Original Article

Measurement Invariance Across Age Groups and Over 20 Years' Time of the Negative and Positive Affect Scale (NAPAS)

Meingold Hiu-ming Chan, Micah Gerhardt, and Xin Feng

Department of Human Sciences, Ohio State University, Columbus, OH, USA

Abstract: The factor structure and measurement invariance across gender of Mroczek's and Kolarz's scales of positive affect (PA) and negative affect (NA) have been examined in past studies; however, little is known about the measurement invariance across age groups and over time, which are important psychometric properties for developmental research. The current study sought to fill this gap using the data from the National Survey of Midlife Development in the United States (MIDUS). Multi-group confirmatory factor analysis (CFA) was used to examine increasing levels of measurement invariance across gender and age groups. Longitudinal CFA was also used to test measurement invariance over three time points using the data from MIDUS 1 (N = 3,748), MIDUS 2 (N = 2,257), and MIDUS 3 (N = 1,414). Results supported full scalar invariance across gender, age groups, and over time. The latent means for NA were significantly different between men and women at time 1 and 2, but not at time 3; the latent means for both PA and NA were also different across age groups. There were no significant differences for PA and only trivial differences for NA over time within individuals. Implications of these results for longitudinal research are discussed.

Keywords: negative affect, positive affect, NAPAS, measurement invariance, longitudinal

It has been well-established that positive affect (PA) and negative affect (NA) are important predictors and correlates of many life outcomes. PA is generally defined as high levels of energy, enthusiasm, and pleasure; in contrast, NA reflects unpleasant moods, such as sadness, anger, and fear (Merz et al., 2013). Reviews of PA and NA suggest that individuals with higher PA have better work achievement (Staw, Sutton, & Pelled, 1994), are more likely to be employed and have higher levels of socioeconomic status (Diener, Nickerson, Lucas, & Sandvik, 2002), have better marital relationships (Lyubomirsky, King, & Diener, 2005), and are more likely to perform prosocial behaviors (Thoits & Hewitt, 2001). In contrast, lower PA and higher NA are associated with worse physical health, higher risk of mental illnesses (e.g., depression), more risky or unhealthy behaviors (e.g., smoking and substance use), and shorter lifespans (Diener & Chan, 2011).

One widely used PA and NA measure is Mroczek's and Kolarz's Negative and Positive Affect Scale (NAPAS) scale. However, the psychometric properties of this scale have not been thoroughly investigated. In particular, longitudinal measurement invariance, which refers to a consistent structure and relation between items and the latent construct being measured across time has yet to be established. Additionally, it is possible that the conceptual framework of psychological constructs, or the interpretation of the same item (s) differs across age or gender groups (Brown, 2006). Therefore, it is important to ensure that the same latent construct is captured across age, gender, and over time even when using the same measure.

Invariance in Longitudinal Contexts

In recent decades, large-scale longitudinal studies have provided generalizable and expansive information that has afforded researchers an opportunity to study PA and NA over time from a developmental perspective. One of the robust findings from longitudinal studies is that there is a developmental change in PA and NA such that there is a decrease in NA throughout adulthood, particularly among older adults (Charles, Reynolds, & Gatz, 2001). Additionally, when comparing adults in different age groups, older adults report less NA and exhibit less variability on PA than younger adults, which suggests an effect of age on PA and NA (Brim, Ryff, & Kessler, 2004; Mroczek & Kolarz, 1998). It is possible that older adults are better at regulating their

https://doi.org/10.1027/1015-5759/a000529

emotions (Mroczek & Kolarz, 1998) and are more focused on maintaining positive relationships (Charles et al., 2001), which results in developmental changes in affect. However, one limitation of these studies is that they assume, often implicitly, that the measurement of PA and NA is equivalent across age groups and over time.

Invariance across age groups or over time is an important issue in studying life span development. In particular, measurement invariance should be tested when using the same measures for adults across developmental stages (e.g., young, middle, and late adulthood) and in longitudinal studies that follow participants over a long period of time. Among the very few studies that examined measurement invariance across ages, two studies found differential item functioning for some PA items. Mackinnon and colleagues (1999) found that older participants have a lower chance of reporting being excited. Merz et al. (2013) showed that older adults are less likely to report feeling "excited", "proud", or "guilty", but more likely to report feeling "inspired" or "interested". As such, the score of these items varied as a function of age. This may be due to differences in interpretation of the same item across ages or changes in the relation between items and the underlying latent construct. For example, older individuals' PA may be less likely to be reflected by the feeling of excitement when compared to younger individuals. Therefore, it seems possible that the latent structure of PA and NA could be different across age or over time as adults continue to progress developmentally. Since changes in affect over time could be due to measurement non-invariance or developmental changes, it is important to understand the psychometric properties of our measures to separate these two potential sources of change in reported affect over time.

Gender Invariance

Measurement invariance across gender is also rarely tested for PA and NA scales; however, it should be established to ensure that differences across gender are a result of differences in the construct being measured and not an issue of poor measurement. In general, research suggested that women have higher NA and more depressive symptoms than men (Nolen-Hoeksema, 2001; Nolen-Hoeksema & Rusting, 2003). Among the few studies that have tested measurement invariance, some items measuring PA or NA are found to be non-invariant across gender. Crawford and Henry (2004) tested measurement invariance in the Positive and Negative Affect Schedule (PANAS) scale, another widely used and well-established measure of PA/NA, and found partial invariance across gender. Similarly, Mackinnon and colleagues (1999) found differential responses across gender to the item "*nervous*" in the NA scale and the item "*excited*" in the PA scale of a short form PANAS scale. These findings seem to suggest partial measurement invariance across gender for measures assessing PA and NA. However, full metric and scalar invariance across gender was found in the NAPAS scale among a sample of US and Iranian respondents (Joshanloo & Bakhshi, 2015). Nonetheless, evidence for gender invariance in the NAPAS scale is still very limited and should be tested in a different sample.

The Present Study

The current study evaluated the psychometric properties, including factor structure and measurement invariance across gender, age groups, and over time, of Mroczek's and Kolarz's NAPAS. The scale correlates with other subjective well-being measures which reveals convergent validity and yields good reliability, Cronbach's α of .87 and .91 for NA and PA, respectively (Mroczek & Kolarz, 1998). Past studies have also evaluated some aspects of the psychometric properties of NAPAS. Studies have established a two-factor structure of affect (i.e., PA and NA) and established measurement invariance across gender and cultures (Joshanloo, 2017; Joshanloo & Bakhshi, 2015). This revealed promising psychometric properties of the measure and provided important information to research using NAPAS in a cross-sectional context. However, Joshanloo and Bakhshi (2015) did not test for invariance across age groups or over time despite the large variation of age (30–84) in their sample; hence, little is known about its psychometric properties from a developmental perspective. Given how widely used the scale is in developmental research (e.g., Mroczek, 2004; Mroczek & Kolarz, 1998), it is essential to establish invariance across age and longitudinally.

We utilized a three-wave national longitudinal dataset called Midlife Development in the United States (MIDUS) 1, 2, and 3 to fill this gap. The goals of this study were to (1) replicate the factor structure and finding of measurement invariance across gender found in Joshanloo and Bakhshi (2015) but with a different sample and at different age ranges, (2) examine measurement invariance across age groups (young, middle-aged, and older adults), (3) examine longitudinal measurement invariance over a 20-year time frame, and (4) explore latent mean differences across gender/age groups/over time if scalar invariance was established. We hypothesized that the scale would be invariant across gender but partially invariant across age groups. We did not have any hypothesis for invariance over time, as there has not been enough work regarding longitudinal invariance to make reasonable hypotheses. Based on past findings, we expected the latent means of both PA and NA to be different across gender, age groups, and time. Specifically, we expected women to have higher NA than men, and older adults as well as individuals in later time points to report less NA. Findings regarding PA are mixed so we could not form specific hypotheses. If findings from this study support good psychometric properties of this measure, future longitudinal studies that aim at measuring affect over time could consider this measure as an option.

Method

Participants

The current study used data from the random-digit-dial sample within the coterminous United States from MIDUS. The mean age of the participants was 46.42 (SD = 13.23, range = 20–75) when the first wave of MIDUS data was collected in 1995–1996 (T1; N = 3,487; 50.6% female). The second wave (T2) was a longitudinal follow-up of the MIDUS 1 study conducted in 2004–2006 (N = 2,257; 52.4% female, M_{age} = 55.84). From T1 to T2, approximately 70% of respondents were retained. The third wave (T3) of data collection was recently completed in 2013–2014 with 1,414 respondents (52.7% female, M_{age} = 64.84). Attrition analyses showed that those who dropped out at T2 or T3 reported lower ratings on PA and higher ratings on NA at T1.

Measures

Positive and Negative Affect

Respondents completed self-administrated questionnaires including NAPAS at all three time points. Respondents reported during the past 30 days, how much of the time they felt certain PA/NA (six items each) on a 5-point Likert scale from 1 (= *all of the time*) to 5 (= *none of the time*). We reverse coded these items from 0 (= *none of the time*), 4 (= *all of the time*) for easier interpretation. Cronbach's α for NA and PA scales in this study ranged from .851 to .868 and .906 to .911, respectively at the three time points.

Data Analysis

To confirm the factorial structure of PA and NA, we conducted a confirmatory factor analysis (CFA) on data from T1. We fixed the factor mean to 0 and variance to 1. To evaluate goodness of fit, we examined the root mean square error of approximation (RMSEA), comparative fit index (CFI) and standardized root mean square residuals (SRMR) in which a model meeting the following criteria was considered a good fit: RMSEA < .06, CFI > .95, and SRMR < .08, while models with RMSEA < .08 and CFI between .90 and .95 was considered acceptable fit (Bentler & Hu, 1995).

Measurement invariance across gender, age groups, and longitudinally was also tested. The sample was divided into three age groups according to the developmental stages of adulthood (Augustus-Horvath & Tylka, 2011; Santrock, 2008): young adulthood (20–40 years), middle adulthood (41–60 years), and late adulthood (61 years or above). Seven respondents were excluded from the analysis of invariance across age groups since their ages were missing. At T1, the number of subjects in the young, middle, and older group was 1,307, 1,551, and 622, respectively (n = 3,480). At both T2 and T3 the young adulthood group was too small to test measurement invariance. Gender invariance was tested at T1, T2, and T3 separately to ensure invariance at different age ranges before proceeding to test longitudinal invariance.

To test measurement invariance using multi-group and longitudinal CFA, a series of models with increasingly restricted invariance were estimated. According to the guidelines of Widaman, Ferrer, and Conger (2010) and the terminology recommended by a frequently cited review on measurement invariance (Vandenberg & Lance, 2000), three levels of measurement invariance are generally tested, namely configural, metric (or weak), and scalar (or strong) invariance. First, configural invariance (i.e., same pattern of factor loading without equating any parameters) was examined across groups/over time. For identification, we fixed factor means to 0 and variances to 1 across groups/time points. Using latent variables for identification allowed us to test whether the first loadings of the factors were also invariant. Next, to assess metric invariance, factor loadings were constrained to be equal across groups/over time. The variances of the factors were freely estimated in all groups/time points but the reference group (i.e., female in gender model, young group in age model, and T1 in longitudinal model), whereas the factor means of all groups were still constrained to 0. Lastly, factor loadings and intercepts were constrained to be equal across groups/over time to test for scalar invariance. The means and the variances of the factors were freely estimated in all groups but the reference group. The path diagram of the longitudinal invariance model is presented in Electronic Supplementary Materials 1 (ESM 1 – Figure 1). Since the factor mean of the reference group was constrained to 0, the freely estimated factor means of the other groups indicated the latent mean differences between the groups. A p value below .05 indicated a significant latent mean difference. If scalar invariance has been established, latent mean differences can be interpreted as resulting from differing levels of PA and NA across groups/time. We also reported the effect sizes of latent mean differences using Cohen's d with 95% confidence intervals. According to guidelines by Cohen (1988), a Cohen's d of 0.2 is considered a small effect size, 0.5 a medium effect size, and 0.8 a large effect size.

Measurement invariance was established if the model fit of the more restricted model was not significantly worse than the less restricted model. We evaluated the relative model fit based on CFI. If the difference in CFI was less than 0.01, the difference in model fit was considered trivial (Cheung & Rensvold, 1999) and measurement invariance was supported. We also tested the differences of chi-square across nested models. Insignificant $\Delta\chi^2$ indicated no significant change in model fit after additional restrictions. However, the χ^2 test is overly sensitive with large sample size in general. Given the large sample size in this study, the χ^2 test may not be the most reliable index; hence, it was not used to determine measurement invariance (Kline, 2010).

All analyses were conducted using the lavaan package (Rossel, 2012; see ESM 2 for code and output) in R (R Core Team, 2016). On average, 13.6% of data were missing at T1, 49.1% at T2, and 66.8% at T3. Little's MCAR test revealed that the data were not missing completely at random. However, we argue that the data were missing at random (MAR). Results of separate variance t-tests showed that demographic variables such as age and sex were significantly associated with missingness (see ESM 1 – Table 1). This supported a systematic pattern of missingness that could be accounted for by other variables in the data (Graham, 2012). Although diagonally weighted least squares with robust standard errors (WLSMV) are usually used as estimator for ordinal variables, we used maximum likelihood estimator (ML) to estimate parameters as ML handles missing data better. Also, ML is shown to perform well with ordinal variables with five or more categories (Beauducel & Herzberg, 2006). Full-information maximum likelihood (FIML) was used to handle missing data. A supplementary analysis was also conducted for longitudinal invariance analysis that included only participants who responded to all waves; the substantive interpretation of the results remained the same (see ESM 1 – Table 2). Furthermore, we tested all the models with WLSMV utilizing pairwise deletion; the substantive interpretation of the results also remained the same (see ESM 1 – Table 3 for detailed results).

Results

Factor Structure

The first goal of this study was to confirm the two-factor structure (a PA factor and a NA factor) of NAPAS found in Joshanloo and Bakhshi (2015) and Joshanloo (2017). We estimated a one-factor and a two-factor model using data at T1 (N = 3,487). A significant chi-square difference test ($\Delta\chi^2$ = 3,010.0, Δdf = 1, p < .001) indicated that the two-factor model (χ^2 = 2,277.1, df = 53, p < .001) fit better than the single-factor model (χ^2 = 5,287.1, df = 54, p < .001). Other model fit indexes of the two-factor model were also superior to those of the one-factor model (Table 1). There was a strong negative correlation between the two factors (r = −.72, SE = 0.014, p < .001). Our finding is consistent with Joshanloo and Bakhshi (2015) who found a two-factor structure in an Iran and US sample. Descriptive statistics of all 12 items can be found in ESM 1 – Table 4.

Table 1. Estimated parameters in the two-factor model

Items	Factor loadings NA	Factor loadings PA	Intercepts	Residual variances
1. Everything was an effort	0.70		0.69	.41
2. Hopeless	0.60		0.32	.19
3. So sad nothing can cheer you up	0.59		0.45	.23
4. Worthless	0.58		0.33	.24
5. Nervous	0.55		0.83	.47
6. Restless or fidgety	0.54		0.80	.51
1. Full of life		0.81	2.23	.41
2. Satisfied		0.76	2.39	.33
3. Extremely happy		0.74	1.91	.40
4. Calm and peaceful		0.71	2.39	.32
5. In good spirits		0.62	2.68	.13
6. Cheerful		0.60	2.59	.16

Notes. NA = Negative affect; PA = Positive affect. Fit indices for one-factor model: CFI = .776, RMSEA = .179, RMSEA 90% CI [.175, .183], SRMR = .086. Fit indices for two-factor model: CFI = .905, RMSEA = .118, RMSEA 90% CI [.114, .122], SRMR = .045.

Invariance Across Gender/Age Groups

We tested the measurement invariance of PA and NA across gender at three different time points. As shown in Table 2, all CFI changes across models were less than .01. Therefore, full scalar invariance appeared to be tenable across gender at all time points. In other words, the factor structures, loadings, and intercepts were equivalent for males and females. Similarly, the same procedures were taken to assess measurement invariance across age groups at T1. As shown in Table 2, again the results showed full scalar invariance across age groups. The results supported the notion that the items in the NAPAS scale measured the same latent construct with equivalent factor structure, factor loadings, and intercepts across the young, middle-aged, and older adults.

Table 2. Measurement invariance using Maximum likelihood

	χ^2	*df*	$\Delta\chi^2$	Δ *df*	CFI	ΔCFI	RMSEA	RMSEA 90% CI	SRMR	Invariance?
Across gender at T1[a] (*n* = 3,029)										
Model CI	2,439.2	106	–	–	.900	–	.121	.116, .125	.047	Yes
Model MI	2,454.1	116	14.83	10	.900	.000	.115	.111, .119	.049	Yes
Model SI	2,585.2	126	131.10**	10	.895	.005	.114	.110, .117	.050	Yes
Across gender at T2[b] (*n* = 1,795)										
Model CI	1,369.1	106	–	–	.901	–	.115	.110, .121	.050	Yes
Model MI	1,385.1	116	16.01	10	.900	.001	.110	.105, .116	.051	Yes
Model SI	1,434.8	126	49.73**	10	.897	.003	.108	.103, .113	.053	Yes
Across gender at T3[c] (*n* = 1,173)										
Model CI	831.28	106	–	–	.912	–	.108	.101, .115	.049	Yes
Model MI	866.60	116	35.32**	10	.909	.003	.105	.099, .112	.054	Yes
Model SI	895.31	126	28.72**	10	.907	.002	.102	.096, .108	.055	Yes
Across age groups at T1 (*n* = 3,027)										
Model CI	2,481.6	159	–	–	.900	–	.120	.116, .125	.047	Yes
Model MI	2,532.8	179	51.21**	20	.899	.001	.114	.110, .118	.051	Yes
Model SI	2,638.4	199	105.61**	20	.895	.004	.110	.106, .114	.052	Yes
Over time (*n* = 3,096)										
Model CI	6,443.3	579	–	–	.877	–	.057	.056, .058	.047	Yes
Model MI	6,466.7	599	23.38	20	.877	.000	.056	.055, .057	.048	Yes
Model SI	6,564.7	619	97.98**	20	.875	.002	.056	.054, .057	.048	Yes

Notes. [a]respondents age ranged from 20 to 75; [b]respondents age ranged from 30 to 84; [c]respondents age ranged from 39 to 93. χ^2 = Chi-square; *df* = degree of freedom; $\Delta\chi^2$ = Chi-square difference; Δ*df* = change in degree of freedom; CFI = comparative fit index; ΔCFI = CFI change; RMSEA = root mean square error of approximation; 90% CI = 90% Confidence interval of RMSEA; SRMR = Standardized root mean square residuals; Model CI = Configural invariance model; Model MI = Metric invariance model; Model SI = Scalar invariance model. ***p* < .01.

Longitudinal Measurement Invariance

To evaluate whether the same construct was measured over time, we tested longitudinal measurement invariance across three time points. Given full scalar gender invariance and full scalar invariance across age groups, male and female participants, as well as young, middle-aged, and older participants were combined for the analysis of longitudinal invariance. Invariance of PA and NA was tested over twenty years using data from T1, T2, and T3. We found full longitudinal scalar invariance for the scale suggesting that the factor structure, factor loadings, and intercepts of the items were equivalent over twenty-year time. As such, the findings supported that this scale measured the same latent constructs, PA and NA, even as the sample aged over time.

Latent Mean Differences

Given the already established scalar invariance longitudinally, as well as across gender and age groups, any latent mean differences can be interpreted as resulting from differing levels of PA and NA across groups and time and not measurement non-invariance. Latent mean differences across gender, age groups, and over time are presented in Table 3. The latent means of NA were significantly different between men and women at T1 and T2, while latent means of PA were only significantly different at T1. Males reported significantly higher PA with a trivial effect size and lower NA with a small effect size at T1 and lower NA with a small effect size at T2 than females. The latent means of NA and PA were also significantly different across all age groups except PA between young and middle-age group; the middle-aged group showed lower NA than the young group with a trivial effect size. The older group showed higher PA and lower NA than the young and middle-aged groups, all with small effect sizes. Although the *p* values suggested that the latent means of NA were different from T1 to T2 and T3, the effect sizes were extremely trivial ranging from −.05 to −.06. The latent means of PA were not significantly different over time.

Discussion

This study utilized a developmental perspective in examining the measurement invariance of NAPAS across adulthood. Invariance across age groups was tested using cross-sectional data and invariance over time was tested with longitudinal data. Our findings suggested that the NAPAS is invariant across young, middle-aged, and older

Table 3. Latent mean differences

Factor	Latent mean differences	*SE*	Cohen's *d*	*p* value
Across gender at T1[a] (female vs. male)				
PA	.108	.037	.11 [.04, .18]	.003
NA	−.205	.032	−.22 [−.29, −.15]	< .001
Across gender at T2[a] (female vs. male)				
PA	.026	.048	.03 [−.07, .12]	.586
NA	−.182	.042	−.20 [−.29,−.11]	< .001
Across gender at T3[a] (female vs. male)				
PA	.010	.060	.01 [−.10, .12]	.873
NA	−.08	.056	−.08 [−.20, .03]	.173
Across age groups at T1[b] (young vs. middle-aged group)				
PA	.032	.043	.03 [−.05, .11]	.458
NA	−.104	.043	−.10 [−.18, −.03]	.015
Across age groups at T1[b] (young vs. old group)				
PA	.287	.051	.30 [.20, .40]	< .001
NA	−.267	.048	−.30 [−.40, −.20]	< .001
Across age groups at T1 (middle-aged vs. old group)				
PA	.247	.048	.26 [.16, .36]	< .001
NA	−.166	.046	−.18 [−.28, −.09]	< .001
Over time[d] (T1 vs. T2)				
PA	.039	.023	.04 [−.01, .09]	.090
NA	−.052	.024	−.05 [−.10, 0]	.025
Over time[d] (T1 vs. T3)				
PA	.025	.028	.03 [−.02, .08]	.359
NA	−.060	.030	−.06 [−.11, −.01]	.041
Over time[e] (T2 vs. T3)				
PA	−.014	.025	−.01 [.−06, .04]	.588
NA	−.008	.026	−.01 [−.06, .04]	.771

Notes. PA = Positive affect; NA = Negative affect. [a]female as reference group; [b]the young group as referent groups; [c]the middle-aged group as referent groups; [d]T1 as referent group; [e]T2 as referent group.

adults as well as longitudinally and therefore is measuring the same construct across age/time. This finding has important implications for future developmental studies using the same scale to study age differences and longitudinal changes of PA and NA in adulthood. Establishing scalar measurement invariance allows researchers to interpret the differences in the latent means of PA and NA between young, middle-aged, and older adults, and across time points, as developmental changes but not as a measurement issue. Furthermore, our finding of full scalar invariance across gender is consistent with Joshanloo and Bakhshi (2015)'s findings despite different samples and age ranges; our findings further supported gender invariance of NAPAS. The findings of full scalar invariance across gender suggested that men and women interpreted the items in the NAPAS scale similarly and responded in a similar fashion. Hence, we could also interpret the latent mean differences between males and females as actual gender differences.

Although the scale was found to be invariant across gender, there were significant latent mean differences between men and women. We found that men reported higher PA and lower NA at T1 and lower NA at T2. This is in line with previous studies suggesting that women tend to report more negative emotions (Joshanloo & Bakhshi, 2015) or are more likely to be depressed than males (Nolen-Hoeksema, 2001). Various biological (e.g., hormonal differences) and social factors (e.g., gender socialization) could help explain the gender differences of PA and NA (Nolen-Hoeksema & Rusting, 2003). However, the latent mean differences across gender at T3 were not significant. This suggests that gender differences may reduce as we age and the effect of gender on PA and NA may vary as a function of age. Hence, developmental factors such as age differences should be taken into account when gender differences in affect are examined.

Contradictory to our hypotheses that there would be non-invariance across age groups and over time, we found full scalar invariance across young, middle-aged, and older adults, as well as over twenty years longitudinally. These findings revealed strong psychometric properties of NAPAS and indicated that individuals tend to understand and respond to the items consistently both across age groups and over time. Given the evidence of strong invariance, the NAPAS scale is a viable option for both cross-sectional research involving adults across developmental stages and longitudinal studies. Furthermore, differences of PA and NA across ages and over time measured by NAPAS can be ascribed to differences in levels of the traits.

Interestingly, we found that the latent means of PA were different across age groups but not within individuals over time. Furthermore, the effect sizes of latent mean differences of NA were bigger across age groups than within individuals over time. The significant mean differences across age groups supported that there are developmental differences in PA and NA. Although the latent mean differences of NA were also significant over time from T1 to T2 and T3, the effect sizes suggested that the differences were extremely trivial and unlikely to have practical implications; in fact, the latent means of NA were not significantly different across T2 and T3. Older-aged individuals having the highest PA and lowest NA was unsurprising as past studies generally suggested that older people are better at emotion regulation and focus more on interpersonal relations; hence, they have lower NA. Our findings further support the idea that PA and NA may change in older adulthood (Charles et al., 2001).

In contrast to the significant latent mean differences across age groups there were no substantive longitudinal differences found as all effect sizes were extremely trivial even when the latent mean differences were significant. This suggests that an individual may have relatively consistent levels of PA and NA but different age groups have an overall difference in PA and NA. The finding of intraindividual stability is consistent with past longitudinal studies of PA and NA (Charles et al., 2001). Another interesting aspect of these results is the seeming contradiction between consistent levels of PA and NA in the same individuals over time but differences across age groups. This implies that there may be a small amount of change in PA and NA in individuals longitudinally, that does not rise to the level of a significant difference, but is sufficient to result in average group differences by age. This further implies that there is a general trend of change in PA and NA that most individuals follow as the only way that there could be average group differences but longitudinal stability in the same construct is through small changes over a large number of individuals. Alternatively, it is possible that a small minority of individuals experience a great deal of change but most individuals experience little to no change. These seemingly contrasting results between the lack of change longitudinally and mean differences across age groups in PA and NA could suggest a new avenue for future research.

The current study took a developmental perspective to explore the psychometric properties of NAPAS and the latent mean differences of PA and NA across adulthood. The results support strong psychometric properties of the scale and reveal highly similar properties when assessing across gender, age groups, and over time. Future large-scale longitudinal studies that aim at examining PA and NA with a short scale could consider the scale evaluated in this study as a potential option.

Electronic Supplementary Materials

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/1015-5759/a000529

ESM 1. Figure 1: Path diagram for longitudinal CFA. Table 1: Separate variance *t*-test. Table 2: Longitudinal invariance analysis using ML without estimating missing data. Table 3: Output for WLSMV. Table 4: Descriptive statistics (.docx).

ESM 2. Combination of the dataset of MIDUS 1, 2, and 3, with only the random-digit-dialling subsample (.pdf).

References

Augustus-Horvath, C. L., & Tylka, T. L. (2011). The acceptance model of intuitive eating: A comparison of women in emerging adulthood, early adulthood, and middle adulthood. *Journal of Counseling Psychology, 58*, 110–125. https://doi.org/10.1037/a0022129

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13*, 186–203. https://doi.org/10.1207/s15328007sem1302_2

Bentler, P. M., & Hu, L. T. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.

Brim, O. G., Ryff, C. D., & Kessler, R. C. (Eds.). (2004). *How healthy are we? A national study of well-being at midlife*. Chicago, IL: University of Chicago Press.

Brown, T. (2006). *Confirmatory factor analysis for applied research* (1st ed.). New York, NY: The Guilford Press.

Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of Personality and Social Psychology, 80*, 136–151. https://doi.org/10.1037/0022-3514.80.1.136

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25*, 1–27. https://doi.org/10.1016/S0149-2063(99)80001-4

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*, 245–265. https://doi.org/10.1348/0144665031752934

Diener, E., & Chan, M. Y. (2011). Happy people live longer: Subjective well-being contributes to health and longevity.

Applied Psychology: Health and Well-Being, 3, 1–43. https://doi.org/10.1111/j.1758–0854.2010.01045.x

Diener, E., Nickerson, C., Lucas, R. E., & Sandvik, E. (2002). Dispositional affect and job outcomes. *Social Indicators Research, 59*, 229–259. https://doi.org/10.1023/A:1019672513984

Graham, J. W. (2012). Multiple imputation with norm 2.03. In *Missing data* (pp. 73–94). New York, NY: Springer.

Joshanloo, M. (2017). Factor structure and criterion validity of original and short versions of the Negative and Positive Affect Scale (NAPAS). *Personality and Individual Differences, 105*, 233–237. https://doi.org/10.1016/j.paid.2016.09.060

Joshanloo, M., & Bakhshi, A. (2015). The factor structure and measurement invariance of positive and negative affect. *European Journal of Psychological Assessment, 32*, 265–272. https://doi.org/10.1027/1015-5759/a000252

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131*, 803–855. https://doi.org/10.1037/0033-2909.131.6.803

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the positive and negative affect schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences, 27*, 405–416. https://doi.org/10.1016/S0191-8869(98)00251-7

Merz, E. L., Malcarne, V. L., Roesch, S. C., Ko, C. M., Emerson, M., Roma, V. G., & Sadler, G. R. (2013). Psychometric properties of Positive and Negative Affect Schedule (PANAS) original and short forms in an African American community sample. *Journal of Affective Disorders, 151*, 942–949. https://doi.org/10.1016/j.jad.2013.08.011

Mroczek, D. K. (2004). Positive and negative affect at midlife. In C. D. Ryff & O. G. Brim (Eds.), *How healthy are we? A national study of well being at midlife* (pp 205–226). Chicago, IL: University of Chicago Press.

Mroczek, D. K., & Kolarz, C. M. (1998). The effect of age on positive and negative affect: A developmental perspective on happiness. *Journal of Personality and Social Psychology, 75*, 1333–1349. https://doi.org/10.1037/0022-3514.75.5.1333

Nolen-Hoeksema, S. (2001). Gender differences in depression. *Current Directions in Psychological Science, 10*, 173–176. https://doi.org/10.1111/1467-8721.00142

Nolen-Hoeksema, S., & Rusting, C. L. (2003). Gender differences in well-being. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 330–350). New York, NY: Russell Sage Foundation.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rossel, Y. (2012). *Lavaan: An R package for structural equation modeling and more* Version 0.5–12 (BETA). Retrieved from https//users.ugent.be/yrosseel/lavaan/lavaanIntroduction.pdf

Santrock, J. W. (2008). *Life-span development*. New York, NY: McGrawHill.

Staw, B. M., Sutton, R. I., & Pelled, L. H. (1994). Employee positive emotion and favorable outcomes at the workplace. *Organization Science, 5*, 51–71. https://doi.org/10.1287/orsc.5.1.51

Thoits, P. A., & Hewitt, L. N. (2001). Volunteer work and well-being. *Journal of Health and Social Behavior, 42*, 115–131. https://doi.org/10.2307/3090173

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

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*, 10–18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

History

Received June 23, 2018
Revision received December 2, 2018
Accepted December 5, 2018
Published online July 2, 2019
EJPA Section/Category Methodological topics in assessment

Funding

The MIDUS 1 study (Midlife in the US) was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development. The MIDUS 2 and MIDUS 3 research was supported by grants from the National Institute on Aging (P01AG020166, R37AG027343).

ORCID

Meingold Hiu-ming Chan
https://orcid.org/0000-0001-9108-2504

Meingold Hiu-ming Chan
Department of Human Sciences
Ohio State University
Columbus, OH 43210
USA
chan.742@osu.edu

