ESEM in comparison to CFA.

Notes

1 Though there are many factor-extracting methods available, the authors use the robust continuous MLR as a method to extract factors in the exploratory factor analyses. Two alternatives are robust least squares (LS) estimation and robust weighted LS (WLS) estimation, with that robust LS and WLS use polychoric correlations whereas MLR uses standard Pearson correlations. All three can result in accurate parameter estimates and test statistics and all three rely on adjustments to the Chi-square test statistic (χ^2; Bentler & Yuan, 1999; Schmitt, 2011).

2 The Chi-square values resulting from MLR method cannot be used for a difference test in a regular way because the difference of Chi-square values from two nested models does not follow a Chi-square distribution (Schermelleh-Engel et?al., 2003). In comparing two nested models, the test of fit and its correction factor are used in computing the test statistic: $= \frac{(F_0c_0 - F_1c_1)(d_0 - d_1)}{c_0d_0 - c_1d_1}$, where F_1 and F_0F0 are robust tests of fit for models M_1M1 and M_0M0 (more restricted model), c_1 and c_0 are the correction factors for F_1 and F_0, d_1 and d_0 are the degrees of freedom for F_1 and F_0. (Asparouhov & Muthén, 2010; Satorra & Bentler, 2001).

3 McDonald (1999) has proposed coefficient omega as an estimate of the general factor saturation of of a test (McDonald, 1999; Zinbarg, Yovel, Revelle, & McDonald, 2006). The coefficient Omega hierarchical (ω_h) can be computed for both the general factor and the specific factors based on the loadings obtained from the bifactor CFA model. h_ω of the general factor for the total score is calculated using the following formula: $\omega_h = \frac{(\Sigma\lambda_{iGEN})^2}{(\Sigma\lambda_{iGEN})^2 + (\Sigma\lambda_{iGRP_1})^2 + (\Sigma\lambda_{iGRP_2})^2 + (\Sigma\lambda_{iGRP_3})^2 \ldots\ldots + (\Sigma\lambda_{iGRP_p})^2 +}$ $(\Sigma\lambda_{iMETHOD})^2 + \Sigma\theta_i^2$, where λ_{iGEN} indicates the standardized loadings of general factor, λ_{iGRP_p} indicates the standardized loadings of specific factors, and θ_i^2 indicates the squared residuals. The ω_s of the specific factor for the subscale score is obtained by the following formula: $\omega_s = \frac{(\Sigma\lambda_{iGRP_p})^2}{(\Sigma\lambda_{iGEN})^2 + (\Sigma\lambda_{iGRP_p})^2 + (\Sigma\lambda_{iMETHOD})^2 + \Sigma\theta_i^2}$. The Omega hierarchical (ω) can also be used as a model-based reliability index and can be compared with Cronbach's *a* (Reise, 2012). The ω reliability for the total score is computed by $\omega = \frac{(\sum\lambda_{iGEN})^2 + (\sum\lambda_{iGRP_1})^2 + (\sum\lambda_{iGRP_2})^2 + (\sum\lambda_{iGRP_3})^2 \ldots\ldots + (\sum\lambda_{iGRP_p})^2}{(\sum\lambda_{iGEN})^2 + (\sum\lambda_{GRP_1})^2 + (\sum\lambda_{GRP_2})^2 + (\sum\lambda_{iGRP_3})^2 \ldots\ldots (\sum\lambda_{iGRPp})^2 +}$ $(\sum\lambda_{iMETHOD})^2 + \sum\theta_i^2$ and for the subscale score is computed by $\omega = \frac{(\Sigma\lambda_{iGEN})^2 + (\Sigma\lambda_{iGRP_p})^2}{(\Sigma\lambda_{iGEN})^2 + (\Sigma\lambda_{iGRP_p})^2 + (\Sigma\lambda_{iMETHOD})^2 + \Sigma\theta_i^2}$.

References

Acton, G. S., & Schroeder, D. H. (2001). Sensory discrimination as related to general intelligence. *Intelligence*, *29*(3), 263–271. doi:10.1016/S0160-2896(01)00066-6

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317–332.

Armon, G., & Shirom, A. (2011). The across-time associations of the Five-Factor Model of Personality with vigor and its facets using the bifactor model. *Journal of Personality Assessment*, *93*(6), 618–627. doi:10.1080/00223891.2011.608753

Asçi, F., Asçi, A., & Zorba, E. (1999). Cross-cultural validity and reliability of Physical Self-Perception Profile. *International Journal of Sport Psychology*, *30*(3), 399–406.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. doi:10.1080/10705510903008204

Asparouhov, T., & Muthén, B. (2010). Computing the strictly positive Satorra-Bentler chi-square test in Mplus. *Mplus Web Notes*, *12*. Retrieved from http://www.statmodel.com/examples/webnotes/SB5.pdf

Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin*, *107*(2), 238–246.

Bentler, P. M., & Yuan, K. H. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, *34*, 181–197.

Booth, T., & Hughes, D. J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, *21*(3), 260–271. doi:10.1177/1073191114528029

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219–251. doi:10.1111/jopy.2012.80.issue-1

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189–225. doi:10.1207/s15327906mbr4102_5

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*(1), 417–440. doi:10.1146/annurev.ps.41.020190.002221

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*(6), 1246–1256. doi:10.1037/0022-3514.73.6.1246

Dombrowski, S. C. (2014). Investigating the structure of the WJ-III Cognitive in early school age through two exploratory bifactor analysis procedures. *Journal of Psychoeducational Assessment*, *32*(6), 483–494. doi:10.1177/0734282914530838

Ebesutani, C., Reise, S. P., Chorpita, B. F., Ale, C., Regan, J., Young, J., & Weisz, J. R. (2012). The Revised Child Anxiety and Depression Scale-Short Version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor. *Psychological Assessment*, *24*(4), 833–845. doi:10.1037/a0027283

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. doi:10.1037/1082-989X.4.3.272

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport & Exercise Psychology*, *11*, 408–430.

Gault, U. (1954). Factorial patterns of the Wechsler intelligence scales. *Australian Journal of Psychology*, *6*(1), 85–89. doi:10.1080/00049535408256079

Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, *48*(5), 639–662. doi:10.1080/00273171.2013.804398

Golay, P. (2011). *Orthogonal higher order structure and confirmatory factor analysis of the French Wechsler Intelligence Scales*. Paper presented at the 12th Congress of the Swiss Psychological Society, Fribourg, Switzerland, 12th–13th September, 2011.

Hagger, M. S., Asci, F. H., & Lindwall, M. (2004). A cross-cultural evaluation of a multidimensional and hierarchical model of Physical Self-Perceptions in three national samples. *Journal of Applied Social Psychology*, *34*(5), 1075–1107. doi:10.1111/jasp.2004.34.issue-5

Hagger, M. S., Biddle, S. J., Chow, E. W., Stambulova, N., & Kavussanu, M. (2003). Physical self-perceptions in adolescence: Generalizability of a hierarchical multidimensional model across three cultures. *Journal of Cross-Cultural Psychology*, *34*(6), 611–628. doi:10.1177/0022022103255437

Hammer, A. (1950). A factorial analysis of the Bellevue intelligence tests. *Australian Journal of Psychology*, *1*(2), 108–114. doi:10.1080/00049535008256022

Harter, S. (1985). *Manual for the self-perception profile for children*. Denver, CO: University of Denver.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York, NY: Guilford Press.

Hayes, S. D., Crocker, P. R., & Kowalski, K. C. (1999). Gender differences in physical self-perceptions, global self-esteem and physical activity: Evaluation of the physical self-perception profile model. *Journal of Sport Behavior*, *22*(1), 1–14.

Hazlett-Stevens, H., Ullman, J. B., & Craske, M. G. (2004). Factor structure of the Penn State Worry Questionnaire: Examination of a method factor. *Assessment*, *11*(4), 361–370. doi:10.1177/1073191104269872

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*(3), 393–416. doi:10.1177/0013164405282485

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*(1), 41–54. doi:10.1007/BF02287965

Holzinger, K., & Swineford, F. (1939). A study in factor analysis: The stability of a bifactor solution. *Supplementary Educational Monograph*. no. 48. Chicago, IL: University of Chicago Press.

Horan, P. M., DiStefano, C., & Motl, R. W. (2003). Wording effects in self-esteem scales: Methodological artifact or response style? *Structural Equation Modeling*, *10*(3), 435–455. doi:10.1207/S15328007SEM1003_6

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*(4), 537–549. doi:10.1007/s11336-011-9218-4

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika*, *77*(3), 442–454. doi:10.1007/s11336-012-9269-1

Jensen, A. R., & Weng, L. J. (1994). What is a good? *Intelligence*, *18*(3), 231–258. doi:10.1016/0160-2896(94)90029-9

Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology*, *85*(5), 751–765. doi:10.1037/0021-9010.85.5.751

Karteroliotis, K. (2008). Validation of the Physical Self-Perception Profile among college students. *Journal of Education and Human Development*, *2*(1), 37–51.

Leue, A., & Beauducel, A. (2011). The PANAS structure revisited: On the validity of a bifactor model in community and forensic samples. *Psychological Assessment*, *23*(1), 215–225. doi:10.1037/a0021400

Lindwall, M., & Hassmen, P. (2004). The role of exercise and gender for physical self-perceptions and importance ratings in Swedish university students. *Scandinavian Journal of Medicine & Science in Sports*, *14*(6), 373–380. doi:10.1046/j.1600-0838.2003.372.x

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, *97*(3), 562–582. doi:10.1037/0033-2909.97.3.562

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*(3), 471–491. doi:10.1037/a0019227

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*, 85–110. doi:10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439–476. doi:10.1080/10705510903008220

Martel, M. M., Roberts, B., Gremillion, M., von Eye, A., & Nigg, J. T. (2011). External validation of bifactor model of ADHD: Explaining heterogeneity in psychiatric comorbidity, cognitive control, and personality trait profiles within DSM-IV ADHD. *Journal of Abnormal Child Psychology*, *39*(8), 1111–1123. doi:10.1007/s10802-011-9538-y

McAbee, S. T., Oswald, F. L., & Connelly, B. S. (2014). Bifactor models of personality and college student performance: A broad versus narrow view. *European Journal of Personality*, *28*(6), 604–619.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. doi:10.1037/0022-3514.52.1.81

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215. doi:10.1111/jopy.1992.60.issue-2

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2015). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometrics multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*. doi:10.1080/10705511.2014.961800

Motl, R. W., & DiStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(4), 562–578. doi:10.1207/S15328007SEM0904_6

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, *105*(3), 430–445. doi:10.1037/0033-2909.105.3.430

Musek, J. (2007). A general factor of personality: Evidence for the big one in the five-factor model. *Journal of Research in Personality*, *41*(6), 1213–1233. doi:10.1016/j.jrp.2007.02.003

Myers, N. D., Martin, J. J., Ntoumanis, N., Celimli, S., & Bartholomew, K. J. (2014). Exploratory bi-factor analysis in sport, exercise, and performance psychology: A substantive-methodological synergy. *Sport, Exercise, and Performance Psychology*, *3*, 258–272. doi:10.1037/spy0000015

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696. doi:10.1080/00273171.2012.715555

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*(6), 544–559. doi:10.1080/00223891.2010.496477

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*(S1), 19–31. doi:10.1007/s11136-007-9183-7

Revelle, W., & Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, *47*(5), 493–504. doi:10.1016/j.jrp.2013.04.012

Rushton, J. P., & Irwing, P. (2009a). A general factor of personality (GFP) from the multidimensional personality questionnaire. *Personality and Individual Differences*, *47*(6), 571–576. doi:10.1016/j.paid.2009.05.011

Rushton, J. P., & Irwing, P. (2009b). A general factor of personality in 16 sets of the Big Five, the Guilford–Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory. *Personality and Individual Differences*, *47*(6), 558–564. doi:10.1016/j.paid.2009.05.009

Rushton, J. P., & Irwing, P. (2009c). A general factor of personality in the Comrey personality scales, the Minnesota multiphasic personality inventory-2, and the multicultural personality questionnaire. *Personality and Individual Differences*, *46*(4), 437–442. doi:10.1016/j.paid.2008.11.015

Saklofske, D. H., Austin, E. J., & Minski, P. S. (2003). Factor structure and validity of a trait emotional intelligence measure. *Personality and Individual Differences*, *34*(4), 707–721. doi:10.1016/S0191-8869(02)00056-9

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*(4), 304–321. doi:10.1177/0734282911406653

Schriesheim, C. A., & Eisenbach, R. J. (1995). An exploratory and confirmatory factor-analytic investigation of item wording effects on the obtained factor structures of survey questionnaire measures. *Journal of Management*, *21*(6), 1177–1193. doi:10.1177/014920639502100609

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, *52*, 333–343.

Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research*, *58*(7), 935–943. doi:10.1016/j.jbusres.2003.10.007

Simms, L. J., Grös, D. F., Watson, D., & O'Hara, M. W. (2008). Parsing the general and specific components of depression and anxiety with bifactor modeling. *Depression and Anxiety*, *25*(7), E34–E46. doi:10.1002/da.v25:7

Steiger, J. H. (1990). Structural model evaluation and modification. *Multivariate Behavioral Research*, *25*(2), 173–180.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (5th ed). New York, NY: Routledge.

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, *10*(4), 297–303. doi:10.1111/1467-9280.00157

Tomas, J. M., & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 84–98. doi:10.1080/10705519909540120

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

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. doi:10.1177/109442810031002

Watkins, M. W. (2010). Structure of the Wechsler Intelligence Scale for Children—Fourth Edition among a national sample of referred students. *Psychological Assessment*, *22*(4), 782–787. doi:10.1037/a0020043

Watkins, M. W., & Beaujean, A. A. (2014). Bifactor structure of the wechsler preschool and primary scale of intelligence—fourth edition. *School Psychology Quarterly*, *29*(1), 52–63. doi:10.1037/spq0000038

Xie, J., Bi, Q., Shang, W., Yan, M., Yang, Y., Miao, D., & Zhang, H. (2012). Positive and negative relationship between anxiety and depression of patients in pain: A bifactor model analysis. *Plos One*, *7*(10), e47577. doi:10.1371/journal.pone.0047577

Xu, X., & Yao, J. X. (2001). Revision and test of Physical Self-perception Profile for college students. *China Sport Science*, *21*(2), 78–81.

Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany, NY: State University of New York Press.

Yuan, K.-H., & Bentler, P. M. (2008). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*(1), 165–200. doi:10.1111/0081-1750.00078

Yung, Y., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128. doi:10.1007/BF02294531

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for omega. *Applied Psychological Measurement*, *30*, 121–144. doi:10.1177/0146621605278814