

The Psychometric Equivalence of the Personal Wellbeing Index for Normally Functioning and Homeostatically Defeated Australian Adults

Ben Richardson · Matthew D. Fuller Tyszkiewicz ·
Adrian J. Tomyn · Robert A. Cummins

Published online: 14 January 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Understanding subjective wellbeing (SWB) at the population level has major implications for governments and policy makers concerned with enhancing the life quality of citizens. The Personal Wellbeing Index (PWI) is a measure of SWB with theoretical and empirical credentials. Homeostasis theory offers an explanation for the nature of SWB data, including the distribution of scores, maintenance and change over time. According to this theory, under normal conditions, the dominant constituent of SWB is Homeostatically Protected Mood (HPMood), which is held within a genetically determined range of values around a set-point. However, in extreme circumstances (e.g., financial hardship, chronic illness), HPMood may dissociate from SWB, as cognitive/emotional reactions to the cause of homeostatic challenge assume control over SWB. This study investigates two groups as people scoring in the positive range for SWB and people who are likely to be experiencing homeostatic defeat/challenge. We test whether the reduced influence of HPMood on SWB due to homeostatic defeat has implications for the validity of SWB measurement. Participants were 45,192 adults (52 % female), with a mean age of 48.88 years (SD = 17.35 years), who participated in the first 23 surveys of the Australian Unity Wellbeing Index over the years 2001–2010. Multiple regression analysis, multiple group confirmatory factor analysis, and Rasch modelling techniques were used to evaluate the psychometric performance of the PWI across the two groups. Results show that while the PWI functioned as intended for the normal group, SWB in the challenged group was lower across all PWI domains, more variable, and the domain scores lacked the strength of inter-correlation observed in the normal, comparison group. These changes are consistent with predictions based on homeostasis theory and one major implication of the findings is that SWB measures may not function equivalently across the entire spectrum of possible domain satisfaction scores.

B. Richardson (✉) · M. D. Fuller Tyszkiewicz · R. A. Cummins
School of Psychology, Deakin University, Burwood, Australia
e-mail: ben.richardson@deakin.edu.au

A. J. Tomyn
Discipline of Psychology, School of Health Sciences, RMIT University, Melbourne, Australia

Keywords Personal Wellbeing Index · Homeostatic defeat · Psychometric equivalence · SWB measures

1 Introduction

The use of subjective wellbeing (SWB) as a social indicator, to complement objective measures of national progress and performance, has risen dramatically over the last decade. This, in part, can be attributed to SWB becoming more widely accepted as a useful and important measure of how people feel about themselves and their lives. While there are many different definitions of SWB, it is generally agreed that this broad, complex construct comprises both affective and cognitive components (e.g., Campbell et al. 1976; Diener and Diener 1996; Steel and Ones 2002; Veenhoven 1994). However, recent evidence suggests that the dominant component of SWB may be mood affect, in a form called Homeostatically Protected Mood (HPMood; Cummins 2010), with cognitive discrepancies playing a significant but subsidiary role (Davern et al. 2007). HPMood is based on Russell's (2003) description of Core Affect, which he defined as a "neurophysiological state consciously accessible as the simplest raw (non-reflective) feelings evident in moods and emotion" (p. 148). In a given moment, a state of core affect is experienced as some combination of Hedonic Valence and Arousal.

The empirical findings of Davern, Cummins and Stokes (2007), and the more recent findings of Blore et al. (2011) and Tomin and Cummins (2011) corroborate this construction. Each study independently explored the relative strength of HPMood (represented by the three affects of happy, content and excited/alert) in three separate structural equation models incorporating cognition (measured according to variables comprising Multiple Discrepancies Theory; MDT, Michalos 1985) and personality (NEO-FFI Personality Inventory; Costa and McCrae 1992). All three studies conclude that HPMood dominates the prediction of SWB over and above the measures of personality and cognition. Collectively, these studies challenge the received wisdom regarding the relationship between SWB and related variables, arguing that HPMood is the driving force behind SWB and not personality or cognition as is generally reported in the literature (e.g., DeNeve and Cooper 1998; Diener et al. 1985; Headey and Wearing 1989, 1992; Vitterso 2001; Vitterso and Nilssen 2002). This contention supports a further idea that, following perturbation, SWB normally returns not only to an 'equilibrium level', as first proposed by Headey and Wearing (Headey et al. 1984a, b; Headey and Wearing 1989), but to a set-point (Cummins et al. 2012, 2014). These findings are consistent with the notion that SWB may be under genetic control, and is thus considered to be an individual difference (see e.g., Lykken and Tellegen 1996).

These propositions form the theoretical underpinnings of SWB Homeostasis Theory (Cummins 1995, 1998, 2010). This theory asserts that SWB is actively maintained around a biologically determined set-point, and that within normal population samples, set-points have a normal distribution of 70–90 points, with an average of 80 points (Cummins et al. 2012, 2014), measured on a standardised 0–100 point scale. This theory further posits that HPMood approximates the set-point and is the construct that homeostasis seeks to defend. In this way, each person can normally maintain a positive level of personal wellbeing that extends some 9–10 % points around their set-point. Thus, homeostasis theory offers a theoretical explanation for both the stability and change in SWB that has been observed at

both the level of individuals (e.g., Headey and Wearing 1989, 1992) and populations (e.g., Cummins et al. 2012).

Homeostasis Theory makes a number of predictions concerning the behavior of SWB. Most notably, that under normal, relatively unthreatening life circumstances, a person's SWB will approximate their biologically determined level of HPMood, which lies in the range of 70–90 points on a scale of life satisfaction. A second prediction is that, if SWB approaches the upper or lower limits of its set-point range, the homeostatic system will resist further change and act to restore SWB back to its set-point through the processes of adaptation (Helson 1964) and other cognitive devices described by Cummins and Wooden (2014). However, strongly aversive conditions, such as extreme financial hardship or generalized anxiety, may defeat the capacity of the homeostatic system to regulate SWB (Cummins 2010; Cummins and Nistico 2002). Once this occurs, and the homeostatic threshold has been exceeded, the theory predicts a considerable drop in SWB below the normal set-point range. Under these circumstances, SWB will lose its association with HPMood, as emotions/cognitions related to the specific cause of homeostatic defeat assume control. Thus, a major implication of this reasoning is that there may be a theoretical difference in the functioning of people who are in normal homeostatic maintenance and those who are defeated. Specifically, the lack of correspondence between the domains of satisfaction and HPMood in challenged/defeated groups is likely to have implications for measurement, such that responses to the domains no longer reflect the underlying SWB construct of interest. These measurement issues threaten the validity of attempts to combine and/or compare data of normally functioning and homeostatically challenged individuals. Consistent with this, a recent Rasch analysis of the PWI measure of SWB (Forjaz et al. 2012) found that respondents who tended to endorse items using the lower end of the scale (i.e., they were reporting lower levels of life satisfaction), displayed higher levels of disordered responding. In particular, they were unable to meaningfully differentiate between response options located at the lower end of the scale suggesting a lack of coherence between their level of SWB and their self-reported satisfaction with specific life domains.

One frequently used measure of SWB, with empirical and theoretical credentials, is the Personal Wellbeing Index (PWI; International Wellbeing Group (IWG) 2006). The PWI is a multiple-item measure of SWB that asks respondents to indicate their level of satisfaction with seven life domains: Standard of Living, Health, Achieving in Life, Relationships, Safety, Community Connectedness, and Future Security. These domains have been selected on the basis that they each contribute significant unique variance to the prediction of the single item measure of General Life Satisfaction (GLS; Andrews and Withey 1976) which asks 'How satisfied are you with your life as a whole?' (IWG 2006). The user manual reports good internal reliability for the PWI across a diverse range of samples and cultural groups (Cronbach's alpha between .70 and .85). Moreover, the combination of unique and shared variances contributed by the domains typically explain between 30 and 60 % of the variance in GLS (45 % in this study), while factor analysis reveals that the seven domains consistently form a single stable factor that accounts for approximately 50 % of variance in Australia and other countries (IWG 2006).

1.1 Summary and Study Aims

This study investigates the validity of the PWI as measured within two distinct groups: (1) people who score in the positive range for SWB (>50); and (2) people who score in the homeostatically challenged/defeated range ($SWB \leq 50$). These ranges have been established through the calculation of set-point ranges (Cummins et al. 2012, 2014). The

theoretical proposition under investigation is that satisfaction judgements used to measure each of the seven domains are normally dominated by HPMood. Moreover, as HPMood is also considered an individual difference, resulting in a normal distribution of set-points within population samples, it is believed to provide the shared variance that causes the domains to correlate with one another. However, this does not apply to people in homeostatic defeat. For people who have a level of SWB ≤ 50 , it is predicted that the domains will correlate less strongly with each other, and with GLS, because reported levels of SWB reflect the dominant emotion causing defeat, rather than HPMood. One implication is that including such individuals in calculations of normative ranges distorts statistical inferences.

To test this methodological issue, the performance of the PWI is examined in both normal and defeated groups using multiple regression analysis, multi-group confirmatory factor analysis and Rasch modelling. It is predicted that the associative processes revealed by the applied analyses will differ between the two groups. Specifically, it is hypothesised that:

1. The correlations between PWI domains and GLS will be higher in the normal group than the challenged/defeated group
2. When regressed against GLS, the domains will contribute greater shared variance in the positive group than the challenged/defeated group
3. The uni-dimensional factorial structure of the PWI will exhibit acceptable fit for the normal group, but poor fit for the challenged/defeated group.
4. Item level responses will exhibit ordered thresholds in the normal range for domain scores (i.e., 6–10), but disordered thresholds for response categories below or equal to 5.

2 Methodology

2.1 Participants

Participants were 45,192 adults (51.9 % female) who responded to the Australian Unity Wellbeing Index (AUWI) over the years 2001–2010 (23 surveys). They are a nationally geographically representative sample of the Australian adult population (with the caveat that they are sufficiently fluent in English to complete the questionnaire). Participants were sampled via telephone and, at the time of responding, had a mean age of 48.88 years ($SD = 17.35$ years).

All respondents completed the Personal Wellbeing Index-Adult (PWI-A; IWG 2006) as the measure of SWB. The PWI generates a composite variable, calculated by averaging satisfaction scores across seven life domains: Standard of Living, Health, Achieving in Life, Relationships, Safety, Community Connection and Future Security. Respondents also completed the single item measure of General Life Satisfaction (GLS) that asks 'How satisfied are you with your life as a whole?' All responses were made using an 11-point, end-defined scale (0 = Completely Dissatisfied; 10 = Completely Satisfied) that were standardised onto a 0–100 point scale. In the current study the correlation between the PWI and GLS single item was .67.

1,658 participants (3.67 %) scored ≤ 50 on the PWI and were classified as the homeostatically defeated group (Cummins et al. 2012, 2014). The remaining 43,534 partic-

ipants (96.33 %) who scored above 50 were classified as the ‘positive’ SWB group.¹ They comprise a mixture of people likely to be functioning normally within their set-point range (scores from 70 to 100 points), and people scoring between 70 and 50. These people may either have a low set-point and be functioning normally, but low in their range, or they may be people with a high set-point who are homeostatically defeated.

2.2 Data Analytic Strategy

SPSS version 21.0 (IBM Corp, 2013) was used for data preparation and screening. To standardise data, all reported values were converted to a Percentage of Scale Maximum (%SM) which converts data onto a 0–100 scale (IWG 2006).

Regression analyses using Mplus version 6.1 (Muthén and Muthén 1998–2011) were then used to evaluate the relationships between General Life Satisfaction (GLS) and domain scores. Separate analyses were run for the homeostatically challenged and positive groups. Each analysis was followed by a separate confirmatory factor analysis (CFA) to evaluate the extent to which a uni-dimensional model of the PWI adequately represented the covariance matrices for members of each group. Adequacy of model fit was examined using the following criteria: Comparative Fit Index (CFI; $>.95$ for good fit, $>.90$ for adequate fit), Root Mean Square Error of Approximation (RMSEA; $\leq.05$ for good fit, $<.08$ for adequate fit), Standardized Root Mean Square Residual (SRMR; $<.05$ for good fit), and factor determinacy scores $\leq.80$ (Byrne 2010; Hu and Bentler 1999; Raykov and Marcoulides 2000). In addition to these global fit indices, local fit was evaluated by inspection of factor loadings for each of the domains of satisfaction. Within the context of factor analysis, Hair et al. (2010) argue that a factor may be considered to be well defined when most of the items have factor loadings in excess of .50 (and ideally .70, with .40 being a minimally acceptable standard).

Although tests of measurement invariance are a logical next step when CFAs are conducted across multiple groups, this step was bypassed in the present study due to poor model fit for the homeostatically challenged individuals (see “Results” section, below). Sources of misfit were instead evaluated using Rasch modeling within the RUMM program (Andrich et al. 2010). Rasch modeling tests the assumption that item scores derive from an individual’s level of the underlying trait (Rasch 1960), such that individuals with high levels of SWB are expected to provide high scores on the domains of satisfaction, whereas individuals with low levels of SWB are expected to provide low domain satisfaction scores.

However, when a scale does not conform to this assumption, Rasch models provide a range of useful diagnostic tools to help determine sources of misfit within the measurement model. In the present study, we were interested in item targeting as well as potential disordered threshold effects. Item targeting reflects the level of match (or mismatch) between the participants’ underlying trait (i.e., subjective wellbeing) and item difficulty. In the context of SWB, a ‘difficult’ item is one that poorly discriminates between participants

¹ A reviewer pointed out that the large sample size discrepancy between the positive ($n = 43,534$) and defeated group ($n = 1,658$) might undermine between group comparisons. To test whether variance restriction was an issue in the defeated group we took random sub-samples of various sizes and plotted the variation to see at which point it stabilized within that group. Even when the random subsample size was 10 % (i.e., $n = 166$), the range of variance was stable (standard deviation only varied by around 1) across random subsamples suggesting that increasing that sample size would have little effect on range of scores and variance in that group. Furthermore, as the sample size increased, the variation in standard deviations was small and seemingly random.

with low and high levels of SWB because both groups give a low score (low level of endorsement) for the item. Similarly, items may poorly discriminate between individuals of varying levels of SWB because most participants respond in an extreme manner to the item, regardless of underlying level of SWB. This is akin to a ceiling effect. Ideally, we want items that allow for scoring across the range of possible response categories (in this case, 0–10), and which allow for response variation across all the levels of the underlying trait (SWB). Mismatch is problematic because it can result in items that are too difficult or too easy for the majority of participants to endorse, either of which scenario prevents researchers from accurately discriminating between individuals with high and low levels of SWB.

Another source of misfit is disordered threshold effects for the categories of the response scale. The PWI provides an 11-point scale with the assumption that, as scores increasingly approximate the maximum (10), they indicate increasing levels of satisfaction. However, when participants struggle to differentiate between response options (e.g., difficulty differentiating between a 1 and a 3 on the rating scale), then the anticipated order of satisfaction across response categories will not be accurately reflected in the data. While the Rasch model predicts that individuals with very low levels of satisfaction are more likely to score 1 than 3, in instances of disordered thresholds it may be found that people are more likely to endorse 3 than 1. Such findings are counter-intuitive and suggest that the response scale may not be functioning as intended.

3 Results

3.1 Data Cleaning and Preparation

SPSS output revealed that the frequency of missing data for all variables was $<2\%$ and that these values appeared to be missing at random. Missing values were replaced using expectation maximization in SPSS. Examination of z-scores revealed univariate outliers on domain satisfaction variables. Comparison of mean scores on these variables with corresponding means trimmed at the upper and lower 5% showed that none of these outliers significantly influenced mean scores on key variables. As a consequence, these univariate outliers were retained in subsequent analyses. Absolute skew and kurtosis values were <2

Table 1 Means and standard deviations for domain scores by group

Domains	Positive SWB (PWI > 50)		Defeated (PWI ≤ 50)		Fmax ^a
	M	SD	M	SD	
1. Standard	79.00	15.47	43.54	21.61	550.76
2. Health	76.23	18.28	41.61	24.03	417.40
3. Achieving	75.12	16.53	36.19	22.99	549.82
4. Relationships	80.98	19.17	38.28	28.97	1,140.26
5. Safety	79.96	16.37	48.47	25.71	931.70
6. Community	72.02	18.50	37.62	22.97	247.70
7. Future	72.40	18.06	32.56	21.57	204.87
PWI	76.72	10.10	41.34	8.90	93.31

^a All Hartley's Fmax results are significant, $p < .001$

and 7, respectively, suggesting that departures from normality for domain scores are unlikely to substantially distort results (Curran et al. 1996). To further ensure accuracy of obtained results, CFA models were run with both the maximum likelihood (ML) and the maximum likelihood robust (MLR) estimators in Mplus, and results did not substantially differ. Consequently, ML results are reported in the present paper.

3.2 Descriptives/Correlations

On average, individuals in the homeostatically defeated group had considerably lower average satisfaction scores for each of the domains as expected (see Table 1). They also exhibited significantly greater within-group variability for each of the domains, as tested using Hartley’s Fmax test (Hartley 1950).

Also of interest is that the removal of defeated scores from the sample has raised the PWI mean by only 1.47 points from the combined data (Cummins et al 2012, 2014: Table A2.21: Mean = 75.27 points) and that all domain mean scores remain within their normal range. This small degree of change is due to the fact that the defeated sample comprises only 3.6 % of the total sample. Also notable, however, are the smaller SDs compared to the defeated group. This is consistent with the dominating influence of set-points, and HPMood as an individual difference, operating most strongly within the positive group.

“Appendix 1” shows the proportion of responses per category for each of the domains, presented separately for the two target populations. Not surprisingly, individuals in the homeostatically defeated group tended to use response options below five, whereas individuals in the healthy group tended to use response options 6–10. However, the distributions are far from clear-cut, with considerable cross-over between categories within each domain.

3.3 Regression Analyses

Separate standard multiple regression analyses were conducted for each of the two groups, as well as for the entire sample, to evaluate the extent to which the domains of the PWI predict GLS. As expected, the domains accounted for approximately half of the variance in

Table 2 Cross-sample differences in regressing life satisfaction onto the domains of satisfaction

Domain	Entire sample			Positive SWB (PWI > 50)			Defeated (PWI ≤ 50)		
	B	Se	Sr ²	B	Se	Sr ²	B	Se	Sr ²
1. Standard	.30*	.004	.063	.30*	.004	.073	.31*	.023	.068
2. Health	.08*	.003	.006	.07*	.003	.006	.11*	.021	.017
3. Achieving	.24*	.004	.044	.23*	.004	.048	.25*	.022	.068
4. Relationships	.16*	.003	.032	.17*	.003	.036	.13*	.018	.032
5. Safety	.01	.004	.000	.00	.004	.000	.02	.020	.000
6. Community	.06*	.003	.004	.06*	.003	.004	.12*	.022	.017
7. Future	.07*	.004	.004	.07*	.004	.005	.11*	.024	.008
R ²	.500			.430			.230		
Shared variance	.347			.258			.020		
Unique variance	.153			.172			.210		

* *p* < .05

life satisfaction scores for the entire sample (see Table 2). All domains, except Safety, made a significant unique contribution to the prediction of GLS. However, according to the user manual for the PWI (IWG 2006), Safety rarely makes a unique contribution to GLS using Australian samples, but is retained in the scale because it does so in other countries. While the unstandardized coefficients and R^2 values for the positive SWB group were largely consistent with the sample as a whole, it is evident that the variance accounted for is substantially lower for the homeostatically defeated subsample than the positive sample ($R^2 = .23$ vs $R^2 = .43$).

Most notable from Table 2 is the changed proportion of shared and unique variance. The shared variance in the defeated group has become negligible, while unique variance has increased. These changes are consistent with increased domain independence.

3.4 Factor Analyses

While a uni-dimensional model provides an adequate fit of the inter-correlations between domain scores for the sample as a whole ($\chi^2 = 38,588.041$; CFI = .933, RMSEA = .063 [90 % CIs: .061–.065], SRMR = .051, factor determinacy = .88, Cronbach's alpha = .77) and also for the positive subsample ($\chi^2 = 25,257.65$; CFI = .901, RMSEA = .063 [90 % CIs: .061–.065], SRMR = .044, factor determinacy = .85, Cronbach's alpha = .72), it is a poor representation of the data for the homeostatically defeated subsample ($\chi^2 = 260.173$; CFI = .249, RMSEA = .085 [90 % CIs: .075–.096], SRMR = .061, factor determinacy = .64, Cronbach's alpha = .38).

Inspection of factor loadings (Table 3) reveals that each of the domain scores is moderately to strongly related to the underlying factor for both the sample as a whole and for the positive subgroup. In contrast, the factor loadings tended to be weaker for the homeostatically defeated group. On the basis of these data alone (i.e., failure to find a common factor structure), we can conclude that the PWI is non-invariant across the two subsamples. However, this does not inform us about the potential reasons for this discrepancy. Accordingly, follow-up Rasch analyses were conducted to evaluate potential problems in the scoring structure of the PWI.

3.5 Rasch Modeling—Item Targeting

Figure 1 shows the level of (mis)match between person and item characteristics for the domains for the sample as a whole. Items are deemed suitable for measuring the full range

Table 3 Factor loadings for one-factor model of the PWI

Domain	Entire sample	Positive SWB (PWI > 50)	Defeated (PWI ≤ 50)
1. Standard	.608	.633	.453
2. Health	.464	.377	.106
3. Achieving	.637	.570	.024
4. Relationships	.497	.404	−.328
5. Safety	.500	.431	.244
6. Community	.526	.457	.067
7. Future	.676	.630	.427
Reliability estimate	.769	.716	.382

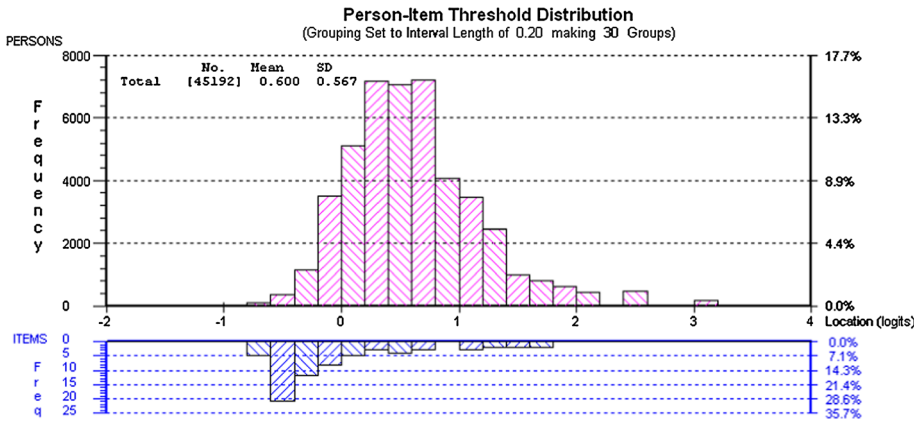


Fig. 1 Location and fit of the item parameters

of individual differences on a given latent variable (in this case, SWB) when the distributions for person (the upward pointing bars in Fig. 1) and item (downward pointing bars) have similar minimum and maximum values. The shape of the distribution for these bars is also instructive and indicates potential skew; a distribution for items that tends towards the left (minimum score) would suggest that the items are easy to endorse (and may signal potential floor effects) whereas a distribution for items (downward bars) that tends towards the right (maximum score) would suggest that the items are too difficult (a potential ceiling effect). Similarly, a distribution of participants (upward bars) that tends towards the left suggests that the majority of the sample have low levels of SWB, whereas skew towards the right (as evident in Fig. 1) reflects a sample that tends towards high levels of SWB. In instances where the person and item distributions overlap in terms of range of scores (transformed into logit scores), this provides assurance that the items cover a range of difficulty to adequately measure wellbeing across the spectrum of SWB exhibited in the sample.

In Fig. 1, there is a positive skew in the person locations (presented in the top portion of Fig. 1), which indicates that the sample as a whole has positive wellbeing. The distribution of item locations (roughly -1 to $+2$ logits) overlaps with the distribution of person

Table 4 Estimates of individual item parameters across the domains of satisfaction

Domain	Location	Thresholds									
		1	2	3	4	5	6	7	8	9	10
Standard	-.12	-.40	-.44	-.49	-.52	-.49	-.38	-.13	.27	.87	1.70
Health	-.01	-.59	-.48	-.42	-.37	-.31	-.21	-.02	.29	.74	1.37
Achieving	.05	-.22	-.46	-.62	-.66	-.60	-.41	-.09	.36	.96	1.72
Relationships	.01	.32	-.02	-.24	-.36	-.39	-.32	-.17	.07	.37	.74
Safety	-.13	-.38	-.38	-.40	-.41	-.37	-.28	-.08	.24	.70	1.35
Community	.11	-.27	-.54	-.66	-.64	-.50	-.26	.07	.47	.92	1.41
Future	.10	-.40	-.58	-.66	-.63	-.51	-.28	.04	.46	.98	1.59

Disordered thresholds appear in bold typeface

locations (range of approximately -1 to $+3$ logits) with the exception of extreme positive scores (right hand-side of the distribution of upward pointing bars) on the person locations. This suggests that the domain items are well-matched for the level of subjective wellbeing except for individuals with very high levels of SWB.

The item location and estimates of threshold parameters are shown in Table 4. The item location value indicates the level of SWB necessary for an individual to have a 50:50 chance of endorsing a particular item; that is, the person is just as likely to endorse as not endorse the item. As these values are mapped on to the underlying trait of SWB, lower (more negative) values indicate items that are easy to endorse (i.e., the individual chooses a relatively high response value for the item even when they have a relatively low level of SWB), whereas higher (positive) values indicate higher level of difficulty for individuals to endorse the item (i.e., the individual chooses a relatively high response value only when they have a relatively high level of SWB). In contrast, the threshold parameters represent the level of the underlying trait (i.e., SWB) at which two adjacent, possible scores on a rating scale are equally likely to be selected by participants. As the PWI items have an 11-point rating scale, there are ten threshold points for each item (0 vs 1, 1 vs 2, 3 vs 4, ... 9 vs 10). The threshold parameter then reflects the score on SWB at which participants are equally likely to choose either of the adjacent rating values.

Threshold values are expressed in logits rather than the traditional 0–10 scale. For instance, the domain ‘standard of living’ has a threshold parameter value of $-.40$ for the first threshold (a satisfaction score of 0 vs a score of 1). This value of $-.40$ implies that individuals with a SWB trait score $<-.40$ (e.g. $-.41$) are more likely to choose response option 0 than 1 whereas an individual with a trait score that is more than $-.40$ (e.g. $-.39$) is more likely to choose response option 1.

When participants use response categories as expected by the scale developer, threshold parameters should become more positive in a non-linear, yet monotonic fashion from lowest threshold (in this instance, threshold for choosing value 1 over value 0) to highest threshold (choosing 10 vs 9). In other words, the threshold parameters for a given domain item should increase in size from left to right in Table 4. Disordered thresholds—that is, instances where the response options are used in a different order to that proposed by the scale developers—are evident when one threshold value on a given item is smaller than any of the previous threshold values to its left.

In our data, there is clear evidence of disordered thresholds for all items. Importantly, this disordering effect is present at the lower levels of the response scales, where homeostasis theory predicts that individuals with homeostatic failure are more likely to reside. For example, the threshold values for Standard of Living do not consistently increase for each response option; an individual with a logit-transformed SWB score of $-.51$ is more likely to give a score of 4 than a score of 1, 2, or 3 whereas individuals with higher SWB scores were more likely to endorse 1, 2 or 3.

4 Discussion

This study investigated the functioning of the Personal Wellbeing Index (PWI) across two groups: people who score in the positive range for Subjective Wellbeing (SWB >50) and people who are likely to be experiencing homeostatic defeat/challenge (SWB ≤ 50). Homeostasis Theory (Cummins 2010) predicts that under normal circumstances, Homeostatically Protected Mood (HPMood) is both the main driver of domain satisfaction judgments and the reason that the domains correlate with one another. In contrast, when

homeostasis is defeated, HPMood loses its association with satisfaction judgments as these fall under the control of cognitions and emotions associated with the cause(s) of challenge. When this happens, satisfaction across the various domains is not dominated by HPMood and, so, becomes less stable and predictable. An implication of this transition is that measurement of SWB becomes less reliable. The pattern of findings reported in this paper is consistent with this notion, and the lines of evidence will now be discussed.

First, the simple correlations and descriptive statistics show that: (1) there is greater variability in scores for the defeated group and yet (2) the correlations are stronger between domain scores within the positive group. Thus, although greater variance in the domain scores would usually be expected to lead to increased effect sizes (particularly when comparing against a subsample with restricted range), this added variation does not seem to vary systematically with other domains for individuals in the homeostatically defeated group. This may indicate randomness of responses or failure to differentiate between various options at the low end of the response scale. Consistent with these results, regression analyses showed that the domain scores accounted for less variance in general life satisfaction for the defeated group compared to the positive group (24 vs. 48 %). This further supports the notion that the variation in the defeated group responses is less systematic and more haphazard compared to the normally positive group.

Second, results from the confirmatory factor analyses (CFA) show that factor loadings are lower for the defeated group, and that the one factor solution is a poor fit in this population. In contrast, the model fits comparatively well for the entire sample as it does for the normally positive group, where the loadings are much higher. Further probing of psychometric properties of the PWI using item response theory shows generally disordered threshold effects across items for the lower scores (i.e., 0–5 on the 0–10 response scale). Given that response options at or below five are typically from individuals in the defeated group, this suggests that these people have difficulty differentiating between these values which comprise the lower half of the response scale. This potential randomness in their choice of a number to represent their level of (dis)satisfaction may also account for the defeated group's poorer regression and CFA results, as well as their greater variability in item scores relative to the positive SWB group.

The implication of these findings is that the PWI seems to function as intended for individuals who respond in the positive half of the response scale (which is around 96 % of individuals). However, the poorer performance of the scale in the defeated group suggests that their data are not reliable and, hence their inclusion in an overall sample will degrade the reliability of the result. The extent to which this phenomenon is likely to seriously impact on the overall reliability of results from a study will depend on the proportion of respondents in the defeated group. In the current general population sample these people comprised 3.6 % of the total, and thus likely had negligible impact on overall results, as shown by the similarity in regression and CFA results for the normally functioning group and the overall sample. However, this proportion will increase markedly in samples where the incidence of depression is high.

What seems clear from these results is that general population samples are not 'normal'. They contain a proportion of respondents who evidence homeostatic defeat and who contribute data that are unreliable. Since the proportion of such people will vary with the sample employed, it is evident that these differences may well represent a systematic source of error in comparisons between study outcomes. One way of increasing comparability is to exclude individuals with $SWB \leq 50$. Another solution, consistent with the recommendation of Forjaz et al. (2012), is to collapse the response

options at the lower end of the scale such that a response that is between 1 and 4 is treated equivalently. This option may be preferable to simply reducing the number of response options on the scale. In this case, it is still possible that homeostatic defeat, and the resulting lack of coherence between domains would cause disordered thresholds at the lower end of the response scale. However, testing the performance of the scale with a smaller number of response options would help to disentangle methodological interpretations of the current results (e.g., too many response options) from more conceptual interpretations (e.g., reduced coherence between life domains due to homeostatic defeat). A longer-term solution would be to develop a scale that functions equivalently across groups. As partially detailed above, there are several reasons to suspect that a simple revision of the PWI (or other existing measures of SWB) may be inadequate. First, our pattern of results suggests that the experience of SWB may be qualitatively, rather than quantitatively, different across the two groups, suggesting creation of a common measure may not be possible. That is, it was predicted and found that the domains assessed by the PWI are coherent only when SWB is within a normal range. However, they should (and do) become less related at lower levels of SWB. Given this, one approach might be to develop a longer scale that includes multiple items to measure each domain. Homeostatic theory does not predict that multiple items measuring the domains should become less coherent at lower levels of SWB. Thus, these domains could be reliably and meaningfully measured within both 'normal' and 'defeated' groups and used in analyses in cases where the total score does not function adequately.

4.1 Study Limitations and Future Directions

An issue not addressed in this study is the extent to which the results observed using the PWI will generalize to other measures. Based on the reasoning of Homeostasis Theory, the reduced coherence of items observed at the lower end of the PWI scale should also be observed in data collected using alternative measures that comprise self-report data. Given the implications of this reasoning for both the measurement and conceptualization of the construct, evaluating this proposition is a key direction for future research.

4.2 Summary and Conclusions

A growing understanding of the SWB construct, and its recognition as a useful and important measure of national performance and progress, has resulted in an increase in the development and implementation of scales that purport to measure this variable in some way. The PWI is one such measure and has been used to measure SWB in the Australian population as part of the Australian Unity Wellbeing Index since 2001. This research has assisted in the development and revision of SWB Homeostasis Theory, which asserts that SWB is normally positive, stable and actively maintained and defended around a narrow, set-point range of values. This theory also asserts that the relative affective/cognitive composition of SWB judgments will vary depending on whether a person is experiencing a level of wellbeing that approximates their set-point, or lower than normal due to external threat/challenge. The results presented in this study highlight the impact that population challenge/defeat has on the reliability of measurement. Most crucially, individuals in the homeostatically defeated range seem to use the rating scale inconsistently. This tendency has been shown to undermine the psychometric properties of the scale and to render meaningful comparisons of SWB between healthy and challenged groups invalid. A replication of these results and further research into the causes of this randomness in response

patterns is clearly required. In the meantime, researchers are encouraged to evaluate the impact that responses from homeostatically defeated individuals has on the results for their whole samples.

Appendix 1

See Table 5.

Table 5 Proportion of responses in different categories across domains, separated by grouping (positive versus homeostatically defeated)

Domain	Response category										
	0	1	2	3	4	5	6	7	8	9	10
<i>1 Standard of living</i>											
Positive	.1	.1	.2	.6	1.3	5.7	6.3	17.3	33.7	18.8	15.8
Defeated	6.2	2.8	8.8	11.8	15.5	27.5	10.5	8.9	5.4	1.6	1.1
<i>2 Health</i>											
Healthy	.3	.3	.8	1.3	2.8	8.2	7.9	17.0	27.5	19.0	14.8
Defeated	7.5	4.3	10.9	13.7	14.4	21.2	7.8	8.3	7.2	2.0	2.8
<i>3 Achievement</i>											
Healthy	.2	.1	.5	.8	2.1	8.5	9.5	22.1	29.5	15.4	11.2
Defeated	11.9	5.3	12.2	13.4	16.3	22.6	7.0	5.8	3.1	1.0	1.4
<i>4 Personal relationships</i>											
Healthy	.9	.3	.7	1.0	1.6	5.9	5.0	12.0	23.5	22.4	26.8
Defeated	17.7	6.6	8.9	10.6	9.7	19.5	6.4	6.3	6.3	3.1	4.9
<i>5 Safety</i>											
Healthy	.2	.1	.3	.6	1.6	5.5	6.6	15.6	27.6	22.9	18.9
Defeated	5.9	4.0	7.7	11.0	11.0	23.4	9.0	10.4	9.2	3.8	4.7
<i>6 Community connectedness</i>											
Healthy	.4	.3	.9	1.6	3.0	12.5	11.9	20.6	25.2	12.7	10.9
Defeated	9.6	6.5	12.9	13.1	13.1	25.1	7.0	6.1	3.8	1.3	1.6
<i>7 Future security</i>											
Healthy	.3	.3	.8	1.5	3.3	11.2	11.7	21.5	26.0	13.0	10.5
Defeated	12.6	8.8	13.6	16.6	14.2	21.8	5.2	4.3	1.7	.4	.9

References

- Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: Americans' perceptions of life quality*. New York: Plenum Press.
- Andrich, D., Sheridan, B., & Luo, G. (2010). RUMM 2030. 4.0 for windows (upgrade 4600.0109) edn. Perth, WA: RUMM Laboratory Pty Ltd.
- Blore, J. D., Stokes, M. A., Mellor, D., Firth, L., & Cummins, R. A. (2011). Comparing multiple discrepancies theory to affective models of subjective wellbeing. *Social Indicators Research, 100*, 1–16. doi:10.1007/s11205-010-9599-2.

- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Taylor and Francis Group, LLC.
- Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations and satisfactions*. New York: Russell Sage Foundation.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, Florida: Psychological Assessment Resources.
- IBM Corp (2013). IBM SPSS Statistics for Windows Version 22.0. Armonk, NY: IBM Corp.
- Cummins, R. A. (1995). On the trail of the gold standard for subjective wellbeing. *Social Indicators Research*, 35, 179–200. doi:10.1007/BF01079026.
- Cummins, R. A. (1998). The second approximation to an international standard for life satisfaction. *Social Indicators Research*, 43(3), 307–334. doi:10.1023/a:1006831107052.
- Cummins, R. A. (2000). Objective and subjective quality of life: An interactive model. *Social Indicators Research*, 52, 55–72. doi:10.2307/27522495.
- Cummins, R. A. (2010). Subjective wellbeing, homeostatically protected mood and depression: A synthesis. *Journal of Happiness Studies*, 11, 1–17. doi:10.1007/s10902-009-9167-0.
- Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive bias. *Journal of Happiness Studies*, 3, 37–69. doi:10.1023/a:1015678915305.
- Cummins, R., & Wooden, M. (2014). Personal resilience in times of crisis: The implications of SWB homeostasis and set-points. *Journal of Happiness Studies*, 15(1), 223–235. doi:10.1007/s10902-013-9481-4
- Cummins, R. A., Woerner, J., Weinberg, M., Collard, J., Hartley-Clark, L., Perera, C., & Horfiniak, K. (2012). Australian unity wellbeing index-report 28-the wellbeing of Australians-the impact of marriage. Retrieved 26 April 2014. <http://www.deakin.edu.au/research/acqol/reports/survey-reports/index.php>.
- Cummins, R. A., Li, L., Wooden, M., & Stokes, M. (2014). A demonstration of set-points for subjective wellbeing. *Journal of Happiness Studies*, 15, 183–206. doi:10.1007/s10902-013-9444-9.
- Curran, P. J., West, G. W., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16–29.
- Davern, M., Cummins, R. A., & Stokes, M. (2007). Subjective wellbeing as an affective-cognitive construct. *Journal of Happiness Studies*, 8, 429–449. doi:10.1007/s10902-007-9066-1.
- DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective wellbeing. *Psychological Bulletin*, 124, 197–229.
- Diener, E., & Diener, M. (1996). Most people are happy. *Psychological Science*, 7, 181–185.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with life scale. *Journal of Personality Assessment*, 49, 71–75.
- Forjaz, M. J., Ayala, A., Rodriguez-Blazquez, C., Prieto-Flores, M.-E., Fernandez-Mayoralas, G., Rojo-Perez, F., et al. (2012). Rasch analysis of the international wellbeing index in older adults. *International Psychogeriatrics*, 24, 324–332.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. New Jersey: Pearson Prentice Hall.
- Hartley, H. O. (1950). The use of range in analysis of variance. *Biometrika*, 37, 271–280.
- Headey, B., Holmstrom, E., & Wearing, A. (1984a). The impact of life events and changes in domain satisfactions on well-being. *Social Indicators Research*, 15, 203–227. doi:10.1007/bf00668671.
- Headey, B., Holmstrom, E., & Wearing, A. J. (1984b). Well-being and ill-being: Different dimensions? *Social Indicators Research*, 14, 115–139. doi:10.1007/bf00293406.
- Headey, B., & Wearing, A. (1989). Personality, life events, and subjective wellbeing: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, 57, 731–739. doi:10.1037/0022-3514.57.4.731.
- Headey, B., & Wearing, A. (1992). *Understanding happiness: A theory of subjective wellbeing*. Melbourne: Longman Cheshire.
- Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.
- Hu, L. A., & 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. doi:10.1080/10705519909540118.
- International Wellbeing Group. (2006). Personal Wellbeing Index-Adult (PWI-A).
- Lykken, D., & Tellegen, A. (1996). Happiness is a stochastic phenomenon. *Psychological Science*, 7(3), 186–189.
- Michalos, A. C. (1985). Multiple discrepancies theory (MDT). *Social Indicators Research*, 16, 347–413.

- Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. *Copenhagen: Danish Institute for Educational Research*.
- Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Hillsdale, NJ: Lawrence A Erlbaum Associates.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 1*, 145.
- Steel, P., & Ones, D. S. (2002). Personality and happiness: A national level analysis. *Journal of Personality and Social Psychology, 83*, 767–781.
- Tomyn, A. J., & Cummins, R. A. (2011). Subjective wellbeing and homeostatically protected mood: Theory validation with adolescents. *Journal of Happiness Studies, 12*, 897–914. doi:[10.1007/s10902-010-9235-5](https://doi.org/10.1007/s10902-010-9235-5).
- Veenhoven, R. (1994). Is happiness a trait'? Test of the theory that a better society does not make people any happier. *Social Indicators Research, 32*, 101–160.
- Vitterso, J. (2001). Personality traits and subjective wellbeing: Emotional stability, not extraversion, is probably the important predictor. *Personality and Individual Differences, 31*, 903–914.
- Vitterso, J., & Nilsen, F. (2002). The conceptual and relational structure of subjective wellbeing, neuroticism, and extraversion: Once again, neuroticism is the important predictor of happiness. *Social Indicators Research, 57*, 89–118.