Soc Indic Res (2009) 93:359–375
DOI 10.1007/s11205-008-9329-1

Investigating the Structural Validity of Ryff's Psychological Well-Being Scales Across Two Samples

Richard A. Burns · M. Anthony Machin

Accepted: 9 September 2008 / Published online: 25 September 2008

Abstract Ryff's (1989b) Psychological Well-Being (PWB) scales measure six related constructs of human functioning. The present paper examined the validity of Ryff's 6-factor PWB model, using data from a life events study ($N = 401$) and an organisational climate study ($N = 679$). Previous validation studies, using Confirmatory Factor Analysis (CFA), have identified alternative PWB models, but limitations include the use of shorter scale versions with items relating to a number of life domains within the same PWB factor, and failure to examine the influence of participants' socio-demographic characteristics on PWB. In this study, Exploratory Factor Analysis (EFA) mostly found consistency in the PWB items and structure between the two studies whereby a 3-factor model delineated between items relating to Autonomy, Positive Relations and a super-ordinate factor comprising the other PWB factors. Using CFA, Goodness of Fit indices reached acceptable levels for the adjusted PWB model identified by the EFA, whilst differences between adjusted models of PWB previously identified in the literature were hardly evident. Post-hoc analysis by gender demonstrated socio-demographic effects on the structure and items that comprise PWB. Further development of PWB measures is needed to reflect its hierarchical and multi-dimensional nature. In the scales' current form, the construct validation of the PWB factors will continue to be problematic and will fail to adequately evaluate the nature and impact of PWB.

Keywords Well-being · Psychological Well-Being · Subjective Well-Being · Structural validity · Well-being measurement

Decades of research have related notions of positive mental and physical health with the absence of such adverse states as depression, anxiety, and physical illness. However, a number of researchers have proposed that well-being is not necessarily the antithesis to these constructs of ill-being (Kahneman 1999; Ryff 1989a). A major challenge is to

R. A. Burns (✉) · M. A. Machin
Department of Psychology, University of Southern Queensland, Toowoomba, QLD, Australia
e-mail: burns@usq.edu.au

identify relevant dimensions of well-being and to understand how these dimensions are shaped by various human experiences.

Ryan and Deci (2001) have described two distinct, yet related approaches with which most psychological theories of well-being could be aligned. The Hedonic, or Subjective Well-Being (SWB) approach, focused on immediate human functioning and experience, and was associated with perceptions of pleasure, displeasure, satisfaction, and happiness. With philosophical roots in antiquity and more recently in Priestley and Bentham's 'greatest happiness principle', models of SWB have perhaps been the most frequently reported on within the well-being literature (Diener et al. 1999). Alternatively, a Eudaimonic or Psychological Well-Being (PWB) model emphasises those mechanisms that are associated with healthy human functioning and adjustment. Whilst daily SWB fluctuates with life experiences (Headey 2000; Headey and Wearing 1989), PWB is a relatively stable construct that captures those aspects of human functioning more likely to lead to adaptive human functioning and positive experiences (Ryan and Deci 2001). Such theories are not new. As with SWB, PWB is grounded in ancient philosophical works (e.g., Epictetus) whilst more recent proponents have included the existential writings of Kierkegaarde and Tillich, both of whose discourses into the experience of melancholia and anxiety led to the conclusion that the good life is one not free of 'angst', but one that is lived in spite of it.

Ryff's (1989b) Psychological Well-Being model drew on gerontological and life-span research and reflects one construct-oriented approach to PWB. Its theoretical underpinnings stemmed from a wide range of influences including Allport's (1961) concept of the mature personality, Rogers' (1961) fully-functioning individual, and Maslow's (1968) self-actualisation, and led to the formulation of six dimensions of PWB: Autonomy, Positive Relations with Others, Environmental Mastery, Personal Growth, Purpose In Life, and Self-Acceptance (Ryff 1989a, b). With intuitive appeal and widespread interest, the use of the PWB model, its domains and items, have been applied to a number of different applied psychological areas (Clarke et al. 2001; Fava et al. 2004), despite unresolved questions relating to its validity (e.g., Springer and Hauser 2006; Springer et al. 2006). Also, Keyes et al. (2002) have identified relationships between PWB and SWB variables which raises questions about the degree to which PWB and SWB are distinct constructs. We propose to investigate the dimensionality of PWB in order to allow a better understanding of how PWB relates to SWB.

Abbot et al.'s (2006) recent review noted that most psychometric analyses of the PWB scales occurred almost a decade after the first publication of the PWB scales (Ryff 1989b) with a number of different findings being reported. Whilst the 'a priori' correlated 6-factor model has received some support (e.g., Ryff and Keyes 1995), a number of studies (e.g., Clarke et al. 2001) have indicated a high degree of correlation between four of the PWB variables: Environmental Mastery (E), Personal Growth (G), Purpose In Life (P), Self-Acceptance (S) (EGPS), such as to warrant analysing these factors as one super-ordinate factor. Further analyses of the PWB scales have supported this structure with separate first order factors for Autonomy and Positive Relations with Others, and one-second order factor containing the EGPS variables (Abbot et al. 2006).

1 Limitations of Existing Studies

A number of the studies reviewing the structure of the PWB constructs (e.g., Kafka and Kozma 2002; Ryff and Keyes 1995; Van Dierendonck et al. 2007) have been fraught with some methodological limitations. Whilst the original model (Ryff 1989b) included 120

items, shorter versions have included 84, 54, 42 and 18 items, generally with equal numbers of items per PWB variable. Most analyses have tested the factorial validity of the PWB model with the smallest 18-item (three items per variable) scale, though two recent analyses used a 42-item scale (Abbot et al. 2006; Springer and Hauser 2006) or an amended 39-item scale (Van Dierendonck et al. 2007). There are issues relating to the validity of these findings since there is a lack of consistency in the items that comprise the shorter versions of the PWB scales. Whilst the 84-item version comprises all items used in the 54-item scales, there is considerably less overlap in the items used between the shorter versions, with only six common items between the 18- and 42-item scales. Van Dierendonck (2004) analysed the 84, 54, and 18 item scale versions and found support for a 6-factor model with a second-order PWB factor. Although internal consistencies were high, Goodness of Fit Indices (GFI) indicated poor fit for the two larger scales.

Further issues relate to the methodology employed in developing the PWB scales. Initial development of the original 120-item version is explained fully elsewhere (Ryff 1989b; Ryff and Singer 2006), but in summary, an initial pool of some 80-items per variable were reduced to 32 items per variable. Ryff (1989b) then analysed the bi-variate correlations of items to their respective variable and retained items (20 per variable) with the strongest correlations, as long as an item's strongest correlation was reported between the item and its parent variable. Even so, this process means that some items which scored most highly on their respective variables will likely fail to discriminate between other variables if they also reported lesser but still very strong correlations with other variables. This process certainly explains why high correlations (e.g., Van Dierendonck et al. 2007) and cross-loading of items across PWB variables (e.g., Springer and Hauser 2006) have been reported, and why internal consistency of the PWB variables is often quite high (e.g., Ryff 1989b).

Most published factor analyses of the Ryff PWB scales have used Confirmatory Factor Analysis (CFA) though one study (Kafka and Kozma 2002) used an Exploratory Factor Analysis (EFA) approach to assess the dimensionality of the original 120-item version, and supported a one general PWB factor. However, the authors first extracted all factors with eigenvalues greater than one and then the 'a priori' 6-factor model using principal components analysis (PCA), with an orthogonal (Variamax) rotation, a process which is generally described as a data reduction process. It is not surprising that most of the PWB items loaded onto the first factor. Given the frequently reported high degree of correlation between the PWB variables, a Principal Axis Factoring (PAF) method with an oblique rotation would seem most appropriate to identify a correlated PWB factor structure. A re-analysis of the original item pool with PAF, using an oblique rotation, should provide a more coherent and defensible set of dimensions.

We propose that the longer scales, at least the 84 and 54 item scale versions, be used to test the validity of Ryff's 6-factor structure of PWB. A significant amount of meaningful data is lost when only three or seven of the original 20 items per variable are used in the data collection as it is likely that the influence of sample characteristics, like gender (Marks and Lambert 1998), age (Ryff and Keyes 1995), and culture (Ryff et al. 2004), all of which have been demonstrated to have some effect on PWB, will be reflected on the PWB factor structure, particularly when using a smaller item pool. It may be that sampling characteristics influence particular response patterns to items of different content. Given these effects, the validity of Ryff's (1989b) original development of the model must also be considered with caution since 60% of the original sample ($N = 321$) were female, and the sample was stratified by three age groups.

However, despite the weaknesses related to certain aspects of the scales' initial construction, and the limitations of some subsequent analyses, considerable evidence (Ryff

and Singer 2006) does relate PWB to a range of outcomes including biological indicators (Ryff et al. 2004), successful transitions in later life (Smider et al. 1996) and better counselling interventions (Fava et al. 2004), supporting the utility of the construct and its operationalisation using Ryff's PWB model.

2 Aims of the Current Study

The current paper seeks to test the factor structure of the PWB model with two of the larger scale versions (84- and 54-items), on two separate studies: a life events study with an Australian community sample ($N = 401$) and a cross-national organisational climate study with teachers ($N = 679$). Unlike previous validation studies, we seek to identify a stable structure underlying Ryff's (1989b) model of PWB using EFA and to determine whether this structure is consistent across our studies. In addition, the availability of a SWB measure in each of these studies allows us to test associations between measures of PWB and SWB across our studies. Finally, using just those items identified in the EFA, a CFA will compare a range of Goodness of Fit Indices (GFI) for the following models: (a) the 'a priori' 6-factor correlated model (Ryff 1989b), (b) a 1-factor model (Kafka and Kozma 2002), (c) the structure identified in the EFA, and (d) a model that combines four of the PWB variables: Environmental Mastery, Personal Growth, Purpose In Life, and Self Acceptance as a second-order factor (Abbot et al. 2006). These models will be tested with two adjustments previously tested in the PWB literature: the inclusion of two method variables (Abbot et al. 2006), and the inclusion of correlated error terms (Springer and Hauser 2006).

3 Method

3.1 Participants and Design

Data from two studies were used for the analyses in this paper. Study 1 was a Life Events Study ($N = 401$) comprising undergraduate students from the Department of Psychology at the University of Southern Queensland (USQ). Participation in departmental projects is a requirement of enrolment in some psychology courses. Participants were predominantly female (83%), and unlike most undergraduate populations, studied part-time (55%) with a relatively equal distribution from late teens to late forties. These sampling characteristics can be attributed to the provision of unique educational services by several universities in Australia, like USQ, which recognise that many do not necessarily follow the traditional route of entering university within a year or so of having completed their high school qualification. With the impediments (e.g., family and work responsibilities) associated with entering higher education later in life, USQ provides opportunities for students to undertake most of their courses on a part-time and external basis, in addition to the traditional full-time and on-campus modes.

Study 2 was an organisational climate study ($N = 679$) comprising three samples of schoolteachers, drawn from privately-funded schools in the Australian Capital Territory, Australia ($n = 253$), school teacher members of the Norwegian teacher union ($n = 250$), and from schools worldwide which designated themselves as being International Schools ($n = 176$). Predominantly female (63%), most participants (46.2%) were aged between 30 and 55 years of age, though 63.2% of the Norwegian sample was aged 45 years and older.

3.2 Procedure

Both studies included two measures of well-being, the data from which were analysed for this study. Ryff's (1989b) Psychological Well-Being scales assesses six dimensions of PWB: Environmental Mastery (E), Personal Growth (G), Purpose In Life (P), Self-Acceptance (S), (EGPS); Autonomy (A); and Positive Relations (PR). An 84-item version was used in Study 1 and a 54-item version was used in Study 2. Unlike shorter versions of the PWB scale, all the items of the 54-item version are included in the larger 84-item version. Therefore, items from the 54-item version were extracted from the 84-item version used in the life events study, to allow for comparison with the 54-item version that was used in the organisational climate study. Individuals indicated their response on a 6-point Likert-type scale, with higher scores on each scale indicating greater well-being on each dimension. The PANAS (Watson et al. 1988) assessed SWB with 20-items relating to positive affect (PA) and negative affect (NA), and was assessed in both studies. Individuals indicated their response on a 5-point Likert-type scale, with higher scores on each scale indicating greater well-being on that dimension.

Both studies were undertaken between June 2006 and June 2007. A high number of participants in both studies did not live in the immediate vicinity of the university, so therefore participants accessed the survey through a secure web facility which is run and monitored by the technical services staff within the Department of Psychology. The University's Human Research Ethics Committee provided approval for both studies.

Analyses were undertaken separately for both studies and then compared. In addition, post-hoc analyses analysed cohort effects in the organisational climate study analysis, and gender effects for both the life events and organisational climate studies.

4 Results

Principal axis factoring with an oblique rotation attempted to differentiate between PWB items. Parallel analysis (O'Connor 2000) was first used to identify the number of factors to be extracted for factoring. For each study, the parallel analyses indicated extracting between 6 and 9 factors, however convergence failed when extracting either 6, 7, 8, or 9 factors. Extracting 4 or 5 factors led to a number of items loading across more than one factor, so items were deleted from the analysis if they loaded above 0.30 on more than one factor, or failed to achieve this level on one factor. For both studies, items now loaded onto three factors: Autonomy (A); Positive Relations (PR); and a first-order factor (EGPS), comprising items relating to Environmental Mastery (E), Personal Growth (G), Purpose In Life (P), and Self-Acceptance (S). This supports Abbot et al.'s (2006) findings of a higher order factor, EGPS, however, in this instance, our EGPS variable reflected a first-order factor. The items and their respective factor loadings are displayed in Table 1 for each study. Inspection of the item loadings reveal mostly moderate loadings and indicate a fair degree of consistency in the items that load onto their respective factor. Few differences in the size of loading scores and in those items identified as significant indicators, are reported.

Further analyses sought to differentiate between PWB and SWB. Therefore, the aforementioned analyses were extended to include 20 items from the PANAS scale. Principal Axis Factoring and an oblique rotation clearly delineated between the three PWB factors and two SWB factors: PA and NA, for both studies. Some cross-loading of the SWB items on PWB factors did occur, but all remained below the 0.30 criterion cut-off.

Table 1 A comparison of the item loadings of the 54 item PWB scale by study

PWB variable	Item #	Life events study			Teacher study*		
		Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Environmental Mastery	02	0.446	–	–	–	–	–
	07	–	–	–	–	–	–
	12	–	–	–	–	–	–
	17	0.706	–	–	0.640	–	–
	20	–	–	–	–	–	–
	29	0.540	–	–	0.408	–	–
	36	0.643	–	–	0.477	–	–
	49	–	–	–	–	–	–
	53	–	–	–	0.520	–	–
Personal Growth	03	–	–	–	–	–	–
	18	–	–	–	–	–	–
	21	0.487	–	–	0.322	–	–
	26	–	–	–	–	–	–
	37	0.776	–	–	0.581	–	–
	41	–	–	–	–	–	–
	45	0.669	–	–	0.458	–	–
	50	0.546	–	–	0.419	–	–
	54	–	–	–	–	–	–
Purpose In Life	8	–	–	–	–	–	–
	13	0.373	–	–	–	–	–
	22	–	–	–	–	–	–
	27	0.444	–	–	–	–	–
	30	0.604	–	–	–	–	–
	33	0.806	–	–	0.511	–	–
	38	0.838	–	–	0.693	–	–
	42	0.630	–	–	0.483	–	–
	46	–	–	–	–	–	–
Self-Acceptance	4	–	–	–	–	–	–
	9	–	–	–	–	–	–
	14	–	–	–	–	–	–
	23	0.395	–	–	0.411	–	–
	28	–	–	–	0.368	–	–
	31	–	–	–	–	–	–
	43	–	–	–	–	–	–
	48	–	–	–	–	–	–
	51	0.460	–	–	0.486	–	–

Table 1 continued

PWB variable	Item #	Life events study			Teacher study*		
		Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
	1	–	–	–	–	–	–
	5	–	0.649	–	–	0.527	–
	10	–	0.747	–	–	0.707	–
	15	–	–	–	–	–	–
Positive Relations	24	–	0.620	–	–	0.754	–
	32	–	0.739	–	–	0.595	–
	34	–	–	–	–	–	–
	39	–	0.728	–	–	0.501	–
	47	–	0.527	–	–	–	–
	6	–	–	–	–	–	0.402
	11	–	–	0.618	–	–	0.309
	16	–	–	0.520	–	–	0.480
	19	–	–	–	–	–	–
Autonomy	25	–	–	0.638	–	–	0.469
	35	–	–	–	–	–	–
	40	–	–	0.501	–	–	0.671
	44	–	–	0.658	–	–	0.434
	52	–	–	0.426	–	–	–

* Teacher study was assessed using Wave 1 data only

'–' Item either cross-loaded onto more than one factor or loaded weakly (<0.30) onto one factor $ *Italics* indicate negatively worded items

Table 2 A comparison of the correlations* between PWB and SWB variables by Study

		Life events study				
		1	2	3	4	5
1. Positive affect	Combined school sample	1	−0.243	0.638	0.317	0.433
2. Negative affect		−0.224	1	−0.405	−0.453	−0.362
3. EGPS[a]		0.589	−0.237	1	0.361	0.527
4. Positive Relations		0.182	−0.386	0.314	1	0.306
5. Autonomy		0.250	−0.293	0.305	0.247	1

* All correlations significant at $p < 0.001$, [a] EGPS comprises items relating to Environmental Mastery, Personal Growth, Purpose In Life, and Self-Acceptance

Although Factor Analysis differentiated between PWB and SWB items, bi-variate correlations (Table 2) indicated mostly moderate to strong correlations between all the well-being factors, with findings generally consistent between the studies. Few differences between studies were reported and related only to the size of association. All correlations were significant ($p < 0.001$).

Confirmatory Factor Analysis (CFA, Table 3) of the items identified in the initial Exploratory Factor Analysis (EFA, Table 1) was undertaken to assess whether the PWB model identified in the EFA reported better Goodness of Fit Indices (GFI) than the

Table 3 Confirmatory factor analysis by study testing several structural models of PWB using the items identified by exploratory factor analysis

Model	Study	C^{a}_{min}	Df	GFI	CFI	RMSEA (95% CI)
Model 1	Life events study	708.32	284	0.877	0.889	0.061(0.055–0.067)
	Combined teacher samples	1010.14	260	0.891	0.808	0.065(0.061–0.070)
Model 2	Life events study	1721.70	300	0.690	0.630	0.109(0.104–0.114)
	Combined teacher samples	1980.81	275	0.778	0.570	0.095(0.091–0.099)
Model 3	Life events study	897.41	296	0.842	0.843	0.071(0.066–0.077)
	Combined teacher samples	1088.19	272	0.882	0.792	0.067(0.062–0.071)
Model 4	Life events study	769.65	292	0.867	0.876	0.064(0.058–0.069)
	Combined teacher samples	1036.65	268	0.898	0.804	0.065(0.061–0.069)
Model 5	Life events study	477.77	257	0.918	0.943	0.046(0.040–0.053)
	Combined teacher samples	820.70	237	0.913	0.851	0.060(0.056–0.065)
Model 6	Life events study	924.38	272	0.829	0.830	0.077(0.072–0.083)
	Combined teacher samples	1085.57	249	0.884	0.786	0.070(0.066–0.075)
Model 7	Life events study	644.04	269	0.883	0.902	0.059(0.053–0.065)
	Combined teacher samples	905.90	248	0.902	0.832	0.063(0.058–0.067)
Model 8	Life events study	548.34	265	0.905	0.926	0.052(0.046–0.058)
	Combined teacher samples	876.00	245	0.906	0.839	0.062(0.057–0.066)
Model 9	Life events study	471.49	258	0.920	0.944	0.045(0.039–0.052)
	Combined teacher samples	547.30	222	0.940	0.917	0.046(0.042–0.051)
Model 10	Life events study	444.87	241	0.923	0.947	0.046(0.039–0.053)
	Combined teacher samples	489.27	216	0.946	0.930	0.043(038–0.048)
Model 11	Life events study	511.27	263	0.913	0.935	0.049(0.042–0.055)
	Combined teacher samples	536.79	232	0.942	0.922	0.044(0.031–0.049)
Model 12	Life events study	496.50	265	0.916	0.940	0.047(0.040–0.053)
	Combined teacher samples	559.637	232	0.939	0.916	0.046(0.041–0.050)

Model 1, 'a priori' six correlated factors; Model 2, 1 PWB Factor; Model 3, EGPS 1st order factor correlated with A and PR; Model 4, EGPS 2nd order factor correlated with A and PR; Model 5 = Model 1 with correlated method variables; Model 6 = Model 2 with correlated method variables; Model 7 = Model 3 with correlated method variables; Model 8 = Model 4 with correlated method variables; Model 9 = Model 1 with significant correlated error terms; Model 10 = Model 2 with significant correlated error terms; Model 11 = Model 3 with significant correlated error terms; Model 12 = Model 4 with significant correlated error terms. [a]All chi square statistics were significant $p = 0.000$

'a priori' 6-factor model, and a number of alternative models identified in the literature. Four main models were tested: Model 1 tested the 'a priori' correlated 6-factor model (Ryff 1989b); Model 2 tested a general 1-factor model (Kafka and Kozma 2002); Model 3 tested for results identified in the EFA reported earlier in this paper whereby a first-order factor (EGPS) comprised items relating to Environmental Mastery, Personal Growth, Purpose In Life, Self-Acceptance; and Model 4 replicated previous findings (Abbot et al. 2006) which identified EGPS as a second-order factor. CFA analyses were performed using the items identified for each study from the original EFA.

In addition to these four main models, we tested additional methodological factors, which have also been identified in the literature. For example, Springer and Hauser (2006) and Abbot et al. (2006) introduced several adjustments to test for methodological effects. Springer and Hauser introduced a latent variable to account for reverse-scored items, which they found significantly improved fit. Abbot et al. found strong support for the introduction

of two method factors reflecting positive (non-reversed scores) and negative (reversed scores) method factors. Springer and Hauser found further support for methodological effects by correlating the error terms of adjacent items and items with similar content. The inclusion of these paths is not usually recommended unless there is a strong theoretical basis for doing so, such as when item content is similar, when there is a likelihood of social response bias/desirability, where a model omits the inclusion of an exogenous variable, and in repeated measures designs where items are measured on two or more occasions (Aish and Jöreskog 1990; Byrne 2001). Based on these findings, we tested the effect of including two method factors, as well as significant error covariances. However, we did not believe testing for correlated adjacent items was warranted since the structure of the PWB scales includes intermittent use of items that require reverse scoring, negatively and positively phrased items, as well the systematic ordering of items so that no item from the same variable is placed adjacent to each other. Our rationale for including significant correlated error terms assumes that Springer and Hauser's findings reflect other artifact such as response bias, which is common in attitude surveys and when items are similar in content. These additional effects were tested in Models 5–8 (two method factors) and Models 9–12 (significant error covariances).

In models 1–4, the pattern of findings was identical across both studies for all models. The six-factor model (Model 1) was clearly a better fitting model than the single factor model and those models with the first and second order EGPS factor, however GFI were far from acceptable. Models 5–8, which included the two latent method variables, performed better than Model 1, though the six-factor model with the additional two method factors (Model 5), performed better than either of the other models. Models 9–12 tested the effect of including significant covariances between correlated error terms. Positive covariances were included if they reported Modification Index values above 4, and if the association was significant ($p = 0.05$). All four of these models performed better than previous models though differences in GFI between models 9 and 12 were less apparent with different GFI indicating different models as best fitting. Many of the models that tested the method variables consisted of paths, between the method variables and the items, that failed to achieve significance ($p > 0.05$), whilst the models with error covariances only included significant associations. The significant correlated error terms included in the analyses varied between the studies and this may reflect differences between participants, where socio-demographic characteristics might be related to different PWB items.

It was surprising that the model with the first order EGPS did not report the best fit considering this was the factor identified in the EFA and that only those items identified in the EFA for each sample were included in the CFA. The authors have previously suggested that sampling characteristics, in particular age and gender, may influence the structure of PWB and this may reflect differences between participants. Unfortunately, any post-hoc analysis of the factor structure by a number of socio-demographic variables was limited by the design of the studies and by the variables that were operationalised for the original purpose of each study. This precluded post-hoc analyses of a number of socio-demographic effects on PWB, such as age. For example, age groupings were not comparable between the Life Events study and the Organisational Climate study as they had been designed to reflect the age range of the targeted study participants. Furthermore, there was a preponderance of young to middle aged adults in the life events study, and middle-aged to late middle-aged adults in the organisational climate study which precluded a sub-groups analysis of PWB by age within each study. However, a sub-groups analysis of the original items (Table 1) by gender (Table 4) for both studies, and by cohort in the organisational climate study (Table 5), was possible. Both of these analyses found some support for the PWB structure

Table 4 A comparison of the item loadings of PWB by gender and by study of the items extracted from the original EFA

PWB variable	Item no.	Teacher study ($N = 679$)						Life events study ($N = 401$)					
		Male ($n = 252$)			Female ($n = 427$)			Male ($n = 68$)			Female ($n = 333$)		
		Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Environmental Mastery	02	–	–	–	–	–	–	0.531	–	–	0.440	–	–
	17	0.690	–	–	0.629	–	–	0.672	–	–	0.634	–	–
	29	0.586	–	–	0.304	–	–	0.453	–	–	0.545	–	–
	36	0.432	–	–	0.418	–	–	0.594	–	–	0.557	–	–
	53	0.595	0.335	–	0.522	–	–	0.540	0.378	–	0.432	–	–
Personal Growth	21	–	–	0.412	0.317	–	–	0.557	–	–	0.436	–	–
	26	–	–	–	–	–	–	0.319	–	–	0.402	–	–
	37	0.588	–	–	0.586	–	–	0.837	–	–	0.741	–	–
	45	0.626	–	–	0.314	–	–	0.726	–	–	0.613	–	–
	50	0.535	–	–	0.338	–	–	0.455	–	0.402	0.557	–	–
Purpose In Life	13	–	–	–	–	–	–	–	0.412	0.430	0.520	–	–
	27	–	–	–	–	–	–	–	–	0.380	0.466	–	–
	30	–	–	–	–	–	–	0.343	–	0.473	0.695	–	–
	33	0.413	0.379	–	0.424	–	–	0.742	–	–	0.777	–	–
	38	0.768	–	–	0.613	–	–	–	–	–	–	–	–
	42	0.420	–	–	0.469	–	–	0.592	–	–	0.586	–	–
	46	–	–	–	–	–	–	–	–	0.306	–	–	–
Self-Acceptance	23	0.414	–	–	0.374	–	–	0.599	–	–	0.341	–	0.311
	28	0.436	–	–	0.416	–	–	0.536	–	–	0.373	–	–
	51	0.400	–	0.492	0.491	–	–	0.529	–	–	0.468	–	–

Table 4 continued

PWB variable	Item no.	Teacher study ($N = 679$)						Life events study ($N = 401$)					
		Male ($n = 252$)			Female ($n = 427$)			Male ($n = 68$)			Female ($n = 333$)		
		Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Positive Relations	05	–	0.575	–	–	−0.478	–	–	0.836	–	–	0.604	–
	10	–	0.694	–	–	−0.766	–	–	0.734	–	–	0.753	–
	24	–	0.643	–	–	−0.768	–	–	0.623	–	–	0.604	–
	32	–	0.311	0.414	–	−0.659	–	–	0.693	–	–	0.747	–
	39	–	–	–	–	−0.547	–	–	0.596	–	–	0.742	–
	47	–	–	–	–	–	–	0.487	0.493	−0.312	–	0.523	–
Autonomy	06	−0.344	0.346	0.785	–	–	0.394	0.444	–	–	–	–	0.551
	11	–	–	–	–	–	0.387	–	−0.302	0.452	–	–	0.662
	16	–	–	0.445	–	–	0.504	–	0.323	0.317	–	–	0.540
	25	–	–	0.502	–	–	0.467	–	–	0.770	–	–	0.569
	40	–	–	0.301	–	–	0.720	–	–	0.545	–	–	0.490
	44	–	–	0.325	–	–	0.447	–	–	0.772	–	–	0.587
	52	–	–	–	–	–	–	–	–	0.475	–	–	0.423

'–' Item either cross-loaded onto more than one factor or loaded weakly (<0.30) onto one factor

Table 5 A comparison of the item loadings of PWB by teacher cohort of the items extracted from the original EFA

PWB variable	Item no.	Teacher study ($N = 679$)								
		International teacher cohort ($n = 176$)			Norwegian teacher cohort ($n = 250$)			Australian teacher cohort ($n = 253$)		
		Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Environmental Mastery	17	0.672	–	–	0.448	0.303	–	0.704	–	–
	29	0.646	–	–	–	–	–	0.528	–	–
	36	0.593	–	–	–	–	–	0.422	–	–
	53	0.549	–	–	0.553	–	–	0.563	–	–
Personal Growth	21	0.448	–	–	–	–	–	0.390	–	–
	37	0.674	–	–	0.333	–	–	0.478	–	–
	45	0.556	–	–	0.435	–	–	0.360	–	–
	50	0.370	–	0.314	0.355	–	−0.385	0.450	–	0.330
Purpose In Life	33	0.571	0.334	–	0.453	–	–	0.334	0.361	–
	38	0.807	–	–	0.545	0.413	–	0.551	–	–
	42	0.559	–	–		–	–	0.441	–	–
Self-Acceptance	23	0.585	–	–	0.468	–	–	0.353	–	–
	28	0.539	–	–		–	–	0.407	–	–
	51	0.557	–	–	0.538	–	–	0.534	–	–
Positive Relations	05	–	0.583	–	0.374	–	–	–	0.605	–
	10	–	0.717	–	–	0.587	–	–	0.657	–
	24	–	0.691	–	–	0.606	–	–	0.657	–
	32	–	0.587	–	–	0.544	–	–	0.505	–
	39	–	0.486	–	–	0.533	–	–	0.406	–
Autonomy	06	–	–	0.359	0.304	–	0.348	–	–	0.414
	11	–	–	0.345	–	–	–	–	–	0.493
	16	–	–	0.496	0.409	–	0.365	–	–	0.526
	25	–	–	0.446	–	–	0.370	–	–	0.440
	40	–	–	0.796	0.351	–	0.328	–	–	0.780
	44	–	–	0.350	–	–	0.311	–	–	0.437

'–' Item loaded weakly (<0.30) onto factor

reported in the initial EFA findings. In particular, the results for the Australian and International teacher cohorts were considerably similar to the findings of the original EFA. However, analysis by the different teacher cohort is still likely to demonstrate the effect of participant characteristics and these findings need to be considered in this light.

Not all items were consistently reported with equal loading by gender and cohort, and several items cross-loaded onto other factors. This may explain why the factor structure with a first-order EGPS variable did not outperform the alternative and 'a priori' models. Whilst females reported items with factor loadings that more closely mirrored the overall results for both samples, this is quite likely a consequence of the greater proportion of

females in both studies. However, consistent differences between males and females for several items (e.g., Environmental Mastery item 53, and Autonomy item 16) between studies, does support the notion that perhaps there are differences between gender on items that comprise PWB.

5 Discussion

In two studies, PAF with oblique rotation delineated three PWB variables: Autonomy, Positive Relations, and EGPS, a first-order factor first previously identified as a second-order factor by Abbot et al. (2006), and comprising the Environmental Mastery, Personal Growth, Purpose In Life and Self Acceptance items. The inclusion of a SWB measure identified two SWB factors: Positive Affect and Negative Affect, which were distinct from the PWB variables, although significant correlations between all the SWB and PWB variables were reported. Whilst some differences between studies in the items constructing the PWB variables were reported, the structure of PWB was consistent between the studies. Post-hoc analysis of the different teacher cohorts within the organisational climate study and by gender for both studies, revealed that sampling characteristics appear to influence both the structure and items that comprise PWB.

These results support a number of previous findings which have postulated either a simple 1-factor model, a correlated 6-factor model, as well as first or second-order factors which incorporated the EGPS variables reported in this study and elsewhere (Abbot et al. 2006). Initial GFI of the unmodified models were poor, though the six-factor model was the preferred model. Two types of adjustments were assessed and included the addition of method factors (Springer and Hauser 2006; Abbot et al. 2006) and reported much better fit. Despite some concern about the methodological, theoretical and statistical implications, a second adjustment expanded on previous findings (Springer and Hauser 2006) which allowed for correlated error terms. Results demonstrated acceptable and comparable fit for all four models where significant paths between correlated error terms were included.

Based on our findings we recommend the use of EFA techniques with larger item pools, and to remove less important items, or items that are related to more than one factor, in subsequent analysis. This would appear to be a happy medium where larger scales improve internal consistency and shorter scales that are suited to factor analysis (Van Dienrendonck 2004). We recognise that the items and structure of PWB will reflect particular characteristics of the sample, but would hypothesise that a larger item pool will increase the likelihood of identifying a consistent structure to the PWB model, though gender, age and other socio-demographic effects on the structure of PWB are to be expected.

In relation to Ryff's (1989b) original model development, Principal Axis Factoring, using an oblique rotation may have proved a more fruitful methodological approach. Whilst the final item pool may have resulted in a multi-dimensional model of PWB, a correlational approach fails to consider item content which enables greater differentiation between highly related constructs, rather leading to the inclusion of items that differ in the extent to which they assess specific versus general judgements of well-being. For example, the Environmental Mastery items cover a wide range of areas of personal control, from control of daily life responsibilities (item #17), to control of time and demands (item #36), to control of personal finances (item #29) and participants' responses will surely reflect the importance of each particular issue for people at different ages or stages in their lives. Items that comprised the Personal Growth variable also reflect a mixture of items that relate to one's personal growth through life to date (item #s 37, 45 and 50), or reflects on

the prospect of continuing to face the challenges to one's growth and development (item 21). Clearly, people of different ages, who are at different stages of their lives, may relate to these questions in quite different ways. For instance, older participants may relate more easily to the reflective questions, whilst younger participants may, in comparison, have shorter temporal contexts within which to reflect on such issues. In contrast, findings for the future-oriented questions may be more important for younger participants.

Such issues are not new to models of self-referent beliefs and attitudes. Decades of research into self-concept failed to consider the implications of generating items within scales that comprise a mixture of items whose content fail to distinguish between global and context-specific judgments. It wasn't until reviews of the existing self-concept measures of the day (e.g., Wylie 1974; Burns 1979), that these weaknesses in self-concept surveys were summarily identified, revealing that they failed to address these very same issues. Consequently, Shavelson et al. (1976) proposed a multidimensional and hierarchical nature of self-concept that reflected this structure, whilst Marsh's (1992) construction of the Self Description Questionnaires (SDQ) operationalised Shavelson et al.'s model, and has since identified the utility of a multi-dimensional and hierarchical structure to self-referent beliefs.

In a similar vein, we would propose a model of well-being that is both multi-dimensional and hierarchical in nature. Whilst well-being variables at a higher order level may certainly have differential predictions on a number of outcomes, we believe that further development of well-being models are needed and must consider the domain and level of specificity that is being assessed. Such a model would incorporate the eudaimonic processes that Ryff has sought to address at a general level, the hedonic states captured by SWB measures (e.g., PANAS), as well as physical and biological health correlates. Support for such a model has previously been indicated (Keyes et al. 2002) where analysis delineated distinct yet related associations between PWB and SWB variables, a finding supported in this study whereby factor analysis of PWB and SWB items delineated differences at the item level, but with high correlations between the variables indicating highly related constructs. We would also propose that the relationship between these different dimensions is reciprocal, though stronger causal paths from PWB to SWB may be expected, and that the strength of this reciprocal nature is reflected by the level of the hierarchy at which the association is investigated.

More recent analyses of the PWB scales (Abbot et al. 2006; Springer and Hauser 2006; Van Dierendonck et al. 2007) have used PRELIS or MPlus to provide polychoric correlation estimates as previous methods have perhaps incorrectly assumed PWB responses to reflect continuous data, which can bias estimates. However, it is common to assume that Likert scales that consist of at least 5 points can be analysed as if reflecting a continuous scale (Dollan 1994). Still, regardless of sample size (Jöreskog and Sörbom 1996), the use of Pearson correlation matrices in Factor Analysis appears to underestimate the degree of association between variables and consequently results in reduced factor loadings (DiStefano 2002).

Approaches that use polychoric correlations may be warranted in some circumstances, but such techniques may only prove to be more stringent. Whilst computing polychoric and tetrachoric matrices is certainly possible in some statistical packages, its use in personality research, for example, where scales frequently comprise Likert scales similar to the PWB scales and where most factor analysis has typically used bivariate correlation matrices, has indicated that this approach generally fails to produce dissimilar results from traditional methods. Holgado-Tello et al. (2007) tested the veracity of the Five-Factor Personality Model using polychoric estimates and concluded that the polychoric estimation approach produced results comparable to previous non-polychoric approaches. It is perhaps for these

reasons that so few commercially available statistical packages allow for these sorts of techniques. Since the use of such methods in personality research has contributed little, we do not think that their use in well-being research would expand our knowledge of the structure of PWB other than, as previously said, to provide a more stringent approach to estimating the correlation matrices. Instead, we believe that the issues relating to the use of the larger scales and the extent to which sample characteristics have influenced the previous results, are much more important issues to consider.

The use of the larger 84- and 54-item scales in this paper is an improvement on previous validation studies that have used the shorter scale versions, which comprise far fewer similar items, and has resulted in considerable confusion about the efficacy of Ryff's PWB scales. As well, previous PWB validation studies have typically reported internal reliability and CFA techniques, whilst the methodology employed in the one EFA study (Kafka and Kozma 2002) has serious limitations. CFA procedures are conceptually different from EFA techniques, being generally theory rather than data driven, and this study has addressed these concerns by using an EFA technique to analyse the larger scale versions, resulting in a revised 3-factor model of PWB, supporting Abbot et al.'s (2006) CFA findings that four of the PWB variables are highly interrelated. To the authors' knowledge, this is the first paper that has undertaken an EFA approach that has correctly used a PAF procedure with an oblique rotation. Importantly, at the item level, we found this procedure could delineate PWB constructs from two broad valence SWB constructs, positive and negative affect, whilst an oblique rotation reported moderate associations at the factor level.

Some limitations to our studies relate to the relatively small sample sizes for the CFA to generate reliable parameter estimates in Structural Equation Modelling. In addition, these studies were designed independently of each other and therefore a number of socio-demographic variables were classified differently from each other. For instance, age groupings reflected the target population of each study and as such a sub-groups analysis of the PWB scales based on age and other demographic variables was not possible. The preliminary sub-groups analysis of gender did reveal some differences and may explain why there is such debate over the validity of Ryff's model of PWB as differences in the findings of previous validation studies may reflect these sampling characteristics. However, the unequal distribution of gender in both our studies prohibits us from placing too much weight to this claim. The post-hoc analysis by cohort within the teacher study revealed greater consistency in the items and structure identified by EFA and suggests that some demographic effects may be consistent amongst Western schoolteachers though differences are still apparent. It should be noted that the Abbot et al. (2006) analysis comprised a birth cohort sample who were all female, therefore age and gender effects could not have been an issue for their findings which identified the second order EGPS variable as a better fitting model than the 6-factor model. Still it does provide support for the multi-dimensional properties purported to be measured by Ryff's (1989) PWB scales, but also a hierarchical structure which we suggest needs to be investigated further.

6 Conclusion

The purpose of this paper has not been to discredit the value of the Ryff PWB scales, nor the previous validation studies cited in this paper. Rather, the authors have sought a back-to-basics approach to test the dimensionality of the PWB scales. The complexity in drawing out a satisfactory conclusion on the structural validity of PWB, leads us to suggest that further development into the nature and structure of well-being, which recognises the

multiple domains and hierarchical structure inherent to self-referent attitudes, is warranted. Further analyses should identify the extent to which socio-demographic characteristics may influence the structural validity of the PWB scales. Finally, we would propose that Ryff's (1989b) PWB scales are limited by item content that comprise both general and context-specific judgements of well-being. However, we recognise the link between PWB in its current form and a number of health outcomes, and would conclude that Ryff's PWB scales are an appropriate tool for assessing distinct aspects of PWB at a general level, though the extent to which this can be replicated across populations will be influenced by sampling characteristics.

Acknowledgements The authors thank the participants from both studies, Mr. Ross Bool and the Technical Services staff for organizing the online questionnaires, and Dr. Tim Windsor for feedback on an earlier manuscript.

References

Abbot, R. A., Ploubidis, G. B., Huppert, F. A., Kuh, D., Wadsworth, M. E. J., & Croudace, T. J. (2006). Psychometric evaluation and predictive validity of Ryff's psychological wellbeing items in a UK cohort sample of women. *Health and Quality of Life Outcomes, 4*, 76. doi:10.1186/1477-7525-4-76.

Aish, A. M., & Jöreskog, K. G. (1990). A panel model for political efficacy and responsiveness: An application of LISREL 7 with weighted least squares. *Quality & Quantity, 24*, 405–426. doi:10.1007/BF00152013.

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

Burns, R. B. (1979). *The self-concept theory in measurement, development and behavior*. London: Longman.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications and programming*. New Jersey: Lawrence Erlbaum Associates.

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*(Supp 1), 79–90. doi:10.1017/S1041610202008013.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*(2), 276–302. doi:10.1037/0033-2909.125.2.276.

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling, 9*, 327–346. doi:10.1207/S15328007SEM0903_2.

Dollan, C. V. (1994). Factor analysis of variables with 2, 3, 4 and 7 response categories: A comparison of categorical variable estimators using simulated data. *The British Journal of Mathematical and Statistical Psychology, 47*, 309–326.

Fava, G. M., Ruini, C., Rafanelli, C., Finos, L., Conti, S., & Grandi, S. (2004). Six-year outcome of cognitive behavior therapy for prevention of recurrent depression. *The American Journal of Psychiatry, 161*, 1872–1876. doi:10.1176/appi.ajp.161.10.1872.

Headey, B. (2000). Subjective well-being: Revisions to dynamic equilibrium theory using national panel data and panel regression methods. *Social Indicators Research, 79*, 369–403. doi:10.1007/s11205-005-5381-2.

Headey, B. W., & Wearing, A. J. (1989). Personality, life events and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology, 57*, 731–739. doi:10.1037/0022-3514.57.4.731.

Holgado-Tello, F. P., Carrasco-Ortiz, M. A., Victroria del Barrio-Gandara, M., & Chacon-Moscoso, S. (2007). Factor analysis of the big five questionnaire using polychoric correlations in children. *Quality and Quantity,* 202–214. doi: 10.1007/s11135-007-9085-3.

Jöreskog, K. G., & Sörbom, D. (1996). *PRELIS 2: User's reference guide*. Chicago: SSI.

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*, 171–190. doi:10.1023/A:1014451725204.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York: Russell Sage.

Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimising well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology, 82*(6), 1007–1022. doi: 10.1037/0022-3514.82.6.1007.

Marks, N. F., & Lambert, J. D. (1998). Marital status continuity and change among young and midlife adults: Longitudinal effects on psychological well-being. *Journal of Family Issues, 19*, 652–686. doi: 10.1177/019251398019006001.

Marsh, H. W. (1992). *Self Description Questonnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and research monograph*. Macarthur: Faculty of Education, University of Western Sydney.

Maslow, A. (1968). *Toward a psychology of being* (2nd ed.). New York: Van Nostrand.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*, 396–402.

Rogers, C. R. (1961). *On becoming a person*. Oxford: Houghton Mifflin.

Ryff, C. D. (1989a). Beyond Ponce de Leon and life satisfaction: New directions in quest of successful aging. *International Journal of Behavioral Development, 12*, 35–55. doi:10.1177/016502548901200102.

Ryff, C. D. (1989b). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*, 1069–1081. doi:10.1037/0022-3514.57.6.1069.

Ryan, R., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*, 141–166. doi:10.1146/annurev.psych.52.1.141.

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. doi:10.1037/0022-3514.69.4.719.

Ryff, C. D., Keyes, C. L. M., & Hughes, D. L. (2004a). Psychological well-being in MIDUS: Profiles of ethnic/racial diversity and life course uniformity. In O. G. Brim, C. D. Ryff, & R. C. Kessler (Eds.), *How healthy are we? A national study of well-being at mid-life, The John D. and Catherine T. MacArthur foundation series on mental health and development: Studies on successful midlife development* (pp. 398–424). Chicago: University of Chicago Press.

Ryff, C. D., & Singer, B. H. (2006). Best news yet on the six-factor model of well-being. *Social Science Research, 35*, 1103–1119. doi:10.1016/j.ssresearch.2006.01.002.

Ryff, C. D., Singer, B. H., & Love, G. D. (2004b). Positive health: Connecting well-being with biology. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 359*, 1383–1394. doi:10.1098/rstb.2004.1521.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407–441.

Smider, N. A., Essex, M. J., & Ryff, C. D. (1996). Adaptation of community relocation: The interactive influence of psychological resources and contextual factors. *Psychology and Aging, 11*(2), 362–372. doi:10.1037/0882-7974.11.2.362.

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*, 1080–1102. doi:10.1016/j.ssresearch.2005.07.004.

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*, 1120–1131. doi:10.1016/j.ssresearch.2006.01.003.

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*, 629–643. doi: 10.1016/S0191-8869(03)00122-3.

Van Dierendonck, D., Díaz, D., Rodríguez-Carvajal, R., Blanco, A., & Moreno-Jiménez, B. (2007). Ryff's six-factor model of psychological well-being, a Spanish exploration. *Social Indicators Research, .* doi: 10.1007/s11205-007-9174-7.

Watson, D., Clark, A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. doi:10.1037/0022-3514.54.6.1063.

Wylie, R. C. (1974). *The self-concept: A critical survey of pertinent research literature*. Lincoln: University of Nebraska Press.