2018

Evidence for a bifactor structure of the Scales of Psychological Well-being using exploratory structural equation modeling

Jose A. Espinoza
Western University, jespinoz@uwo.ca

John P. Meyer
Western University, meyer@uwo.ca

Brittney K. Anderson
Western University, bander27@uwo.ca

Chelsea Vaters
Western University, cvaters@uwo.ca

Christina Politis
christina.politis25@gmail.com

Evidence for a bifactor structure of the Scales of Psychological Well-being using exploratory structural equation modeling

Jose A. Espinoza [a], John P. Meyer [ab], Brittney K. Anderson [a], Chelsea Vaters [a], Christina Politis [a]

[a] Department of Psychology, The University of Western Ontario

[b] Curtin Business School, Curtin University

Corresponding Author:
Jose A. Espinoza
Department of Psychology
The University of Western Ontario,
London, Ontario, Canada N6A 5C2
Email: jespinoz@uwo.ca

Author ORCID & Contact Information:
Jose A. Espinoza: 0000-0002-6352-6830; jespinoz@uwo.ca
John P. Meyer: 0000-0001-8377-2003; meyer@uwo.ca
Brittney K. Anderson: bander27@uwo.ca
Chelsea Vaters: cvaters@uwo.ca
Christina Politis: cpoliti2@uwo.ca

Acknowledgements:

This research was supported by a research grant from the Social Sciences and Humanities Research Council (435-2014-0956) awarded to the second author.

This is a prepublication version of a manuscript accepted for publication on April 8, 2018, please cite as:

Espinoza, J.A., Meyer, J.P., Anderson, B.K., Vaters, C., & Politis, C. (in press, 2018). Evidence for a bifactor structure of Scales for Psychological Well-being using exploratory structural equation modeling. *Journal of Well-being Assessment*

Abstract

This research investigates the much-debated factor structure of the 54-item version of Ryff's (1989) Scales of Psychological Well-being (SPWB). Using two samples (n_1 = 573; n_2 = 449) of undergraduate university students, we apply confirmatory factor analysis (CFA) along with recently developed exploratory structural equation modeling (ESEM) techniques to evaluate several unidimensional and multidimensional models identified in previous research, as well as a new bifactor model. In a bifactor model, items load directly on both a global and a specific factor; when tested using ESEM, cross-loadings on other specific factors are also permitted and are targeted to be as close to zero as possible. After comparing various ESEM and traditional CFA models, the results indicate that a bifactor model estimated using ESEM provided the best fit to the data. Most items were found to reflect the global factor, but some items failed to reflect the intended specific factor. Thus, the 54-item version of the SPWB appears to be a good measure of overall psychological well-being, but may need refinement as a measure of the intended specific factors, at least among young adults. The benefits of applying ESEM to investigate the factor structure of the SPWB in other populations are discussed.

KEYWORDS (4 to 6): psychological well-being, measurement, measure evaluation, scale evaluation, exploratory structural equation modeling, ESEM

WORD COUNT (up to 250): 200

1. Introduction

In 1989, Carol Ryff introduced a model of psychological well-being to extend thinking and research beyond the traditional focus on hedonic well-being (happiness; life satisfaction; positive affect). The model had its roots in Aristotelian philosophy (Aristotle, trans. 1985) and was intended to integrate concepts from existential and humanistic psychology (e.g., Allport, 1961; Frankl & Lasch, 1959/1992; Maslow, 1968; Rogers, 1962) to capture the essence of what is now commonly referred to as eudaimonic well-being (Ryan & Deci, 2001; Waterman, 2011). In contrast to hedonic well-being, eudaimonic well-being emphasizes the realization of an individual's unique potential. Ryff proposed a multi-dimensional framework and developed a measure with six subscales: *purpose in life* (i.e., feeling that one's life has meaning and purpose), *autonomy* (i.e., living in accord with one's personal convictions), *personal growth* (i.e., making use of one's talents and potential), *environmental mastery* (i.e., effective management of ones' life situation), *positive relations* (i.e., deep connections with significant others), and *self-acceptance* (i.e., knowledge and acceptance of one's strengths and limitations).

Ryff's (1989) Scales of Psychological Well-being (SPWB) have been translated into multiple languages and used extensively in research around the world (see Ryff, 2014, for a review). Despite the extensive use, there is some controversy regarding the dimensionality of the SPWB. Indeed, this issue was the focus of a debate by Ryff and Singer (2006) and Springer and colleagues (Springer & Hauser, 2006; Springer, Hauser, & Freese, 2006) in *Social Science Research*. Springer and Hauser acknowledged that a 6-factor solution fit best (albeit not well) for several data sets, but noted that the correlations among some of the factors were sufficiently strong (.72 – .97) that they could not be considered distinct constructs. Ryff acknowledged the strong correlations but cited a large body of evidence demonstrating that even the more highly correlated scales relate differently to other variables. For example, in several cross-sectional studies it has been found that environmental mastery increases with age, whereas purpose in life and personal growth decline, and self-acceptance does not change (Clarke, Marshall, Ryff, & Rosenthal, 2000; Ryff, 1989; Ryff & Keyes, 1995). These differences, she argued, provide evidence that the distinctions among the six dimensions are worth making.

Several other investigators have also addressed the dimensionality issue using a variety of analytic techniques, including principal components analysis (PCA; e.g., Kofka & Kozma, 2002), exploratory factor analysis (EFA; e.g., Burns & Machin, 2009), and confirmatory factor analysis (CFA; e.g., Abbott, Ploubidid, Huppert, Kuh, & Croudace, 2006). The findings have been mixed, as have the conclusions concerning how the SPWB should be scored, interpreted, and applied. The analytic techniques used in these studies all have their limitations. Therefore, in the present investigation, we evaluated the dimensionality of the SPWB using a relatively new analytic technique, exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009; Marsh, Morin, Parker, & Kaur, 2014), that capitalizes on the strengths of previous exploratory and confirmatory techniques. Specifically, we used ESEM to test several competing models, including those proposed in the previous literature. We describe ESEM and its relative advantages over traditional EFA and CFA below, following a more detailed review of previous research on the dimensionality of the SPWB.

1.1 Factor Structure of the SPWB

The SPWB were developed by Ryff (1989) using a construct-oriented approach (Wiggins, 1973). Starting with theoretical descriptions of individuals at the high and low ends of six well-being dimensions, Ryff wrote a large set of items for each construct. The content of these items varied widely to capture their full breadth and relevance to diverse populations. The original measure comprised 20 items per scale, but these have subsequently been reduced to 14-, 9-, 7-, and 3-item versions. The measures are typically administered in an anonymous self-report format, but have also been administered in face-to-face and telephone interviews. Differences in the observed factor structure of the SPWB are undoubtedly affected by the scale length and administration strategy but, as discussed in the next section, might also reflect differences in the analytic strategies used by different investigators.

Although Ryff and her colleagues (Clarke, Marshall, Ryff, & Wheaton, 2001; Ryff, 1989; Ryff & Keyes, 1995) provided some support for a 6-dimensional structure of the SPWB, others have challenged these findings. For example, Kafka and Kozma (2002) reported evidence for a single dimension, and other researchers have suggested that the strong correlations among the factors justifies inclusion of a global second-order factor (van Dierendonck, 2004). Still others suggest that there may be sufficient justification for treating autonomy and personal relations as independent factors, but that environmental mastery, self-acceptance, purpose in life, and personal growth can be combined or represented by a second-order factor (e.g., Abbott et al., 2006; Burns & Machin, 2009). Other researchers have argued that a 6-factor model provides the best fit to the data, but that the factors are highly correlated (Springer & Hauser, 2006). Springer and Hauser also argued that, even though it fits the data best, the 6-factor model does not fit the data well. To obtain acceptable fit, they included method factors to account for keying direction and allowed correlated errors among adjacent items.

As noted previously, Ryff and Singer (2006) acknowledged the high correlations among the dimensions, but argued that there is sufficient evidence to suggest that the dimensions relate differently to other variables. They also argued that the content of the dimensions is quite distinct (i.e., the dimensions have face validity) and individual dimensions might be of particular interest in applied contexts. Thus, all things considered, Ryff and Singer argued that the SPWB are best treated as comprising six distinguishable dimensions.

One noteworthy anomaly regarding scale length is the finding that the short 3-item versions of the SPWB tend to have the lowest reliability, but also yield among the strongest support for a 6-factor structure. In contrast, the longer versions tend to be more reliable, but often fail to support a 6-factor structure. The low reliability of the 3-item versions is not surprising, as reliability tends to increase with scale length (Cronbach, 1970). In this case, reliabilities are particularly likely to be affected by attempts to maintain content breadth in the selection of items. It is also possible, however, that the selection of such a small set of items allowed for minimization of content overlap across dimensions. In contrast, aggregation across a larger number of items with diverse content can yield high reliabilities even when inter-item correlations are quite low. However, as item content becomes more varied, the likelihood of content overlap increases. This might help to explain why it is more difficult to find support for the 6-dimensional structure using longer scales. As we discuss below, overlap in item content is incompatible with the assumptions underlying parameter estimation in CFA.

1.2 Analytic Procedures

A variety of analytic techniques have been employed in the investigation of the factor structure of the SPWB. Kafka and Kozma (2002) used PCA with orthogonal rotation. PCA is an exploratory technique, but is best suited to the purpose of data reduction rather than identification of an optimal latent-factor structure (Bandalos & Boehm-Kaufman, 2009). Moreover, the use of orthogonal rotation is incompatible with the theoretical structure of Ryff's (1989) model. Other authors have used exploratory factor analyses with oblique rotation (e.g., Burns & Malchin, 2009). EFA can be very useful when the factor structure is unknown, but it is less well-suited to testing theory. Initial extraction of factors is intended to account for maximum variance among the variables. Rotation tends to spread the variance more evenly across factors, but the criteria used for this purpose (e.g., simple structure; Thurstone, 1945) are not theory-driven. Consequently, EFA with oblique rotation will not always produce a structure compatible with theory, even when such a structure would provide a good fit to the data.

The most common technique used to investigate the factor structure of the SPWB has been CFA. With CFA, the investigator specifies one or more models (i.e., factor structures) in advance. Each model is specified by indicating which items will define each factor. That is, each item is presumed to reflect only one latent factor; loadings on all other factors are set to zero. One of the major advantages of CFA is the ability to measure and compare the fit of alternative models. A model is said to 'fit' the data well when the parameters (e.g., loadings, disturbance terms, factor correlations) estimated for a model can be used to reproduce a covariance matrix that is a close approximation of the original. Arguably, one of the disadvantages of CFA is the often-unrealistic assumption that the loadings of an item on all latent factors but one are zero (Asparouhov & Muthén, 2009; Marsh et al., 2014). This assumption is particularly unrealistic for multidimensional constructs like psychological well-being where there is likely to be some natural overlap. This overlap is clearly seen in the non-zero cross-loadings in EFA.

Another advantage of CFA over EFA is the ability to test hierarchical models, where the first-order latent factors define one or more higher-order factors. As noted earlier, several studies found evidence for a higher-order global well-being factor defined by the six first-order well-being dimensions (e.g., Abbott et al., 2006; Springer & Hauser, 2006). However, it has been noted recently that this approach involves assumptions that might not be consistent with the data (e.g., Morin et al., 2016; Reise, 2012). Specifically, in a hierarchical model, the influence of the higher-order factor on the indicator variables (e.g., items) is presumed to be indirect through the first-order factors (i.e., the effect is a product of the loading of the first-order factor on the second-order factor and the loading of the indicator on the first-order factor). Because the indicators of a first-order factor all share that factor's loading on the second-order factor, it is assumed that the influence of the second-order factor on the indicators, albeit different, are proportional (Gignac, 2016; Morin et al., 2016). This may or may not be true, and could lead to misinterpretation of the data.

A relatively new analytic strategy that combines the advantages of EFA and CFA is ESEM (Asparouhov & Muthén, 2009; Marsh et al., 2014). Like CFA, ESEM is model-based and can be used to evaluate and compare the fit of multiple solutions. However, like EFA, ESEM allows for modest cross-loadings of items on factors other than its designated factor. This is accomplished by using target rotation procedures whereby cross-loadings are freed with the specification that they be as close to zero as possible. Also, like CFA, ESEM can be used to test models where it is assumed that an indicator reflects both a specific and a global factor. For example, in the case of the SPWB, an item intended to measure autonomy might also be expected to reflect a global sense of psychological well-being. This model is tested by specifying what is referred to as a *bifactor* model. In the bifactor model, each indicator loads directly on its designated specific factor and on a global factor. By allowing each indicator to load directly on the global factor, the assumption of proportionality inherent in hierarchical models is unnecessary. The resulting factors are also orthogonal and allow for the assessment of the unique variance in each item that is attributable to the global factor and the corresponding specific factor (Gignac, 2016).

Although it is relatively new, ESEM has been used with success in the evaluation of other established measures. For example, Guay, Morin, Litalien, Valois, & Vallerand (2015) used ESEM to evaluate the factor structure of the Academic Motivation Scale (AMS; Vallerand et al., 1992). Previous studies using CFA were often unable to find good fit for a model reflecting the underlying theory. Guay et al. found that by allowing cross-loadings, the expected 7-factor model fit the data well. One of the problems with the AMS that contributed to difficulties in the CFA was the strong correlations among the latent factors. This is the same problem that has been observed for the SPWB (e.g., Springer & Hauser, 2006). When cross-loadings are fixed (unrealistically) at zero in CFA, the variance shared by indicators of the different constructs in the model is accounted for by inflated correlations among the latent factors. When cross-loadings are permitted, the correlations among the latent factors decrease (Asparouhov & Muthén, 2009). More recently, Litalien, Morin, Gagné, Vallerand, Losier, and Ryan (2017) also examined the AMS and found that a bifactor ESEM structure was the best fitting model even when compared to the 7-factor ESEM solution supported in Guay et al.'s (2015) work. Litalien and colleagues' finding is an apt example of how ESEM often results in reduced latent factor correlations, but can also result in inflated cross-loadings when the presence of a global factor is ignored, leading to suboptimal fit (Morin, Arens, & Marsh, 2016). Other researchers have also applied ESEM and found a bifactor structure to fit the data best for a variety of measures with similar issues as the SPWB (e.g., Howard, Gagné, Morin, & Forest, 2016; Morin et al., 2016; Sánchez-Oliva, et al., 2017), suggesting that it might be important to investigate both the theorized 6-factor structure of the SPWB as well as a bifactor structure.

1.3 Present Research

Our objectives in the present research were to use ESEM to investigate the factor structure of the SPWB with data obtained from two university student samples. Although exclusive use of university samples limits generalizability to other populations, it is a population that has been used in several previous investigations (e.g., Gallagher, Lopez, & Preacher, 2009; Kafka & Kozma, 2002; van Dierendonck, 2004) and for whom psychological well-being is an important variable (e.g., psychological well-being is negatively associated with various health risk behaviours in young adults attending college; Schwartz et al., 2011). Moreover, having multiple samples from the same population allows us to evaluate consistency within a population. We hope that by introducing ESEM in the investigation of the dimensionality of the SPWB with this sample, we will stimulate others to apply it to a wider range of populations.

We used ESEM to compare the fit of various models that have been proposed and or detected using alternate analytic techniques in previous research. These included a 1-factor (global well-being) model, a 3-factor model (based on evidence for strong overlap between self-acceptance, environmental mastery, purpose in life, and personal growth), and a 6-factor model. We also tested for the co-existence of specific and global factors by testing 3- and 6-bifactor models.[1] To assess the efficacy of ESEM for investigating the dimensionality of the SPWB, we

---

[1] For completeness, hierarchical 3- and 6-factor CFA models were also estimated. However, these models provided a poorer fit to the data than the corresponding non-hierarchical CFA models. For this reason, and because these models often involve unrealistic assumptions (e.g., Morin et al., 2016; Reise, 2012), they are not reported in the Results but are available from the first author upon request.

also conducted a series of CFA on the same models estimated using ESEM. For illustrative examples of prototypical CFA and ESEM models see Figures 1–2.

2. Method

2.1 Participants

Data were collected from two samples of university students enrolled in a large Canadian university. Individuals were eligible to participate if they were enrolled in the introductory psychology course at the university and if they were fluent in English. All eligible students received the recruitment information for the study online, which was posted on the online platform that the university uses to recruit students to participate in various research studies. Data for Sample 1 were collected during the 2014 and 2015 academic years (*N* = 604), and data for Sample 2 were collected during the 2016 academic year (*N* = 497). All of these data were collected as part of a larger study on student well-being. Data for the 2014 and 2015 years were combined to achieve a sufficient sample size for purposes of analysis as these data were collected prior to the decision to employ ESEM to examine the factor structure of the SPWB. Ethical approval for these studies were obtained from the Non-Medical Research Ethics Board at the University of Western Ontario (Sample 1 Reference Number: 106297/107341; Sample 2 Reference Number: 107804).

Similar data cleaning procedures were conducted for both samples prior to analyses. First, data from participants were excluded based on their responses to direct-check careless responding items interspersed through out the survey (Kam & Meyer, 2015). The careless responding items indicated to participants which response option to select for that item (e.g., “Please select Strongly Disagree for this item”). In Sample 1, participants who responded correctly to at least five out of six careless responding items were retained (*n* = 590). In Sample

2, participants who responded correctly to at least four out of five careless responding items were retained (n = 451). Second, given that average survey completion time for both samples was approximately 30 minutes, we omitted data from those who completed the survey in less than 10 minutes and may have given insufficient attention to the task. This resulted in the elimination of 17 cases in Sample 1 and two cases in Sample 2. Final Ns for purposes of analysis were 573 and 449 for Samples 1 and 2, respectively.

The average age in Sample 1 was 18.38 years, with 65% of the population identifying as female. A majority of the participants indicated they were Caucasian (53%), while 17% indicated they were Chinese, 12% were South Asian, nearly 5% indicated they were Arab, and less than 4% each indicated they were Black, Korean, Southeast Asian, West Asian, Japanese, Filipino, Aboriginal or Latino. In Sample 2, the average age was 18.2 years old, and the sample was 81% female. Again, approximately half of the sample identified as Caucasian (52%), while 22% indicated they were Chinese, 12% were South Asian, and 4% or less each reported they were Arab, Black, Korean, Southeast Asian, West Asian, Japanese, Filipino, Aboriginal or Latino. Because of its broad appeal, the introductory psychology course draws students from a variety of disciplines. For both samples, approximately one third of the participants were from the Faculty of Social Science, one third were from the Faculty of Science, and one fifth were from the Faculty of Health Sciences, with the remaining students distributed across almost all of the other smaller faculties.

2.2 Measures

Psychological well-being. We measured psychological well-being using the 9-item versions of the SPWB developed by Ryff (1989). We opted to use the 9-item versions of the scales to achieve a balance between length and content breadth. Participants responded to the measure on a 6-point Likert-type scale from 1 (strongly disagree) to 6 (strongly agree). When referring to specific SPWB items in our results, we use item codes and these can be found in the Appendix along with the full content of the scales. Descriptive statistics for all items can be found in Online Resource 1.

2.3 Analyses

Analyses were conducted using Mplus (version 7.3, Muthén & Muthén, 2014). Preliminary analyses revealed that the responses to the SPWB items were slightly skewed (Sample 1: -1.519 to .469; Sample 2: -1.581 to .526). Therefore, all models were estimated using robust Maximum Likelihood (MLR). MLR calculates fit indices and standard errors robust to violations of normality assumptions, as well as to the nature of Likert response scales with five or more response categories (Finney & DiStefano, 2013). We evaluated five models using CFA and four models using ESEM. In both cases, we evaluated 3- and 6-factor models; a 1-factor model was also tested using CFA.[2] The 3-factor model was created by combining environmental mastery, self-acceptance, purpose in life, and personal growth. We also tested bifactor CFA and ESEM models in which all items loaded on both the specific and global factors. Bifactor models

---

[2] In the past, some investigators have reported that fit improves in models allowing correlations among the residuals for adjacent items and/or by allowing negatively worded items to load on a method factor (e.g., Springer & Hauser, 2006). Due to the complexity of the ESEM models tested here (e.g., 6-bifactor model), we did not use either of these procedures to account for potential method variance. However, inspection of the modification indices provided by Mplus suggested that applications of these procedures would be unlikely to improve model fit substantially. Details regarding these modification indices can be found in Online Resource 2.

with both three and six specific factors were tested. All CFA models were specified so that items were only allowed to load on their corresponding factors and all cross-loadings were constrained to zero. In bifactor CFA models, the factors were specified to be orthogonal and items were only allowed to load on their corresponding specific factor and the global factor. ESEM models were specified using target rotation, allowing all items to load on all factors with cross-loadings for each item on non-corresponding specific factors estimated to be as close to zero as possible. Bifactor ESEM models were estimated using orthogonal target rotation where latent correlations between factors were constrained to zero.

Considering the oversensitivity to sample size and minor changes in specifications of the chi-square difference test (Marsh, Hau, & Grayson, 2005), recent research employing ESEM (e.g., Howard et al., 2016; Morin et al., 2016; Sánchez-Oliva et al., 2017) has instead relied on goodness-of-fit indices and inspection of parameter estimates to compare models. Consistent with these earlier studies, all models were evaluated using multiple goodness-of-fit indices and information criteria: the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), the Akaike Information Criteria (AIC), the Bayesian Information Criteria (BIC), and the sample-size adjusted BIC (ABIC). Following guidelines by Hu and Bentler (1999), models with values greater than .90 on the CFI and TLI suggest acceptable fit, while values greater than .95 denote an excellent fit. Values on the RMSEA below .08 reflect adequate fit, whereas values below .05 reflect good fit (Browne & Cudeck, 1993). When comparing models, we were guided by Chen's (2007) suggestion that an increase in CFI of .005–.010 and a decrease in RMSEA of .010–.015 indicates a significant improvement in fit to the data. AIC, BIC, and ABIC values were also used to compare models, with lower values indicating a better fit to the data. Although the guidelines above were originally established for CFA, previous research has used them to evaluate the fit of models derived using ESEM (e.g., Howard et al., 2016; Morin et al., 2016; Sanchéz-Oliva et al., 2017). When comparing models using CFA and ESEM, it is important to keep in mind that fit will generally be better for comparable ESEM models because they involve less restrictive assumptions for cross-loadings (i.e., the loadings are targeted to be as close to zero as possible rather than being fixed at zero). Morin et al. (2016) point out that a more important consideration when comparing CFA and ESEM models is to examine the parameter estimates (e.g., correlations among the latent factors) to determine which are more in line with theory.

3. Results

Model values on goodness-of-fit indices and information criteria for both samples can be found in Table 1. Bifactor CFA models did not converge after 5,000 iterations in Sample 2 and, therefore, are not reported here. Results of the factor analyses indicated that CFI and TLI values for CFA and ESEM models fell below acceptable levels in both samples, whereas the RMSEA suggested that these models provided good fit. The exception was the 6-bifactor ESEM model in Sample 1, which did meet the cut-off on the CFI for acceptable fit (CFI = .914). Generally, comparisons of the models based on the information criteria supported the fit of the ESEM models over the corresponding CFA models as expected, given the number of additional parameters that are freely estimated in ESEM solutions. Comparisons of the models also supported the fit of all multidimensional models over the unidimensional solution. These results suggest that the presence of cross-loadings that are not allowed by CFA are likely one of the sources of poor fit for these models. Supporting this conclusion, the ESEM models also reduced the latent factor correlations as expected when compared to the corresponding CFA models (see Online Resources 3–6 in the supplementary materials). For example, the average correlation between latent factors in the 6-factor ESEM models (Sample 1 $|M_r|$ = .273; Sample 2 $|M_r|$ = .265) were less than half of the average correlation in the corresponding 6-factor CFA models (Sample 1 $|M_r|$ = .643; Sample 2 $|M_r|$ = .627). Finally, the 6-factor solutions were all found to fit the data better than their corresponding 3-factor solutions. Based on the results above, only the detailed comparisons between the 6-factor ESEM model and the 6-bifactor ESEM model are reported here.

The 6-bifactor ESEM model resulted in a better fit than the 6-factor ESEM model in both samples (Sample 1: ΔCFI = +.017; ΔTLI = +.016; ΔRMSEA = -.002; lower AIC and ABIC; Sample 2: ΔCFI = +.039; ΔTLI = +.042; ΔRMSEA = -.006; lower AIC and ABIC). Notably, the 6-bifactor ESEM model was not an improvement over the 6-factor ESEM according to Chen's (2007) guidelines for RMSEA (Sample 1: ΔRMSEA = -.002; Sample 2: ΔRMSEA = -.006) although it did meet the criteria for the CFI (Sample 1: ΔCFI = +.017; Sample 2: ΔCFI = +.039). Overall, these results suggest that the specification of a global factor improved the fit of the ESEM model as expected. The 6-bifactor ESEM model provided the best fit across all solutions, likely because it allows the small cross-loadings for each item present in the 6-factor ESEM model while simultaneously specifying the global factor necessary to account for the variance shared by most of the items in the SPWB. However, according to Morin et al. (2016), model selection should be based on the examination of parameter estimates as well as on goodness-of-fit values, so we examined item loadings for these two solutions more closely to determine if the 6-bifactor ESEM model could be considered a better fit to the data than the 6-factor ESEM when parameter estimates were also considered. Parameter estimates for all 6-factor solutions are reported in Online Resources 7–10 in the supplementary materials.

Item loadings for the 6-factor ESEM solutions indicate that the specific factors were not clearly defined by the intended items. Indeed, the average item loadings per factor in both samples were below .5 (Sample 1 target $|M_\lambda|$ = .437; Sample 2 target $|M_\lambda|$ = .434) and two of the specific factors had at least one item with a loading below .25 (i.e., less than 7% of the item variance was explained by its corresponding specific factor). The cross-loadings were small on average (Sample 1 cross $|M_\lambda|$ = .105; Sample 2 cross $|M_\lambda|$ = .120), but ranged in absolute value from .003 to .502 in Sample 1 and .000 to .557 in Sample 2.

In comparison, parameter estimates for the 6-bifactor ESEM solutions revealed support for a global factor. Notably, this evidence is weaker in Sample 1 (global $|M_\lambda|$ = .439) than it is in Sample 2 (global $|M_\lambda|$ = .535), but item loadings on the global factor in both samples were often larger than loadings for these items on their corresponding specific factors (Sample 1 target $|M_\lambda|$ = .310; Sample 2 target $|M_\lambda|$ = .315). Similar to the 6-factor ESEM solutions, there were also small item cross-loadings on average (Sample 1 cross $|M_\lambda|$ = .100; Sample 2 cross $|M_\lambda|$ = .100). The absolute range of cross-loadings for the 6-bifactor ESEM solutions indicate that cross-loadings were reduced in comparison to the 6-factor ESEM solutions (Sample 1: .000–.386; Sample 2: .001–.421). These results suggest that the 6-bifactor ESEM solution is the best-fitting model to the data and a global factor is crucial in accounting for the variance shared by most of the items in the SPWB.

4. Discussion

4.1 Multidimensionality of the SPWB

Our objective in the present research was to investigate the factor structure of the SPWB using ESEM in addition to more traditional CFA. ESEM addresses some of the limitations inherent in CFA (e.g., forcing zero cross-loading) and has been used with success in clarifying the structure of measures with similar issues as the SPWB (e.g., Howard et al. 2016; Sánchez-Oliva et al., 2017). Based on our results with data collected from two samples of university students, the SPWB appears to be multidimensional. All the multidimensional solutions in our analyses provided a better fit to the data than a unidimensional solution. Relatedly, it appears that the 6-factor structure originally proposed by Ryff (1989) provides a better fit than a more parsimonious 3-factor structure proposed by some researchers (Abbott et al., 2006; Burns & Machin, 2009).

Some researchers have noted that, while a 6-factor solution produced the best fit, the factors were highly correlated, indicating the need for a second-order factor (van Dierendonck, 2004), or that they had to include method factors and correlated errors among some items to achieve acceptable fit (Springer & Hauser, 2006). Our results suggest that, to some extent, these issues might be caused by limitations inherent in CFA. When applied to measures with overlap in item content across correlated constructs, in the absence of a global factor and cross-loadings, CFA is forced to account for the shared item variance by inflating the latent factor correlations (Asparouhov & Muthén, 2009). The best-fitting solution in the present research was the 6-bifactor ESEM solution, likely because it: 1) reduces latent factor correlations by allowing small cross-loadings to account for overlap in item content; 2) specifies a global factor to account for shared variance resulting from experiences that contribute to overall well-being; and 3) contains specific factors reflecting meaningful distinctions as articulated by Ryff (1989) in her theory of psychological well-being.

4.2 Item Quality in the SPWB

Although the results of our analyses indicate that a 6-bifactor ESEM solution best reflects the structure underlying the SPWB, this does not attest to the quality of the SPWB as a multidimensional measure of well-being. Closer inspection of the parameter estimates for the 6-bifactor ESEM solution in both samples suggests that there are items that do not adequately reflect their intended specific factor or the global factor, at least for a university student population; see Table 2 for a summary of parameter estimates. While an in-depth analysis of the quality of the SPWB on an item-by-item basis is beyond the scope of this paper, we identify general categories of items that illustrate the issues that might be addressed in the future.

First, there are items on the SPWB that appear to be solely indicative of a global experience of psychological well-being rather than the specific factor they were intended to measure. These items were characterized by large loadings on the global factor and substantially smaller loadings on all other factors, including the target factor. Items in this category could be problematic in that they do not reflect the unique aspects of their corresponding subscale and could lead to erroneous interpretations of the scale scores, as well as relations between those scores and other variables. That is, the scale scores might be contaminated by overall well-being and this can contribute to over- or under-estimation of the specific scale's relations with other variables. One example is the item "In general, I feel I am in charge of the situation in which I live" from the Environmental Mastery scale (EM2; Sample 1: Global λ = .497; Target λ = .139; Sample 2: Global λ = .521; Target λ = -.070).

Second, there are items that have cross loadings on one or more factors that are substantially higher than the target-factor loading. For example, the item "I do not fit very well with the people and the community around me" has a higher loading on the Personal Relations factor than on the intended Environmental Mastery factor (EM14R; Sample 1: .358 vs. -.082; Sample 2: .421 vs. -.076). Again, inclusion of such items in the computation of scale scores can result in misinterpretation of the scores and their relations with other variables. In cases like the example given here, it might be better to re-classify the item than to delete it. Indeed, the content of item ER14 seems more indicative of “having warm, satisfying trusting relationships with others” (Positive Relations; Ryff, 1989, p. 1072) than of “being able to choose or create contexts suitable to personal needs and values” (Environmental Mastery; Ryff, 1989, p. 1072).

Third, some items are characterized by modest loadings on all factors. These items are problematic because they do not appear to be good indicators of any of the specific factors, or of the global factor. Consequently, they too are likely to contribute to misinterpretation of scale scores and their relations with other variables. One example is the item “there is truth to the saying you can't teach an old dog new tricks,” (PG51R) for which the highest loading was .241 on the Global factor in Sample 1 and .328 on the Personal Growth factor in Sample 2. Items such as these might best be replaced in subsequent revisions of the SPWB.

4.3 Limitations and Future Research Directions

Our study contributes to the on-going debate regarding the dimensionality of the SPWB by applying a new analytic technique, ESEM, to data obtained from two samples of university students. The consistency in findings across the two samples adds to our confidence that the best fitting factor structure for the SPWB is one including a global factor and six specific factors corresponding to the dimensions of well-being identified by Ryff (1989). However, our study is not without limitations, the most notable being our exclusive use of university students. Although university students, and young adults in general, constitute a large segment of the population for whom psychological well-being is very important, we cannot attest to the generalizability of our findings to other segments of the population. We strongly encourage additional research applying the new ESEM methodology using more diverse samples and/or other segments of the population where the SPWB might be applied in research and practice (e.g., older adults; disadvantaged youth).

One reason to question whether the results of our study will generalize to other segments of the population is item content. Although all items on the SPWB are likely to be interpreted somewhat differently depending on the individual's life circumstances, some items in particular might seem less relevant to young as opposed to middle-aged or older adults. The Personal Growth item "I gave up trying to make big improvements in my life a long time ago" is one example. If students interpreted and responded differently to such items than would older adults, this might have affected our results. Going forward, it is important to consider whether it would be best to eliminate items that might be interpreted differently by large segments of the population, or to develop different versions tailored to segments of the population that might be of special interest for research or practical purposes.

Our analysis was also limited to the 54-item version of the SPWB with its 9-item subscales. As noted earlier, longer versions (84 or 120 items) of the instrument include more diverse content and tend to have higher reliability than shorter versions with more restricted content but often do not provide support for the theorized 6-factor structure. The 54-item version was chosen for the present research because it provides a nice balance between length and content breadth. However, it is possible that our results might have been different had we used longer or shorter versions. Again, we encourage additional applications of ESEM with the other versions. The findings with longer versions could be of particular interest because they provide data for a wider range of items that would be helpful in scale refinements.

We also encourage future research to examine the factor structure of the SPWB using other methods, most notably, Bayesian CFA. Bayesian estimation, similarly to ESEM, overcomes the unrealistic constraints of traditional CFA by allowing the researcher to specify a prior distribution that models small but non-zero item cross-loadings alongside freely estimated loadings for items on their corresponding scale factors (Muthén & Asparouhov, 2012). The Bayesian approach has an advantage over other approaches based on maximum likelihood, including ESEM, when there is information available to specify an a priori distribution. Alternatively, in the absence of such information (as in the present case), one can estimate an initial prior distribution based on theory; however, these priors are likely to be highly subjective and should be evaluated using sensitivity analysis, wherein the results of multiple prior distributions are compared to determine the impact of the prior on the posterior distribution (van de Schoot et al., 2014; Zyphur & Oswald, 2015). Although the application of such an approach goes beyond the objectives of the present study, our findings using ESEM might guide the estimation of an a priori distribution for future applications of Bayesian CFA.

4.4 Implications

Our findings have important implications for applications of the SPWB in research and practice involving university students (and perhaps young adults more generally). The evidence we provide for the superior fit of a model with six specific factors and one global factor suggests that future researchers might begin by repeating this analysis to create factor scores for use in subsequent analyses linking well-being to other variables. Among other things, studies might be conducted to determine whether specific factors add incrementally beyond the global factor to the prediction of outcome variables of interest.

In contrast, practitioners are more likely to interpret the raw score composites for the six specific scales and/or the global measure of well-being. Our findings suggest that such interpretations, particularly for the individual scales, could be problematic. The creation of composite scores involves equal weighting of all designated items, each of which is assumed to be content relevant. Our findings suggest that this is not the case. Indeed, some items have very low loadings on the designated specific scale and/or on the global factor and should therefore not be included in creation of composites. This raises the question of whether there might be some reduced number of items that could legitimately be combined to produce meaningful composite scores. To address this question, we examined the pattern of loadings for each item in Table 2, and used the following criteria to generate recommendations that are summarized in the Appendix. Items with loadings of .40 or above on its designated factor in both samples were deemed appropriate for inclusion in a composite scale for that factor. Similarly, items with loadings of .40 or above on the global factor for both samples were deemed appropriate for inclusion on an overall well-being measure. Items with loading loadings of .40 or above on both the designated specific factor and on the global factor were deemed appropriate for inclusion on both the specific and global composites (assuming they are not used together in comparative analyses).

As can be seen from the Appendix, if we assume that a minimum of three items is sufficient to form a composite scale, only the Autonomy and Positive Relations factor qualify, with four and three items, respectively. The alpha coefficient for these two scales were .767 and .728, respectively, in Sample 1, and .734 and .767 in Sample 2. In contrast, there are many items that would qualify for inclusion on a composite global well-being scale, but these are spread unevenly across the specific factors, varying from two (Autonomy) to nine (Self-Acceptance). The alpha for such a global measure is .932 in Study 1 and .936 in Study 2. Taking the best two items per factor to create a balanced scale yields an alpha of .852 in Study 1 and .862 in Study 2.

Again, these recommendations and the findings upon which they are based, must be interpreted with caution given that they are derived from two samples, both consisting of university students. But it is important to be equally cautious in interpreting composite scores based on the item assignments in the existing 54-item version of the SPWB. There appears to be a clear need for revisions to the SPWB, perhaps beginning with additional items for several of the existing scales. Once these items are available, it will be necessary to re-evaluate the factor structure to determine whether it is consistent with the multidimensional structure proposed by Ryff (1989). We strongly recommend the use of ESEM or Bayesian CFA for this purpose. Although our findings will be most useful as a starting point for the refinement of measures for use with young adults, the procedures we used in the present study can also be applied in the evaluation of measures for use with other populations.

5. Conclusion

Our findings suggest that the Scales of Psychological Well-being (Ryff, 1989) are best represented by a multidimensional structure. Specifically, the best-fitting solution is one that contains factors for each scale indicated by their respective items and a global psychological well-being factor indicated by all items in the measure. This conclusion derives from our application of a relatively new analytic procedure, exploratory structural equation modeling, and we strongly encourage additional use of this procedure to evaluate the factor structure of the SPWB with a more diverse population and/or a wider range of subpopulations.

Conflicts of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Abbott, R. A., Ploubidis, G. B., Huppert, F. A., Kuh, D., Wadsworth, M. E., & Croudace, T. J. (2006). Psychometric evaluation and predictive validity of Ryff's psychological well-being items in a UK birth cohort sample of women. *Health and Quality of Life Outcomes*, 4(1), 76–92.

Allport, G.W. (1961). *Pattern and growth in personality*. New York, NY: Holt, Rinehart, and Winston.

Aristotle (1985). *Nicomachean ethics* (T. Irwin, Trans.). Indianapolis, IN: Hackett.

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.

Bandalos, D.L., & Boehm-Kaufman, M.R. (2009). Four common misconceptions in exploratory factor analysis. In. C.E. Lance & R.J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends* (pp. 61-87). New York: Routledge.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation models* (pp. 136–162). Newbury Park, CA: Sage.

Burns, R. A., & Machin, M. A. (2009). Investigating the structural validity of Ryff's psychological well-being scales across two samples. *Social Indicators Research*, *93*(2), 359–375.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464-504.

Clarke, P. J., Marshall, V. W., Ryff, C. D., & Rosenthal, C. J. (2000). Well-being in Canadian seniors: Findings from the Canadian study of health and aging. *Canadian Journal on Aging/La Revue Canadienne du Vieillissement*, *19*(2), 139–159.

Clarke, P. J., Marshall, V. W., Ryff, C. D., & Wheaton, B. (2001). Measuring psychological well-being in the Canadian study of health and aging. *International Psychogeriatrics*, *13*(1), 79–90.

Cronbach, L. J. (1970). *Essentials of psychological testing*. Macmillan: New York, NY: Macmillan.

Frankl, V. E., & Lasch, I. (1992). *Man's search for meaning: An introduction to logotherapy*. Boston, MA: Beacon Press. (Originally work published 1959)

Gallagher, M. W., Lopez, S. J., & Preacher, K. J. (2009). The hierarchical structure of well-being. *Journal of Personality*, *77*(4), 1025–1050.

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, *55*, 57-68.

Guay, F., Morin, A. J., Litalien, D., Valois, P., & Vallerand, R. J. (2015). Application of exploratory structural equation modeling to evaluate the Academic Motivation Scale. *The Journal of Experimental Education*, 83(1), 51–82.

Finney, S. J., & DiStefano, C. (2013). Non-normal and categorical data in structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 439–492) (2nd ed.). Greenwich, CO: IAP.

Howard, J. L., Gagné, M., Morin, A. J., & Forest, J. (2016). Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation. *Journal of Management*, 1–27.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis:

Conventional criteria versus new alternatives. *Structural equation modeling: A multidisciplinary Journal*, *6*(1), 1–55.

Litalien, D., Morin, A. J. S., Gagné, M., Vallerand, R. J., Losier, G. F., & Ryan, R. M. (Accepted, 29 June 2017). Evidence of a continuum structure of academic self-determination: A two-study test using a bifactor-ESEM representation of academic motivation. *Contemporary Educational Psychology*.

Kafka, G. J., & Kozma, A. (2002). The construct validity of Ryff's scales of psychological well-being (SPWB) and their relationship to measures of subjective well-being. *Social Indicators Research*, *57*(2), 171–190.

Kam, C. C. S., & Meyer, J. P. (2015). How careless responding and acquiescence response bias can influence construct dimensionality: The case of job satisfaction. *Organizational Research Methods*, *18*(3), 512-541.

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics. A Festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah NJ: Erlbaum.

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.

Maslow, A. H. (1968). *Toward a psychology of being*. New York: D. Van Nostrand Company.

Morin, A. J. S., Arens, A., & Marsh, H. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, *23*, 116–139.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*, 313–335.

Muthén, L. K., & Muthén, B. O. (2014). Mplus user's guide. Los Angeles, CA: Muthén & Muthén.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696.

Rogers, C. R. (1962), The interpersonal relationship: The core of guidance. *Harvard Educational Review*, *32*, pp. 416–429.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, *52*(1), 141–166.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 069–1081.

Ryff, C. D. (2014). Psychological well-being revisited: Advances in the science and practice of eudaimonia. *Psychotherapy and Psychosomatics*, *83*(1), 10–28.

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69*(4), 719–727.

Ryff, C. D., & Singer, B. H. (2006). Best news yet on the six-factor model of well-being. *Social Science Research*, *35*(4), 1103–1119.

Sánchez-Oliva, D., Morin, A. J., Teixeira, P. J., Carraça, E. V., Palmeira, A. L., & Silva, M. N. (2017). A bifactor exploratory structural equation modeling representation of the structure of the basic psychological needs at work scale. *Journal of Vocational Behavior*, *98*, 173–187.

Springer, K. W., & Hauser, R. M. (2006). An assessment of the construct validity of Ryff's scales of psychological well-being: Method, mode, and measurement effects. *Social Science Research*, *35*(4), 1080–1102.

Springer, K. W., Hauser, R. M., & Freese, J. (2006). Bad news indeed for Ryff's six-factor model of well-being. *Social Science Research*, *35*(4), 1120–1131.

Schwartz, S. J., Waterman, A. S., Vazsonyi, A. T., Zamboanga, B. L., Whitbourne, S. K., Weisskirch, R. S., . . . Ham, L. S. (2011). The association of well-being with health risk behaviors in college-attending young adults. *Applied Developmental Science*, *15*(1), 20-36.

Thurstone, L.L. (1945). *Multiple factor analysis*. Chicago: University of Chicago Press.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, *52*(4), 1003–1017.

van Dierendonck, D. (2004). The construct validity of Ryff's Scales of Psychological Well-being and its extension with spiritual well-being. *Personality and Individual Differences*, *36*(3), 629–643.

van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development*, *85*, 842–860.

Waterman, A. S. (2011). Eudaimonic identity theory: Identity as self-discovery. In *Handbook of Identity Theory and Research* (pp. 357–379). New York, NY: Springer.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Menlo

Park, CA: Addison-Wesley

Zyphur, M. J., & Oswald, F. L. (2015). Bayesian estimation and inference: A user’s guide. *Journal of Management*, *41*, 390–420.

Table 1

Results of Factor Analyses of the Scales of Psychological Well-being for Sample 1 and Sample 2

Model	χ^2, *df*	*CFI*	*TLI*	*RMSEA [90% CI]*	*AIC*	*BIC*	*ABIC*
Sample 1							
1. 1-Fac CFA	5168.609, 1377	.608	.593	.069 [.067, .071]	90335.347	91040.191	90525.911
2. 3-Fac CFA	4275.400, 1374	.700	.688	.061 [.059, .063]	89327.129	90045.025	89521.221
3. 6-Fac CFA	3841.390, 1362	.744	.731	.056 [.054, .058]	88851.471	89621.578	89059.679
4. 3-Bifactor CFA	3886.666, 1324	.735	.714	.058 [.056, .060]	88929.925	89865.365	89182.833
5. 6-Bifactor CFA	3563.699, 1324	.769	.750	.054 [.052, .056]	88581.765	89517.206	88834.673
6. 3-Fac ESEM	3534.566, 1272	.766	.737	.056 [.054, .058]	88596.995	89758.681	88911.071
7. 6-Fac ESEM	2119.036, 1122	.897	.869	.039 [.037, .042]	87299.214	89113.533	87789.738
8. 3-Bifactor ESEM	2883.849, 1221	.828	.799	.049 [.046, .051]	87964.790	89348.372	88338.859
9. 6-Bifactor ESEM	1910.509, 1074	.914	.885	.037 [.034, .040]	87134.148	89157.310	87681.135
Sample 2							
1. 1-Fac CFA	5007.594, 1377	578	.562	.077 [.074, .079]	71364.030	72029.368	71515.243
2. 3-Fac CFA	4286.376, 1374	.662	.648	.069 [.066, .071]	70564.032	71241.690	70718.045
3. 6-Fac CFA	3764.711, 1362	.721	.707	.063 [.060, .065]	69992.887	70719.730	70158.101
4. 3-Bifactor CFA	No convergence after 5000 iterations						
5. 6-Bifactor CFA	No convergence after 5000 iterations						
6. 3-Fac ESEM	3723.488, 1272	.715	.680	.066 [.053, .068]	69949.187	71045.762	70198.408
7. 6-Fac ESEM	2480.337, 1122	.842	.799	.052 [.049, .055]	68765.345	70488.973	69154.578
8. 3-Bifactor ESEM	3115.896, 1221	.780	.742	.059 [.056, .061]	69408.822	70714.855	69705.647
9. 6-Bifactor ESEM	2099.261, 1074	.881	.841	.046 [.043, .049]	68571.210	70480.976	69005.248

Note. Sample 1 n = 573; Sample 2 n = 449. All models estimated using MLR. Fac = Factor (e.g., 1-fac = 1-factor); CFA = Confirmatory Factor Analysis; ESEM = Exploratory Structural Equation Modeling; df = Degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation, 90% CI = 90% confidence interval for RMSEA; AIC = Akaike Information Criterial; BIC = Bayesian Information Criteria; ABIC = Sample-sized Adjusted BIC.

Table 2

Summary of Standardized Loadings (λ) and Uniquenesses (δ) for 6-Factor Bifactor ESEM Solution in Sample 1 and Sample 2

	Sample 1				*Sample 2*			
Items	Global λ	Target λ	Cross λ Range	δ	Global λ	Target λ	Cross λ Range	δ
Autonomy								
AU1	.300*	.624*	-.007–.027	.520*	.428*	.496*	-.324–.239	.401*
AU7	.285*	.295*	-.159–.113	.763*	.382*	.511*	-.276–.139	.517*
AU13R	.346*	.320*	-.235–.016	.715*	.248*	.425*	-.081–.215	.662*
AU19	.237*	.289*	-.233–.360	.665*	.503*	.210*	-.134–.209	.656*
AU25R	.287*	.335*	-.027–.189	.751*	.207*	.518*	-.133–.090	.661*
AU31	.414*	.587*	-.080–.129	.437*	.454*	.532*	-.194–.121	.445*
AU37R	.407*	.606*	-.106–.094	.433*	.407*	.570*	.056–.284	.485*
AU43R	.360*	.467*	-.193–-.056	.590*	.268*	.568*	.004–.194	.580*
AU49	.342*	.330*	-.011–.291	.632*	.513*	.416*	-.140–.179	.541*
	*M*G |λ| = .331	*M*T |λ| = .428	*M*C |λ| = .069		*M*G |λ| = .379	*M*T |λ| = .472	*M*C |λ| = .115	
Environmental Mastery								
EM2	.497*	.139*	-.026–.183	.682*	.521*	-.070	-.160–.200	.611*
EM8R	.600*	.055	-.281–.040	.547*	.400*	.423*	-.031–.259	.579*
EM14R	.562*	-.082	-.056–.358	.538*	.600*	-.076	-.072–.421	.514*
EM20	.495*	.625*	-.066–.087	.345*	.669*	.515*	-.155–.033	.344*
EM26R	.465*	.075	-.304–-.021	.633*	.296*	.485*	.026–.246	.590*
EM32	.302*	.235*	-.037–.116	.818*	.451*	.163	-.231–-.002	.726*
EM38	.415*	.598*	-.069–.099	.453*	.543*	.535*	-.199–-.008	.460*
EM44R	.724*	.050	-.164–.040	.436*	.740*	.371*	-.029–.193	.454*
EM50	.574*	.070	-.022–258	.594*	.672*	.065	-.110–.245	.551*
	*M*G |λ| = .515	*M*T |λ| = .214	*M*C |λ| = .082		*M*G |λ| = .544	*M*T |λ| = .300	*M*C |λ| = .114	
Personal Growth								
PG3R	.158*	.410*	-.077–.144	.759*	.264*	.268	-.080–.069	.856*
PG9R	.171*	.564*	-.326–.055	.540*	.159	.314	-.169–.291	.730*
PG15	.159*	.592*	-.015–.122	.595*	.340*	.248	-.245–.280	.584*
PG21R	.458*	.308*	-.105–.168	.628*	.663*	.321*	-.072–.110	.581*
PG27	.453*	.129	.028–.241	.582*	.648*	.245	-.141–.138	.482*
PG33R	.313*	.404*	-.254–.143	.787*	.299*	.186	-.027–.275	.811*
PG39	.495*	.183*	-.007–.260	.505*	.560*	.275	-.155–-.027	.517*

PG45R	.605*	.215*	-.031–.311	.499*	.705*	.289*	.019–.139	.512*
PG51R	.241*	-.003	-.076–.091	.875*	.282*	.328	-.032–.181	.810*
	*M*G |λ| = .339	*M*T |λ| = .312	*M*C |λ| = .086		*M*G |λ| = .436	*M*T |λ| = .275	*M*C |λ| = .094	
Positive Relations								
PR4	.242*	.288*	-.150–.330	.701*	.465*	-.102	-.320–.035	.625*
PR10R	.503*	.388*	-.125–-.050	.559*	.536*	.516*	-.123–.066	.559*
PR16R	.541*	.481*	-.192–.010	.393*	.748*	.575*	-.075–.127	.407*
PR22	.233*	.309*	.018–.357	.637*	.351*	-.020	-.250–.105	.716*
PR28R	.509*	.503*	-.137–.152	.439*	.668*	.477*	-.045–.025	.519*
PR34R	.500*	.332*	-.217–.017	.555*	.689*	.544*	-.080–.190	.453*
PR40	.266*	.281*	-.074–.386	.668*	.457*	-.123	-.226–-.023	.652*
PR46R	.402*	.455*	-.094–.049	.611*	.776*	.449*	-.083–.065	.503*
PR53	.332*	.489*	.001–.101	.619*	.430*	.162	-.163–-.057	.695*
	*M*G |λ| = .392	*M*T |λ| = .392	*M*C |λ| = .108		*M*G |λ| = .569	*M*T |λ| = .330	*M*C |λ| = .091	
Purpose in Life								
PL5R	.037	.559*	-.092–.068	.667*	.064	.239	-.252–.140	.845*
PL11R	.330*	.418*	-.213–.009	.669*	.315*	.399	-.019–.213	.683*
PL17R	.631*	.085	-.149–.095	.555*	.596*	.141	-.127–.246	.637*
PL23R	.507*	.257*	-.089–.027	.658*	.719*	.243	-.049–.185	.628*
PL29R	.570*	.383*	-.046–.104	.514*	.696*	.210	-.085–.283	.504*
PL35	.411*	.394*	-.085–.261	.572*	.548*	.291	-.172–.127	.588*
PL41	.558*	.212*	.006–.347	.501*	.745*	.100	-.142–.206	.411*
PL47	.544*	.278*	-.042–.106	.613*	.619*	.146	-.074–.191	.679*
PL54R	.270*	.211*	-.027–.286	.790*	.338*	.098	-.026–.400	.739*
	*M*G |λ| = .429	*M*T |λ| = .311	*M*C |λ| = .073		*M*G |λ| = .516	*M*T |λ| = .207	*M*C |λ| = .104	
Self-Acceptance								
SA6	.647*	.307*	-.141–.032	.455*	.799*	.339*	-.187–.073	.388*
SA12	.737*	.177*	-.129–.077	.391*	.810*	.508*	-.079–.085	.272*
SA18R	.695*	-.032	-.105–.008	.499*	.876*	.260*	.012–.275	.436*
SA24	.626*	.258*	-.122–.130	.490*	.634*	.194*	-.159–.111	.514*
SA30	.590*	.299*	-.019–.079	.551*	.717*	.142	-.138–.084	.519*
SA36R	.721*	.160*	-.088–.034	.443*	.931*	.295*	-.054–.178	.380*
SA42R	.636*	.007	-.143–.069	.551*	.735*	.501*	.005–.130	.459*
SA48	.514*	.366*	.000–.116	.586*	.807*	.133	-.280–.046	.566*
SA55	.493*	.202*	-.153–.000	.675*	.573*	.372*	-.156–.052	.630*
	*M*G |λ| = .629	*M*T |λ| = .200	*M*C |λ| = .060		*M*G |λ| = .765	*M*T |λ| = .305	*M*C |λ| = .079	

Note. Global-λ = standardized loading on Global factor; Target-λ = standardized loading on corresponding specific factor; Cross-λ Range = range of standardized loading on non-corresponding specific factors; δ = uniqueness; *M*G |λ| = absolute average standardized loading on global factor; *M*T |λ| = absolute average standardized loading on corresponding specific factor; *M*C |λ| = absolute average standardized loading on non-corresponding factors; AU = Autonomy; EM = Environmental Mastery; PG = Personal Growth; PR = Positive Relations; PL = Purpose in Life; SA = Self-Acceptance; Reverse-coded items end in R. * $p < .05$

Appendix

Proposed Item Selection for Refined Versions of Ryff's (1989) 9-item Scales of Psychological Well-Being

Item Code	Item Content	Inclusion
AU1	I am not afraid to voice my opinions, even when they are in opposition to the opinions of most people.	AU
EM2	In general, I feel I am in charge of the situation in which I live.	Global
PG3R	I am not interested in activities that will expand my horizons.	–
PR4	Most people see me as loving and affectionate.	–
PL5R	I live life one day at a time and don't really think about the future.	–
SA6	When I look at the story of my life, I am pleased with how things have turned out.	Global
AU7	My decisions are not usually influenced by what everyone else is doing.	–
EM8R	The demands of everyday life often get me down.	Global
PG9R	I don't want to try new ways of doing things--my life is fine the way it is.	–
PR10R	Maintaining close relationships has been difficult and frustrating for me.	Global
PL11R	I tend to focus on the present, because the future nearly always brings me problems.	–
SA12	In general, I feel confident and positive about myself.	Global
AU13R	I tend to worry about what other people think of me.	–
EM14R	I do not fit very well with the people and the community around me.	Global
PG15	I think it is important to have new experiences that challenge how you think about yourself and the world.	–
PR16R	I often feel lonely because I have few close friends with whom to share my concerns.	Global* & PR
PL17R	My daily activities often seem trivial and unimportant to me.	Global
SA18R	I feel like many of the people I know have gotten more out of life than I have.	Global*
AU19	Being happy with myself is more important to me than having others approve of me.	–
EM20	I am quite good at managing the many responsibilities of my daily life.	Global & EM
PG21R	When I think about it, I haven't really improved much as a person over the years.	Global*
PR22	I enjoy personal and mutual conversations with family members or friends.	–
PL23R	I don't have a good sense of what it is I'm trying to accomplish in life.	Global
SA24	I like most aspects of my personality.	Global
AU25R	I tend to be influenced by people with strong opinions.	–
EM26R	I often feel overwhelmed by my responsibilities	–
PG27	I have the sense that I have developed a lot as a person over time.	Global
PR28R	I don't have many people who want to listen when I need to talk.	Global & PR
PL29R	I used to set goals for myself, but that now seems like a waste of time.	Global*

SA30	I made some mistakes in the past, but I feel that all in all everything has worked out for the best.	Global
AU31	I have confidence in my opinions, even if they are contrary to the general consensus.	Global* & AU
EM32	I generally do a good job of taking care of my personal finances and affairs.	–
PG33R	I do not enjoy being in new situations that require me to change my old familiar ways of doing things.	–
PR34R	It seems to me that most other people have more friends than I do.	Global*
PL35	I enjoy making plans for the future and working to make them a reality.	Global
SA36R	In many ways, I feel disappointed about my achievements in life.	Global*
AU37R	It's difficult for me to voice my own opinions on controversial matters.	Global* & AU
EM38	I am good at juggling my time so that I can fit everything in that needs to get done.	Global & EM
PG39	For me, life has been a continuous process of learning, changing, and growth.	Global
PR40	People would describe me as a giving person, willing to share my time with others.	–
PL41	I am an active person in carrying out the plans I set for myself.	Global*
SA42R	My attitude about myself is probably not as positive as most people feel about themselves.	Global
AU43R	I often change my mind about decisions if my friends or family disagree.	AU
EM44R	I have difficulty arranging my life in a way that is satisfying to me.	Global*
PG45R	I gave up trying to make big improvements or changes in my life a long time ago.	Global*
PR46R	I have not experienced many warm and trusting relationships with others.	Global & PR
PL47	Some people wander aimlessly through life, but I am not one of them.	Global
SA48	The past had its ups and downs, but in general, I wouldn't want to change it.	Global
AU49	I judge myself by what I think is important, not by the values of what others think is important.	–
EM50	I have been able to build a home and a lifestyle for myself that is much to my liking.	Global*
PG51R	There is truth to the saying you can't teach an old dog new tricks.	–
VC52	Please choose strongly disagree for this item.	–
PR53	I know that I can trust my friends, and they know they can trust me.	–
PL54R	I sometimes feel as if I've done all there is to do in life.	–
SA55	When I compare myself to friends and acquaintances, it makes me feel good about who I am.	–

Note. AU = Autonomy; EM = Environmental Mastery; PG = Personal Growth; PR = Positive Relations; PL = Purpose in Life; SA = Self-Acceptance; VC = Item used in validity check for data cleaning. Reverse-coded items end in R. Reproduced from Ryff (1989). Inclusion = Indicates item had a loading of at least .40 in both samples on the noted factor; recommended for inclusion in the corresponding composite. * = Items within a scale (e.g., Autonomy) with the highest average loading on the Global factor based on both samples; recommended for inclusion in a short composite of global psychological well-being (two items per scale).

Figure Captions

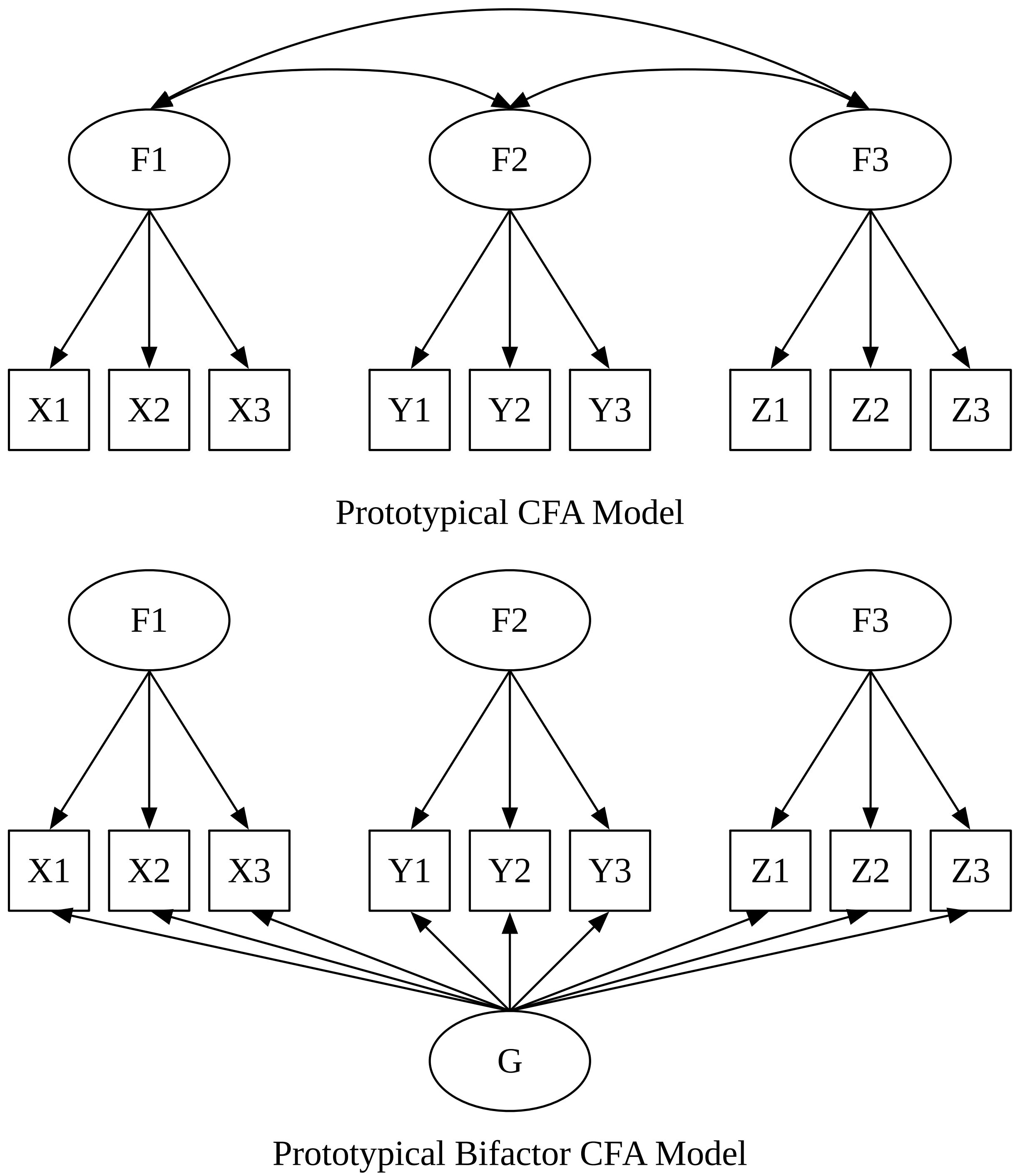
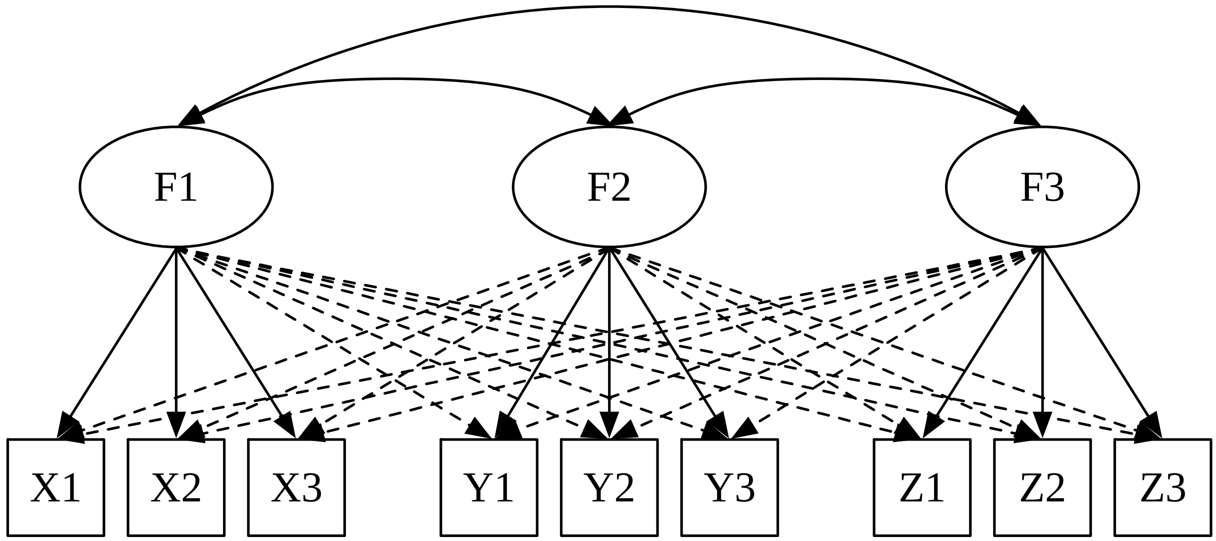
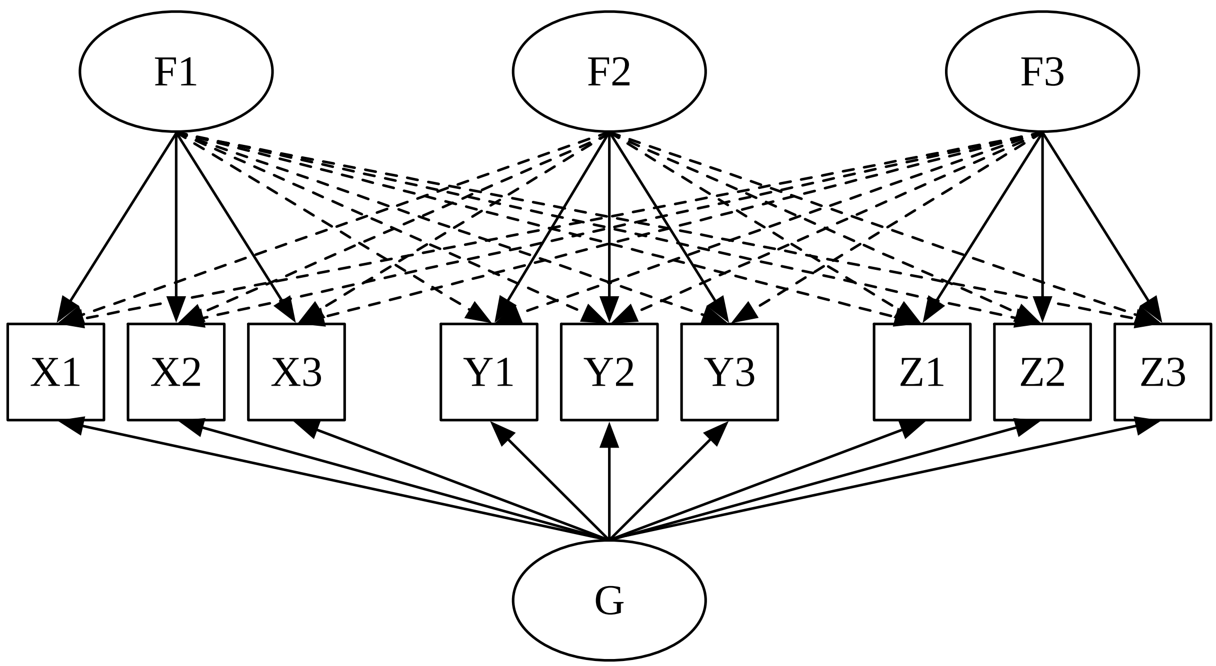


Fig.1 Graphical representation of prototypical confirmatory factor analysis (CFA) models. G = global latent factor; F1–F3 = specific latent factors; X1–X3, Y1–Y3, and Z1–Z3 = items. Solid arrows indicate factor loadings. Variances are omitted for clarity.



Prototypical ESEM Model



Prototypical Bifactor ESEM Model

Fig.2 Graphical representation of prototypical exploratory structural equation modelling (ESEM) models. G = global latent factor; F1-–F3 = specific latent factors; X1–X3, Y1–Y3, and Z1–Z3 = items. Solid arrows indicate target factor loadings. Dotted arrows indicate cross-loadings. Variances are omitted for clarity.