



The PERMA + 4 Short Scale: A Cross-Cultural Empirical Validation Using Item Response Theory

Scott I. Donaldson¹ · Stewart I. Donaldson² · Michelle McQuaid³ · Margaret L. Kern³

Accepted: 9 June 2023
© The Author(s) 2023

Abstract

A substantial body of empirical research has used the positive emotions, engagement, relationships, meaning, accomplishment (PERMA) framework to measure building blocks of well-being across diverse samples and cultures, with most studies using the 23-item PERMA-Profiler (Butler & Kern, 2016) or a workplace variant. Donaldson and Donaldson (2021a) added four additional domains (physical health, mindset, environment, economic security; PERMA+4). Psychometric development and testing of the original, translated, and variant versions of the measure have relied on Classical Test Theory approaches, such as factor analytic methods. In the workplace, valid, brief measures are critical. The current study used item response theory to analyze data from a large sample of Canadian ($n=1,003$) and Australian ($n=942$) employees to create a 9-item short scale of PERMA+4. A graded response model showed good item discrimination ($a>1.40$), and similar test information compared to the full measure. A short scale of PERMA+4 will be useful for future studies of the building blocks of well-being and positive functioning, as well as for evaluating well-being programs and interventions within the workplace.

Keywords Well-being · PERMA · Item response theory · Measurement · Workplace wellbeing

✉ Scott I. Donaldson
scottdon@usc.edu

¹ Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, 1845 N Soto Street, 3rd Floor, SSB 310-23, Los Angeles, CA 90032, USA

² Department of Psychology, Claremont Graduate University, Claremont, CA, USA

³ Centre for Wellbeing Science, University of Melbourne, Victoria, Australia

Seligman (2011) proposed that flourishing was comprised of five components: **P**ositive emotion, **E**ngagement, **R**elationships, **M**eaning, and **A**ccomplishment (i.e., **PERMA**). Seligman's framework has been commonly measured by the PERMA-Profiler (Butler & Kern, 2016), along with a number of variants (e.g., for the workplace and in childhood/adolescence; e.g., Kern et al., 2014, 2015) and translations of the original measure (e.g., Giangrasso, 2021; Lai et al., 2018). While this is only one measure amongst others that have captured positive mental health, the model and the measure have been used in numerous studies, research programs, and applied programs.

Seligman (2018) noted that PERMA is not an exhaustive list of the building blocks of well-being, and encouraged researchers to explore additional evidence-based building blocks that may improve the PERMA framework (Seligman, 2018). Donaldson and Donaldson (2021a) suggested that there was a theoretical rationale for testing four additional building blocks of well-being (PERMA+4): (1) physical health (i.e., I typically feel physically healthy), (2) mindset (i.e., setbacks are opportunities to grow), (3) environment (i.e., access to nature, natural light), and (4) economic security (i.e., stable resources).

First, physical health was included to measure the impact of health on well-being above and beyond the absence of disease (Seligman, 2008). Physical health focuses on measuring health assets by determining factors that can lead to a variety of positive health outcomes, such as lower health care expenditures, better prognosis when illness does strike, and higher quality of life. Second, mindset was included to measure an open, developable "future-oriented" construct characterized by prospection, growth mindset, and a proclivity towards persevering in the face of setbacks, especially over long periods of time (Duckworth et al., 2007; Dweck, 2006; Luthans et al., 2007). Third, environment was added to measure physical, restorative factors that have been found to maximize our best selves at work (Hartig et al., 1997). Elements may include an abundance of natural light, access to nature, assurance of physiological safety, and organization in the physical arrangement of the workplace (Hartig et al., 1997). Fourth, scholars in the field of positive psychology have long examined the impact of income on well-being. Diener and Seligman (2004) found that individuals who are well-off financially are on average happier than poor people. However, Diener and Seligman (2004) found that differences in income made the most dramatic impact on well-being at varying levels of poverty, presumably as it relates to meeting basic needs. To further examine the impact of income on well-being, economic security was added and defined as an individuals' perception of dimensions they believe are critical to their economic security, such as income stability, job security, and buffers against medical spending shocks.

Donaldson and Donaldson (2021a) developed a 29-item measure of PERMA+4 and found that PERMA+4 was predictive of important work outcomes, such as individual, team, and organizational adaptivity, proactivity, and proficiency. A recent systematic review further supported the original work by Donaldson and Donaldson (2021a) by showing strong associations between PERMA+4 and desirable workplace outcomes (Cabrera & Donaldson, 2023). It has been suggested that PERMA+4 may serve as robust framework for the design, measurement, and evaluation of work-related well-being programs and interventions (Donaldson et al., (2022).

In the workplace, brief measures are critical, both to allow repeated measurement to occur (e.g., assessing the “pulse” of employees), and because time available for employees to complete surveys is limited (Lang & Tay, 2021). Past research has shown that findings from short surveys can be more valid, reliable, and result in higher response rates compared to longer surveys (Kost & Correa da Rosa, 2018). In addition, studies have suggested that short scales create opportunities for more advanced research designs (e.g., ecological momentary assessment) and theory building (Ziegler et al., 2014).

To date, measurement studies on PERMA and PERMA+4 have used Classical Test Theory (CTT), with methods such as factor analysis to develop and validate the PERMA measures (cf. Butler and Kern, 2016; Giangrasso, 2021; Iasiello et al., 2017; Kern and Khaw, 2015; Ryan et al., 2019; Umucu et al., 2020), including the research that validated the four additional building blocks (Donaldson and Donaldson, 2021b). While CTT is useful for understanding the reliability and validity of latent traits for new tests, interpretations of respondents abilities are dependent on the test used (Diener et al., 2018). In other words, if items with poor discrimination between participants are used to assess a construct, interpretations of such findings may misclassify those who are high versus low on the latent trait. Item response theory (IRT) can ensure that a scale is suitable for respondents possessing different levels of the construct under scrutiny (i.e., sufficient information for very high or low scorers). Additionally, CTT assumes that errors of measurement are constant for all respondents and does not account for respondents that possess varying levels of the latent traits, and thus varying levels of error (Zanon et al., 2016). To compliment CTT, IRT can provide ample information across the latent spectrum, which enables predictive validity with desired outcome measures.

Item response theory is a scale development tool that has been used to shorten existing scales, and has been shown to measure the well-being and performance of employees (Lang & Tay, 2021; Nima et al., 2020). Item response theory compliments CTT by providing a person-by-item interaction, showing the quality of measurement at specific points on the latent trait (Oishi, 2006). Unlike CTT that applies measurement error across the entire sample, measurement error in IRT varies depending on the latent score (Oishi, 2006). Past research has shown that IRT has been successfully used to create short versions of existing instruments (Donaldson et al., 2021; Petersen et al., 2006; Sekely et al., 2018). However, to our knowledge, IRT has not been applied to the measurement of PERMA or PERMA+4.

The current study used IRT to develop and evaluate the psychometric properties of a short scale of PERMA+4, including an examination of item difficulty and discrimination, test information, and a direct comparison of test information in the short version of PERMA+4 compared to the longer version of PERMA+4. Findings from this study may be used to help workplace well-being coaches and practitioners accurately screen for nine dimensions of well-being in the workplace to best inform programs and interventions.

1 Methods

1.1 Participants and Procedure

Participants were recruited from two independent data sources, representing full-time employees in Canada ($n=1,003$) and Australia ($n=942$). Table S1 provides a sociodemographic breakdown of the Australian and Canadian samples. Participants were recruited using an online panel agency, which collected a representative sample based on sociodemographic characteristics. Participants were paid between \$2.00–\$3.00 for completing the survey. Participants were provided with a link and were directed to an online survey. After providing informed consent, participants completed survey items assessing their work-related well-being. The last part of the survey asked participants to report demographic characteristics. The survey took approximately 20 min to complete. All research materials and procedures were approved by Claremont Graduate University's Institutional Review Board.

1.2 Measures

PERMA+4. One item from each of the PERMA dimensions of the PERMA-Profiler (Butler & Kern, 2016) were included, along with an additional item on physical health. Three items on each additional PERMA building block—mindset, environment, and economic security—were adapted from the Positive Functioning at Work Scale (PF-W) (Donaldson & Donaldson, 2021a), resulting in a 15-item measure. Past research has supported the internal consistency, reliability, and convergent and discriminant validity of the PERMA+4 (Donaldson & Donaldson, 2021a). For each item, participants were asked to consider how well they have felt and functioned at work over the past two weeks. They were instructed to indicate the extent to which they agreed with each statement, ranging from 0 (not at all, 0%) to 10 (completely, 100%).

Demographic variables. Gender was coded as male, female, non-binary, or not-applicable. Age was coded as 25–34, 35–44, 45–54, 55–65, or 66–99. Education was coded as still at school, less than a high school diploma, high school diploma, some college, but no degree, vocational training, Bachelor's, Master's, Doctorate, or other. Race/ethnicity was measured as White/Caucasian, Asian, Indigenous, Black, Hispanic/Latinx, Middle Eastern or North African, Pacific Islander, Bi-Racial/Multi-Racial, other, or not applicable. Work sector was coded as privately funded organization, publicly listed organization, government funded organization, not for profit, or other.

1.3 Analytic Strategy

All analyses were conducted in the R statistical program using the *mirt*, *psych*, *lordif*, *mokken*, and *stats* packages (Chalmers, 2012; Chambers et al., 1990; Choi et al., 2011; R Core Team, 2021; Revelle, 2017; van der Ark, 2007). Descriptive statistics, including means, standard deviations, response frequencies, and skewness and kurtosis were computed. Byrne (2010) argued that data is considered to be normal

if skewness is between -2 to +2 and kurtosis is between -7 to +7. Missing data was handled using listwise deletion.

Prior to implementing a parametric IRT model, a Mokken Scale Analysis and differential item functioning procedure were performed to examine the dimensionality and item invariances of PERMA+4 (Choi et al., 2011; Mokken, 2011). To assess unidimensionality, item-pair scalability coefficients were computed using H. A strong unidimensional scale is denoted by $H > 0.5$, $0.4 < H < 0.5$ denotes a medium scale, and $H < 0.40$ denotes a weak scale (Mokken, 2011). An automated item selection procedure was used to calculate inter-item covariances and the relationship between items and the latent trait (Meijer & Baneke, 2004). Latent monotonicity for each item was examined using a visual plot of item step response function by rest score group. The rest score group is achieved by summing the overall score minus the score on each item (Junker & Sijtsma, 2000). Per the recommendation of Robinson et al. (2019), a residual correlation matrix with a cutoff < 0.25 was used to assess local independence between the PERMA+4 items. A one-way analysis of variance test was conducted to check for balance between the age and gender categories on PERMA+4. Nonsignificant differences were found for age ($p = 0.189$) and gender ($p = 0.450$), suggesting that there were no pre-existing differences in age and gender on PERMA+4.

To test for measurement invariance between the Canadian and Australian samples before performing parametric IRT on the combined samples, a differential item functioning procedure was used to detect uniform and non-uniform differential item functioning (Choi et al., 2011). A chi-square likelihood ratio test was used as the detection criterion at the alpha level of 0.01, and McFadden's pseudo (R^2) was used as the magnitude measure. Four plots were produced for each item to visualize item response models between the two groups. The first plot showed the item characteristic curves (ICCs) between the two groups, the second plot showed the absolute differences between the ICCs for the two groups, the third plot showed the item response functions for the two groups, and the fourth plot showed the absolute differences between the ICCs weighted by the score distribution.

A Samejima's graded response model was used to estimate item discrimination and ability on the latent trait of the PERMA+4 (Woods, 2006). The graded response model is an extension of the two-parameter logistic model for items with two or more response categories (Samejima, 1997). Three parameters were estimated and used to evaluate each item of the 15-item measure of PERMA+4: 1) the ability level of PERMA+4 was denoted as θ (Θ); 2) an index of difficulty was denoted as b_i , with items that represent the top end of the PERMA+4 being considered more difficult; 3) the discrimination of the item was denoted as a . Items that accurately differentiate an individual that is low on dimensions of PERMA+4 from individuals that are high on dimensions of PERMA+4 received a higher discrimination score. Test information was used to estimate the items precision across the latent trait, and item and option characteristics curves were used to screen the ability of items to progress along the latent trait given the response categories (i.e., 0–10). Conditional reliability was used to compare the reliability of the 15-item and 9-item measure of PERMA+4 along the latent trait.

Based on past research (Sekely et al., 2018), two approaches were used to select the best item that represented mindset, environment, and economic security: a the-

ory-based approach (i.e., item content was evaluated by the research team) and an empirically driven approach (i.e., items with the highest test information, discrimination, and difficulty). Using both approaches allowed for the selection of items that possessed the most information along the latent trait (PERMA+4) and maintained the construct validity of each dimension as defined by Donaldson and Donaldson (2021a).

2 Results

2.1 Item Descriptive Statistics and Response Frequencies

Item means, standard deviations, and response frequencies for the 15-item measure of PERMA+4 are presented in Table 1. Skewness and kurtosis values were considered normal for the 15-item measure of PERMA+4. The mean scores ranged from 4.8 to 7.2. Response frequencies across PERMA+4 items tended to favor the higher response options (i.e., above response option 5 compared to below 5). However, the environment (nature) item and all three economic security items had a considerable number of responses at the lower end of the Likert-scale. All response options were included for item response modeling.

2.2 Dimensionality

A Mokken Scale Analysis of the 15-item measure of PERMA+4 conducted separately for the Australian and Canadian samples showed a medium item-pair scalability coefficient ($H=0.420$). An automated item selection procedure and visual inspection of plots further supported the dimensionality of the 15-item measure of PERMA+4, by showing a monotonic relationship with the rest score group. All residual correlations between items were below the recommended <0.25 (Robinson et al., 2019), suggesting local independence. A differential item functioning analysis found that seven items were flagged for uniform and non-uniform differential item functioning. However, a graphical display of the differential item functioning on each item of the 15-item measure of PERMA+4 showed that the absolute differences between the item characteristic curves mainly occurred at low levels (theta values <-2) of PERMA+4, suggesting minimal impact. The two samples (i.e., Canadian and Australian) were then combined for item response modeling.

2.3 Item Response Modeling

Table 2 shows the graded response model item parameters for the 15-item measure of PERMA+4. The discrimination parameters for the five items used to measure PERMA ranged from 2.16 to 3.34, and were classified as possessing very high discrimination (Baker, 2001). The discrimination parameters for the four additional PERMA factors ranged from 0.80 (moderate) to 2.80 (very high) (Baker, 2001). In terms of item difficulty thresholds (i.e., the point at which each item is 50% likely to be endorsed from the preceding response options along the continuum of theta) 0

Table 1 15-item measure of PERMA+4 with item means, standard deviations, and response frequency values

Dimension (subdimension)	Item	M (SD)	Response frequency										
			0	1	2	3	4	5	6	7	8	9	10
Positive emotion	I felt positive at work.	6.5 (2.4)	71	24	42	69	106	265	252	359	407	159	191
Engagement	I was deeply engaged and interested in my work.	6.6 (2.4)	59	25	53	67	101	269	248	345	365	206	207
Relationships	I was encouraging and supportive of others.	7.2 (2.2)	55	11	17	33	50	213	197	365	465	265	274
Meaning	I felt that the work I did was valuable and worthwhile.	6.8 (2.4)	62	22	40	62	83	240	207	357	390	197	285
Accomplishment	I set and achieved clear goals.	6.8 (2.3)	50	14	37	57	67	258	248	365	422	209	218
Health	I felt physically healthy and strong.	6.4 (2.4)	59	19	60	100	114	278	272	355	330	179	179
Mindset (growth mindset)	I believed I could improve my job skills through hard work.	7.1 (2.1)	42	7	14	38	54	252	229	380	422	232	275
Mindset (prospectron job)	I believed my job would allow me to develop in the future.	6.4 (2.6)	102	36	52	66	95	275	254	335	342	182	206
Mindset (prospectron work)	I have a bright future at my current workplace.	6.0 (2.8)	147	42	69	79	95	313	241	308	296	166	189
Environment (focus)	My physical work environment (e.g., office space) allowed me to focus on my work.	6.6 (2.4)	73	24	34	71	81	269	248	349	387	193	216
Environment (light)	There was plenty of natural light in my workplace.	6.6 (2.8)	117	28	50	85	89	233	171	319	336	212	305
Environment (nature)	I could conveniently access nature in my work environment (e.g., parks, oceans, mountains, etc.).	5.4 (3.1)	243	72	117	121	106	284	188	261	219	123	211
Economic security (income)	I was comfortable with my current income.	5.5 (2.8)	170	55	102	116	163	278	247	273	264	133	144
Economic security (medical)	I could lose several months of pay due to serious illness, and still have my economic security.	4.8 (3.3)	363	97	123	128	110	258	184	216	192	108	166
Economic security (savings)	In the event of a financial emergency, I had adequate savings.	5.1 (3.2)	293	84	127	95	120	269	194	241	229	115	179

Note. M = mean; SD = standard deviation; N = 1,945.

versus 1 ranged (b_1) from -2.96 to -1.64 , 1 to 2 (b_2) ranged from -2.84 to -1.48 , 2 to 3 (b_3) ranged from -2.63 to -1.25 , 3 to 4 (b_4) ranged from -2.25 to -0.85 , 4 to 5 (b_5) ranged from -1.91 to -0.53 , 5 to 6 (b_6) ranged from -1.07 to 0.19 , 6 to 7 (b_7) ranged from -0.65 to 0.74 , 7 to 8 (b_8) ranged from -0.07 to 1.5 , 8 to 9 (b_9) ranged from 0.7 to 2.42 , and 9 to 10 (b_{10}) ranged from 1.19 to 3.21 .

Option characteristic curves were plotted for the 15-item measure of PERMA+4 (see Fig. 1). The item discrimination coefficient (a) reflects the steepness or the slope of the curve. Items that had higher discrimination (e.g., positive emotion [$a=3.06$], engagement [$a=3.34$]) were endorsed along theta as the response options increased. On the other hand, items that had lower discrimination (e.g., economic security (medical) [$a=0.80$], economic security (savings) [$a=0.86$]) were not able to distinguish between response options as theta increased.

2.4 Test and Item Information Functions

Based on results from the item response modeling and an examination of item content, a 9-item short version of PERMA+4 (see Table 3) was tested against the 15-item measure of PERMA+4: positive emotion, engagement, relationships, meaning, accomplishment, health, mindset (prospection work), environment (focus), and economic security (income). Figure 2 shows a conditional reliability plot comparing the 15-item measure of PERMA+4 and 9-item measure of PERMA+4. Visual inspection of Fig. 2 shows that score estimates were most reliable in the -3 to $+2$ theta range. The item information function for the 9-item measure of PERMA+4 shows that items positive emotions, engagement, meaning, and accomplishment possessed the most test information relative to the total PERMA+4 score (Fig. 3).

3 Discussion

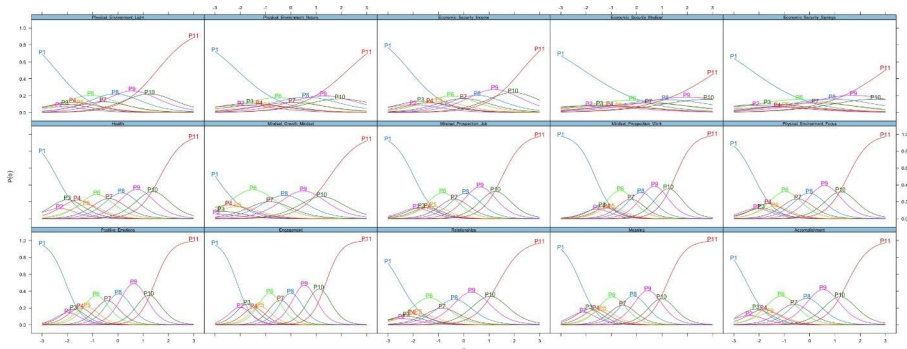
This study developed and tested a 9-item short scale of PERMA+4 using item response theory from two large samples of employees in Canada and Australia. The findings showed that 9 items, one item representing each building block of PERMA+4, had good item discrimination and provided comparable test information compared to a 15-item measure of PERMA+4. Taken together, these findings may be useful to well-being researchers who are interested in using a short, robust, and reliable measure of PERMA+4, or for practitioners who want to reduce participant burden while maintaining a comprehensive metric of well-being in “real world” settings.

The present study found that most respondents in the Canadian and Australian workplace reported relatively high levels of PERMA. However, reports on the additional PERMA+4 building blocks of well-being, including on economic security and environment, were found to be much lower than the five PERMA elements. Employees’ perceptions of their economic security and immediate physical environment may have conjured up feelings of uncertainty or negative emotions. Past research has shown that individuals fear economic losses, and when they experience such losses, their well-being also suffers (Hacker & Jacobs, 2008). Global crises, such as the COVID-19 pandemic and the war in Ukraine may have exacerbated feelings of

Table 2 Graded response model item parameters for the 15-item measure of PERMA+4

Item (subitem)	a	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
Positive emotion	3.06	-2.05	-1.86	-1.62	-1.36	-1.08	-0.61	-0.23	0.27	0.97	1.44
Engagement	3.34	-2.12	-1.88	-1.58	-1.33	-1.07	-0.59	-0.24	0.22	0.83	1.37
Relationships	2.16	-2.48	-2.34	-2.16	-1.91	-1.66	-1.03	-0.65	-0.07	0.70	1.35
Meaning	2.79	-2.21	-1.99	-1.73	-1.46	-1.21	-0.71	-0.37	0.13	0.75	1.19
Accomplishment	2.36	-2.46	-2.29	-1.97	-1.66	-1.41	-0.81	-0.39	0.16	0.90	1.49
Health	2.24	-2.40	-2.19	-1.79	-1.40	-1.10	-0.57	-0.14	0.41	1.07	1.66
Mindset (growth mindset)	1.70	-2.96	-2.84	-2.63	-2.25	-1.91	-1.07	-0.59	0.08	0.89	1.52
Mindset (prospection job)	2.48	-1.96	-1.76	-1.54	-1.32	-1.08	-0.56	-0.18	0.32	0.96	1.51
Mindset (prospection work)	2.80	-1.64	-1.48	-1.27	-1.08	-0.90	-0.40	-0.06	0.40	0.97	1.51
Environment (focus)	2.40	-2.19	-1.99	-1.78	-1.48	-1.24	-0.68	-0.28	0.24	0.94	1.49
Environment (light)	1.44	-2.42	-2.22	-1.96	-1.62	-1.35	-0.77	-0.42	0.20	0.92	1.54
Environment (nature)	1.11	-2.11	-1.81	-1.41	-1.08	-0.82	-0.18	0.25	0.91	1.63	2.22
Economic security (income)	1.42	-2.13	-1.87	-1.50	-1.17	-0.80	-0.24	0.23	0.81	1.59	2.27
Economic security (medical)	0.80	-2.08	-1.68	-1.25	-0.85	-0.53	0.19	0.74	1.50	2.42	3.21
Economic security (savings)	0.86	-2.29	-1.91	-1.43	-1.13	-0.78	-0.05	0.49	1.24	2.21	2.96

Note. a =discrimination; b =difficulty; N =1,945.

**Fig. 1** Option characteristic curves for 15-item measure of PERMA+4

Note. P1=the probability of respondents endorsing response option 0 (not at all); P11=the probability of respondents endorsing the response option 10 (completely)

Table 3 9-item short measure of PERMA+4

Dimension (subdimension)	Item
Positive emotion	I felt positive at work.
Engagement	I was deeply engaged and interested in my work.
Relationships	I was encouraging and supportive of others.
Meaning	I felt that the work I did was valuable and worthwhile.
Accomplishment	I set and achieved clear goals.
Health	I felt physically healthy and strong.
Mindset (prospection work)	I had a bright future at my current workplace.
Environment (focus)	My physical work environment (e.g., office space) allowed me to focus on my work.
Economic security (income)	I was comfortable with my current income.

Fig. 2 Conditional reliability plot comparing the 15-item and 9-item measure of PERMA+4
 Note. θ =the level of PERMA+4; r_{xx} = reliability value from 0 to 1

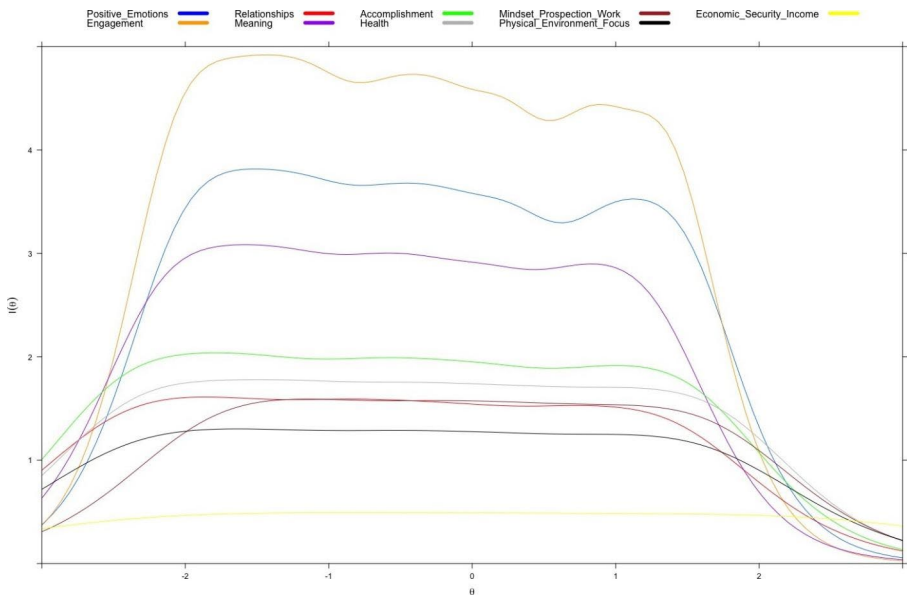
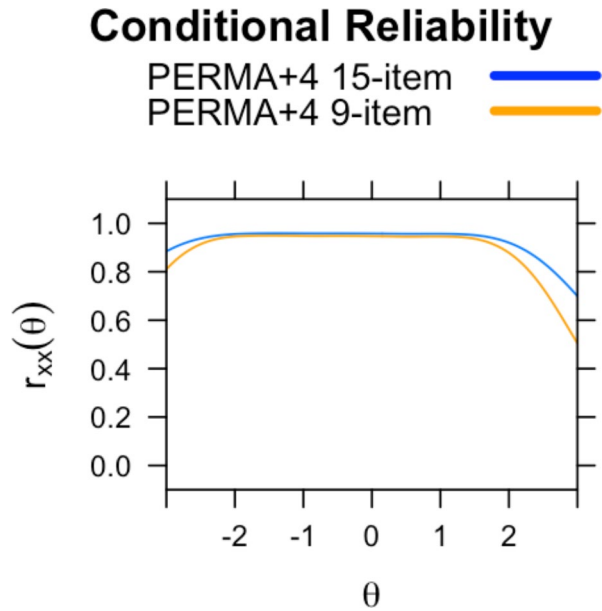


Fig. 3 Item information function for the 9-item measure of PERMA+4
 Note. θ =the level of PERMA+4; I =information; information is an arbitrary value estimated using the latent trait

economic insecurity and threats to the physical environment for participants in the current study. Nonetheless, research supports that perceptions of economic security and restorative physical features of our environment can impact employee well-being (Bellini et al., 2015; Diener & Seligman, 2004; Easterlin, 2003). Scholars and practitioners that design work-related well-being programs and interventions should consider economic and environmental aspects to support wellbeing.

Numerous studies have validated measures of PERMA across countries and populations (Lai et al., 2018; Pezirkianidis et al., 2021; Watanabe et al., 2018). These studies have primarily relied on CTT methods, whereas findings from the present study advanced measurement precision in this literature by using item response theory. Findings from the item response model showed that the 9-item measure of PERMA+4 had good item discrimination. Thresholds estimates and option characteristic curves also showed acceptable separation along the latent trait. Most respondents endorsed items higher on the scale (i.e., 6 and beyond), even if their scores were found to be average on the latent trait. These findings suggest that this 9-item measure of PERMA+4 may accurately reflect the workplace well-being levels of typical employees.

Past systematic reviews have shown that employees with higher levels of PERMA+4 were associated with increased job satisfaction and job performance (Donaldson et al., 2019; Roll et al., 2019). Likewise, the present study found that the 9-item measure of PERMA+4 was conditionally reliable between theta values of -3 to 2, representing most of the sample. These findings suggest that practitioners and researchers may use the total score across the nine items to measure the overall construct of PERMA+4 in the workplace. Future research should assess the predictive ability of the 9-item measure of PERMA+4 with other work outcomes, such as work performance, and assess convergent validity with similar well-being measures (e.g., psychological capital). Studies might also test the 9-item measure of PERMA+4 in various workplace industries and validate the scale among other populations like college students or school children. Further longitudinal studies are needed to assess how PERMA+4 varies within individuals, between individuals, and the subsequent impact on work performance.

3.1 Limitations

There were restrictions on the number of items that could be administered from the original 29-item PERMA+4 developed by Donaldson and Donaldson (2021a)). As such, the 15 items selected to represent PERMA+4 may have impacted the findings. The data were collected using a cross-sectional, self-reported survey instrument and may be prone to social desirability bias (Donaldson & Grant-Vallone, 2002). Age was measured using a band response option (i.e., a range of ages such as 25–34) and thus specific age-related differences may not have been detected than if a continuous variable was collected.

4 Conclusion

The present study found support for a 9-item measure of PERMA+4 using item response theory. The item level information builds on past research that used CTT to develop and validated PERMA-related measures, adding an IRT analysis and reduction of items needed to adequately capture the PERMA+4 construct. Scholars and practitioners are encouraged to use this short scale to design, measure, and evaluate workplace programs and interventions that assess well-being.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s41042-023-00110-9>.

Funding No funding was obtained for this study.
Open access funding provided by SCEL, Statewide California Electronic Library Consortium

Data Availability All data have been made publicly available at the [https://github.com/SDonaldsonUSC/PERMA4_IRT] and can be accessed at [https://github.com/SDonaldsonUSC/PERMA4_IRT/blob/main/McQuaid_IRT.Rmd].

Declarations

Ethical Approval Claremont Graduate University Institutional Review Board approved this study.

Competing Interests None declared.

Informed Consent Not Applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baker, F. B. (2001). *The basics of item response theory*. For full text: <https://eric.ed.gov/?id=ED458219>
- Bellini, D., Ramaci, T., & Bonaiuto, M. (2015). The restorative effect of the environment on organizational cynicism and work engagement. *Journal of Human Resource and Sustainability Studies*, 3(3), <https://doi.org/10.4236/jhrss.2015.33017>
- Butler, J., & Kern, M. L. (2016). The PERMA-Profiler: A brief multidimensional measure of flourishing. *International Journal of Wellbeing*, 6(3), 1–48. <https://doi.org/10.5502/ijw.v6i3.526>
- Byrne, B. M. (2010). *Structural equation modeling with amos: Basic concepts, applications, and programming*, 2nd ed (pp. xix, 396). Routledge/Taylor & Francis Group.
- Cabrera, V., & Donaldson, S. I. (2023). PERMA to PERMA + 4 building blocks of well-being: A systematic review of the empirical literature. *The Journal of Positive Psychology*, 0(0), 1–20. <https://doi.org/10.1080/17439760.2023.2208099>

- Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the r environment. *Journal of Statistical Software*, 48, 1–29. <https://doi.org/10.18637/jss.v048.i06>
- Chambers, J., Hastie, T., & Pregibon, D. (1990). Statistical Models in S. In K. Momirović & V. Mildner (Eds.), *Compstat* (pp. 317–321). Physica-Verlag HD. https://doi.org/10.1007/978-3-642-50096-1_48
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An r package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte carlo simulations. *Journal of Statistical Software*, 39(8), 1–30.
- R Core Team (2021). *R: A language and environment for statistical computing*. <https://www.R-project.org/>
- Diener, E., & Seligman, M. E. P. (2004). Beyond money: Toward an economy of well-being. *Psychological Science in the Public Interest*, 5, 1–31. <https://doi.org/10.1111/j.0963-7214.2004.00501001.x>
- Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and open questions in the science of subjective well-being. *Collabra: Psychology*, 4(1), 15. <https://doi.org/10.1525/collabra.115>
- Donaldson, S. I., & Donaldson, S. I. (2021a). The Positive Functioning at Work Scale: Psychometric assessment, validation, and measurement invariance. *Journal of Well-Being Assessment*, 4(2), 181–215. <https://doi.org/10.1007/s41543-020-00033-1>
- Donaldson, S. I., & Donaldson, S. I. (2021b). Examining PERMA + 4 and work role performance beyond self-report bias: Insights from multitrait-multimethod analyses. *The Journal of Positive Psychology*, 1–10. <https://doi.org/10.1080/17439760.2021.1975160>
- Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, 17(2), 245–260. <https://doi.org/10.1023/A:1019637632584>
- Donaldson, S. I., Lee, J. Y., & Donaldson, S. I. (2019). Evaluating positive psychology interventions at work: A systematic review and meta-analysis. *International Journal of Applied Positive Psychology*, 4(3), 113–134. <https://doi.org/10.1007/s41042-019-00021-8>
- Donaldson, S. I., Strong, D., & Zhu, S. H. (2021). A brief measure of problematic smartphone use among high school students: Psychometric assessment using item response theory. *Computers in Human Behavior Reports*, 3, 100091. <https://doi.org/10.1016/j.chbr.2021.100091>
- Donaldson, S. I., van Zyl, L. E., & Donaldson, S. I. (2022). PERMA + 4: A framework for work-Related wellbeing, performance and positive organizational psychology 2.0. *Frontiers in Psychology*, 12, 817244. <https://doi.org/10.3389/fpsyg.2021.817244>
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92, 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>
- Dweck, C. S. (2006). *Mindset: The new psychology of success* (pp. x–276). Random House.
- Easterlin, R. A. (2003). Building a better theory of well-being. *SSRN Scholarly Paper No 392043*. <https://doi.org/10.2139/ssrn.392043>
- Giangrasso, B. (2021). Psychometric properties of the PERMA-Profiler as hedonic and eudaimonic well-being measure in an italian context. *Current Psychology*, 40(3), 1175–1184. <https://doi.org/10.1007/s12144-018-0040-3>
- Hacker, J., & Jacobs, E. (2008, May 27). *Income volatility: Another source of growing economic insecurity*. Economic Policy Institute. https://www.epi.org/publication/webfeatures_snapshots_20080528/
- Hartig, T., Korpela, K., Evans, G. W., & Gärling, T. (1997). A measure of restorative quality in environments. *Scandinavian Housing and Planning Research*, 14(4), 175–194. <https://doi.org/10.1080/02815739708730435>
- Iasiello, M., Bartholomaeus, J., Jarden, A., & Kelly, G. (2017). Measuring PERMA + in South Australia, the state of wellbeing: A comparison with national and international norms. *Journal of Positive Psychology and Wellbeing*, 1(2), Article 2.
- Junker, B. W., & Sijsma, K. (2000). Latent and manifest monotonicity in item response models. *Applied Psychological Measurement*, 24(1), 65–81. <https://doi.org/10.1177/01466216000241004>
- Kern, M., & Khaw, D. (2015). A cross-cultural comparison of the PERMA model of well-being. *Undergraduate Journal of Psychology at Berkeley*, 8. <https://findanexpert.unimelb.edu.au/scholarlywork/1052981-a-cross-cultural-comparison-of-the-perma-model-of-well-being>
- Kern, M. L., Waters, L., Adler, A., & White, M. (2014). Assessing employee wellbeing in schools using a multifaceted approach: Associations with physical health, life satisfaction, and professional thriving. *Psychology*, 5, 500–513. <https://doi.org/10.4236/psych.2014.56060>
- Kern, M. L., Waters, L. E., Adler, A., & White, M. A. (2015). A multidimensional approach to measuring well-being in students: Application of the PERMA framework. *The Journal of Positive Psychology*, 10(3), 262–271. <https://doi.org/10.1080/17439760.2014.936962>

- Kost, R. G., & da Correa, J. (2018). Impact of survey length and compensation on validity, reliability, and sample characteristics for ultrashort-, short-, and long-research participant perception surveys. *Journal of Clinical and Translational Science*, 2(1), 31–37. <https://doi.org/10.1017/cts.2018.18>
- Lai, M. K., Leung, C., Kwok, S. Y. C., Hui, A. N. N., Lo, H. H. M., Leung, J. T. Y., & Tam, C. H. L. (2018). A multidimensional PERMA-H positive education model, general satisfaction of school life, and character strengths use in Hong Kong senior primary school students: Confirmatory factor analysis and path analysis using the APASO-II. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01090>
- Lang, J. W. B., & Tay, L. (2021). The science and practice of item response theory in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 8, 311–338. <https://doi.org/10.1146/annurev-orgpsych-012420-061705>
- Luthans, F., Avolio, B. J., Avey, J. B., & Norman, S. M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psychology*, 60, 541–572. <https://doi.org/10.1111/j.1744-6570.2007.00083.x>
- Meijer, R. R., & Baneke, J. J. (2004). Analyzing psychopathology items: A case for nonparametric item response theory modeling. *Psychological Methods*, 9, 354–368. <https://doi.org/10.1037/1082-989X.9.3.354>
- Mokken, R. J. (2011). A theory and procedure of scale analysis: With applications in political research. *A Theory and Procedure of Scale Analysis*. De Gruyter Mouton. <https://doi.org/10.1515/9783110813203>
- Nima, A. A., Cloninger, K. M., Persson, B. N., Sikström, S., & Garcia, D. (2020). Validation of subjective well-being measures using item response theory. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.03036>
- Oishi, S. (2006). The concept of life satisfaction across cultures: An IRT analysis. *Journal of Research in Personality*, 40(4), 411–423. <https://doi.org/10.1016/j.jrp.2005.02.002>
- Petersen, M. A., Groenvold, M., Aaronson, N., Blazeby, J., Brandberg, Y., de Graeff, A., Fayers, P., Hammerlid, E., Sprangers, M., Velikova, G., Bjorner, J. B., & European Organisation for Research and Treatment of Cancer Quality of Life Group. (2006). Item response theory was used to shorten EORTC QLQ-C30 scales for use in palliative care. *Journal of Clinical Epidemiology*, 59(1), 36–44. <https://doi.org/10.1016/j.jclinepi.2005.04.010>
- Pezirkianidis, C., Stalikas, A., Lakioti, A., & Yotsidi, V. (2021). Validating a multidimensional measure of wellbeing in Greece: Translation, factor structure, and measurement invariance of the PERMA profiler. *Current Psychology*, 40(6), 3030–3047. <https://doi.org/10.1007/s12144-019-00236-7>
- Revelle, W. R. (2017). *Psych: Procedures for personality and psychological research*. Software. <https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research>
- Robinson, M., Johnson, A. M., Walton, D. M., & MacDermid, J. C. (2019). A comparison of the polytomous rasch analysis output of RUMM2030 and R (ltm/eRm/TAM/lordif). *BMC Medical Research Methodology*, 19(1), 36. <https://doi.org/10.1186/s12874-019-0680-5>
- Roll, L. C., Van Zyl, L. E., & Griep, Y. (2019). Brief positive psychological interventions within multi-cultural organizational contexts: A systematic literature review. In L. E. Van Zyl & S. Rothmann Sr. (Eds.), *Theoretical Approaches to Multi-Cultural Positive Psychological Interventions* (pp. 523–544). Springer International Publishing. https://doi.org/10.1007/978-3-030-20583-6_24
- Ryan, J., Curtis, R., Olds, T., Edney, S., Vandelandotte, C., Plotnikoff, R., & Maher, C. (2019). Psychometric properties of the PERMA profiler for measuring wellbeing in Australian adults. *PLOS ONE*, 14(12), e0225932. <https://doi.org/10.1371/journal.pone.0225932>
- Samejima, F. (1997). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 85–100). Springer. https://doi.org/10.1007/978-1-4757-2691-6_5
- Sekely, A., Taylor, G. J., & Bagby, R. M. (2018). Developing a short version of the Toronto structured interview for Alexithymia using item response theory. *Psychiatry Research*, 266, 218–227. <https://doi.org/10.1016/j.psychres.2018.03.002>
- Seligman, M. E. P. (2008). Positive health. *Applied Psychology*, 57(s1), 3–18. <https://doi.org/10.1111/j.1464-0597.2008.00351.x>
- Seligman, M. (2018). PERMA and the building blocks of well-being. *The Journal of Positive Psychology*, 13(4), 333–335. <https://doi.org/10.1080/17439760.2018.1437466>
- Umucu, E., Wu, J. R., Sanchez, J., Brooks, J. M., Chiu, C. Y., Tu, W. M., & Chan, F. (2020). Psychometric validation of the PERMA-profiler as a well-being measure for student veterans. *Journal of American College Health*, 68(3), 271–277. <https://doi.org/10.1080/07448481.2018.1546182>

- van der Ark, L. A. (2007). Mokken scale analysis in r. *Journal of Statistical Software*, 20(11), 1–19. <https://doi.org/10.18637/jss.v020.i11>
- Watanabe, K., Kawakami, N., Shiotani, T., Adachi, H., Matsumoto, K., Imamura, K., Matsumoto, K., Yamagami, F., Fusejima, A., Muraoka, T., Kagami, T., Shimazu, A., & Kern, M. L. (2018). The Japanese workplace PERMA-Profil: A validation study among Japanese workers. *Journal of Occupational Health*, 60(5), 383–393. <https://doi.org/10.1539/joh.2018-0050-OA>
- Woods, C. M. (2006). Ramsay-curve item response theory (RC-IRT) to detect and correct for nonnormal latent variables. *Psychological Methods*, 11, 253–270. <https://doi.org/10.1037/1082-989X.11.3.253>
- Zanon, C., Hutz, C. S., Yoo, H., Henry, & Hambleton, R. K. (2016). An application of item response theory to psychological test development. *Psicologia: Reflexão e Crítica*, 29(1), 18. <https://doi.org/10.1186/s41155-016-0040-x>
- Ziegler, M., Kemper, C. J., & Kruey, P. (2014). Short scales – five misunderstandings and ways to overcome them. *Journal of Individual Differences*, 35, 185–189. <https://doi.org/10.1027/1614-0001/a000148>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.