Personality and Individual Differences 217 (2024) 112467

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Assessing the temporal stability of a measure of trait emotional intelligence: Systematic review and empirical analysis

Bogdan S. Zadorozhny [a,*], K.V. Petrides [a], Joran Jongerling [b], Stephen Cuppello [c], Dimitri van der Linden [d]

[a] Department of Psychology, University College London (UCL), United Kingdom
[b] Department of Methodology and Statistics, Tilburg School of Social and Behavioral Sciences, the Netherlands
[c] Thomas International, Marlow, United Kingdom
[d] Department of Psychology, Erasmus University Rotterdam, the Netherlands

ARTICLE INFO

Keywords:
Emotional intelligence
Trait EI
TEIQue
Test-retest reliability
Temporal stability
Validity

ABSTRACT

Determining the temporal stability of a construct is crucial to establishing its validity and utility in real-world scenarios. To-date, few studies have investigated the test-retest reliability of trait emotional intelligence (trait EI), particularly over extended periods of time. The present study presents relevant data from the Trait Emotional Intelligence Questionnaire (TEIQue) over variable intervals, ranging from 30 days (one month) to 1444 days (≈ four years). Results provide support for trait EI theory, demonstrating strong temporal stability at all levels of the construct (global, factor, and facet). Future research may focus on extending the test-retest intervals at both ends (i.e., below one month and over four years) as well as on comparisons between different trait EI measures.

1. Introduction

Trait emotional intelligence (trait EI) is defined as a constellation of emotion-related self-perceptions that describes how an individual assesses their own emotional and social effectiveness (Petrides, Pita, et al., 2007). Although several instruments have been developed that purport to measure trait EI, amongst these measures, the Trait Emotional Intelligence Questionnaire (TEIQue) is the only inventory that directly and comprehensively operationalizes the trait EI theory (Austin et al., 2008; Petrides et al., 2016).

With respect to the various indicators of reliability and validity, test-retest reliability is of particular importance (McCrae, 2015; Oostrom et al., 2019). Nevertheless, a large number of trait studies fail to report test-retest reliability, despite the fact that "*only test–retest reliability is necessarily relevant to studies of longitudinal stability or change*" (McCrae et al., 2011, p. 29). In other words, for accurate measure of personality, a test must perform consistently over time (Dave et al., 2021; Davies et al., 2010). Additionally, test-retest reliability provides an estimate of the maximum strength with which a measure can correlate with other variables, thereby directly impacting its construct validity (Assaad et al., 2022).

Demonstrating test-retest reliability is particularly important for trait EI as the theory conceptualizes EI as a stable personality construct (Petrides, Furnham, et al., 2007). Notably, there are presently few measures of EI that have been demonstrated to be test-retest reliable, internally consistent, and valid in terms of robust factor structure and predictive ability (Davis & Wigelsworth, 2018).

To-date, there have been few investigations into the test-retest reliability of TEIQue (Perazzo et al., 2021), most of which failed to truly assess its temporal stability due to their utilization of insufficient test-retest intervals (Costa & McCrae, 1998; Wood et al., 2022). It has been argued that any period of less than a year can be considered short-term in the context of assessing the stability of personality traits (Murray et al., 2003; Schuerger et al., 1989). Therefore, utilizing shorter test-retest intervals prevents investigation of longitudinal change while also potentially exposing studies to confounds such as memory effects (Sovet et al., 2014).

In addition to the overall scarcity of studies investigating the temporal stability of personality and EI constructs, the majority thereof have used contrived experimental designs within laboratory settings, which raises concerns surrounding the transferability of their results to real-world applications, such as the career context (Lievens et al., 2005). Prior studies were also predominantly performed in cohort paradigms wherein groups were tested en masse at two times, the results of which

* Corresponding author at: University College London, Gower Street, London WC1E 6BT, United Kingdom.
E-mail address: bogdan.zadorozhny.21@ucl.ac.uk (B.S. Zadorozhny).

https://doi.org/10.1016/j.paid.2023.112467
Received 11 April 2023; Received in revised form 10 October 2023; Accepted 30 October 2023
Available online 3 November 2023

can notably only demonstrate temporal stability over the selected time interval. Additionally, whereas personality had been found to be reasonably consistent in adulthood, results concerning adolescents and early adults have been less conclusive (Dave et al., 2021). Therefore, it is plausible that demographic differences between study and target populations may lead to erroneous conclusions, if response patterns of college students and older adults diverge.

Regarding the statistical methods utilized to examine test-retest reliability, most past studies used the Pearson correlation coefficient (*r*) which has been argued to be unsuitable due to its inherent insensitivity to systematic error, it being unaffected by changes in means and standard deviations, and its inability to analyze more than two scores per participant (Baumgartner, 2000; Russo & Lathan, 2015). Consequently, both Baumgartner (2000) as well as Russo and Lathan (2015) instead recommend the use of Intraclass Correlations (ICCs), but as Koo and Li (2015) cautioned, it is then vital to carefully report the form of ICC utilized.

1.1. The present study

Given the importance of establishing test-retest reliability in the context of trait EI theory and the scarcity of relevant extant literature, particularly covering longer temporal periods, with the exception of Parker et al. (2021), the current study was designed to examine the temporal stability of trait EI, while addressing the aforementioned pitfalls. All study participants were active in the workforce, were tested separately using test-retest intervals that were unique to each individual, and completed the assessment as part of a career recruitment or development task. In order to enable comparison with existing literature, the dataset was clustered into groups based on test-retest durations utilized in past studies (Wood et al., 2022).

The first goal of the present study was to examine the temporal stability of trait EI at the global and factor levels. It was hypothesized that the correlations within temporal clusters would be moderate-to-strong (Petrides, 2009; Shahzad et al., 2014; Siegling et al., 2015; Stassart et al., 2019). In addition, an exploratory goal of the study was to investigate ICC patterns across the clusters. It was predicted that no significant moderating effects of test-retest interval durations on the difference of TEIQue profile scores as evaluated at test and retest would be obtained. It was also expected that at least configural and metric model invariance would be obtained across test administrations, thereby supporting the temporal stability of the construct by indicating that its factor structure and intercepts are equal across time.

2. Materials and methods

2.1. Participants

The sample comprised 1490 respondents (35 % male) who had completed the TEIQue twice as part of occupation-related activities. Therefore, the present data were obtained in genuine high stakes testing with real-life outcomes. All participants were based in the United Kingdom and completed the instrument in English.

2.2. Procedure

Participants were recruited from several companies across a variety of industries and were selected to undergo testing for genuine occupational purposes such as candidate recruitment and internal personnel development. Once participants completed the first test, they were sent follow-up email invitations after a variable duration of time had elapsed to again undergo the TEIQue assessment. Participants completed the test twice for various reasons, including annual corporate performance and development appraisals, applying to multiple jobs that required undergoing the assessment, and internal promotion procedures that mandated recompleting the test.

In all cases, assessments were administered electronically, and participants were provided with instructions and information prior to initiation. Individuals were informed that in case of any questions, they ought to contact the requisitioning individual.

2.3. Measures

The Trait Emotional Intelligence Questionnaire (TEIQue) (Petrides, Pita, et al., 2007) is an inventory composed of 153 items rated on a 7-point Likert scale, ranging from 1 (completely disagree) to 7 (strongly agree). It comprises 15 facets at its lowest interpretable level, of which 13 cluster into 4 intercorrelated intermediate-level factors (Well-being, Self-control, Emotionality, and Sociability), with two facets (viz., Adaptability and Self-motivation) feeding directly into the global trait EI score located at the apex of this hierarchy.

2.4. Data cleaning and banding

The data were cleaned prior to analysis, with a total of 647 cases removed for various valid reasons, resulting in a final sample size of $n = 843$ (see supplementary material for full details). Since the temporal delay intervals between administrations were unique for each participant, ranging from 30 to 1444 days ($\approx$ 4 years) in the cleaned dataset, it was decided to band the data prior to analysis.

An extensive review of prior studies reporting the test-retest reliability of various personality related constructs was conducted. The results of this review, which included 160 unique effect sizes, described in 108 articles, are presented in supplementary Table S1.

The present dataset was banded based on the typical test-retest intervals utilized in past studies and considering the data available. The 8 clusters that emerged as a result are summarized in Table 1.

2.5. Data analysis plan

Statistical analyses were performed using SPSS v.28, Stata v.17, and Mplus v.8.8. Measures of internal consistency (Cronbach's alpha) were reported and were used to obtain attenuated and disattenuated test-retest reliabilities.

ICC estimates and their 95 % confidence intervals were calculated based on a mean-rating (k = 2), absolute-agreement, 2-way mixed-effects model [ICC(A,k)], in accordance with recommendations in Koo and Li (2015) and Qin et al. (2019). The guidelines for interpreting the reliabilities were adopted from Koo and Li (2015), wherein values <0.50, between 0.50 and 0.75, between 0.75 and 0.90, and > 0.90 were considered to be indicative of poor, moderate, good, and excellent reliability, respectively.

Additionally, potential interaction effects of retest intervals on TEIQue profile scores were investigated using the PROCESS plugin v.4.1 for SPSS, Model 1.

Next, prior to initiating multigroup Structural Equation Modeling (SEM), four tests of measurement invariance (configural, metric, scalar, and residual) were carried out to confirm the equivalence of the factor structure and regression coefficients between the test and retest administration groups. Said tests were carried out in accordance with

Table 1
List of clusters utilized in the present study.

Cluster No.	N	Duration Between Tests	Descriptive Name
1	69	30–90 days	1–3 months
2	89	91–180 days	3–6 months
3	92	181–270 days	6–9 months
4	114	271–365 days	9–12 months
5	122	365–540 days	12–18 months
6	157	541–730 days	18–24 months
7	144	731–1095 days	2–3 years
8	56	1096–1460 days	3–4 years

Keefer et al. (2013) and Perazzo et al. (2021), through a series of Confirmatory Factor Analyses utilizing Stata.

Finally, SEM (with Maximum Likelihood estimation) was used to test a correlational and a hierarchical model. The former consisted of a test and retest group, each containing four intercorrelated factors. The latter was composed of the same two groups containing the same four factors in addition to a latent apex factor representing global trait EI. Model goodness-of-fit was assessed by the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA) and its 90 % confidence intervals, and the Standardized Root Mean Square Residual (SRMR). Fit thresholds were selected from Hoyle (1995) and were as follows: $CFI > 0.90$, $TLI > 0.90$, $RMSEA < 0.08$, and $SRMR < 0.08$ indicating good fit. Additionally, the chi-square (χ^2), degrees of freedom (df), and p-value statistics were reported.

3. Results

3.1. Clustered Pearson correlations and basic descriptives

The descriptive statistics, reliabilities, and attenuated and disattenuated Pearson correlations for the global trait EI score and its four factors across clusters have been included in Table A1 and Tables S2-S5 respectively. The data for the 15 facets are included in Tables S6-S20. Cronbach's α values exceeded 0.87 for global trait EI in all clusters for both administrations, but were somewhat lower at the factor level, ranging between 0.72 and 0.89. Significant moderate-to-strong test-retest (Pearson) correlations were obtained within all clusters, ranging from $r = 0.62$ to $r = 0.77$ ($p < .001$) for global trait EI and between $r = 0.59$ to $r = 0.82$ ($p < .01$) for the factors. The facet-level results ranged more greatly which was expected considering the high number of facets relative to the clustered N sizes. The cluster-aggregated test-retest reliabilities were also moderate-to-strong, $r = 0.68$ ($p < .001$) for global trait EI, $r = 0.70$ ($p < .001$) for Well-being, $r = 0.72$ ($p < .001$) for Self-control, $r = 0.65$ ($p < .001$) for Emotionality, and $r = 0.77$ ($p < .001$) for Sociability.

3.2. Intraclass correlations

Table A2 reports the intraclass correlations (ICCs) that were obtained for global trait EI and its factors. The facet-level ICC results are included in Table S21. For global trait EI and Emotionality, the 95 % confidence intervals (CIs) of the ICC estimates indicated good to excellent reliability (between 0.78 and 0.99 and between 0.79 and 0.99, respectively). For Well-being and Sociability, the CIs indicated moderate to excellent reliability (0.66 to 0.99 and 0.74 to 0.99, respectively). However, the Self-control factor results varied to a greater extent, ranging between 0.43 and 0.99. The results at the facet level were inconsistent, with reliability ranging from poor to excellent, depending on the specific facet.

3.3. Moderator effect analysis

An analysis investigating whether overall change in the trait EI variables was moderated by the duration of time elapsed between administrations was conducted. The TEIQue profile at initial test was modeled as the independent variable, the profile at retest as the outcome, and the delay interval as the moderator. For global trait EI, no significant interaction effect was obtained, $F(1, 839) = 0.002$, $p = .96$, and no significant interaction effects were obtained on either the factor or the facet levels.

3.4. Measurement invariance

Further investigation into the measurement invariance of the hierarchical trait EI model – with a latent global EI loading on the four EI factors – was conducted using a CFA-based comparison across test and retest administrations. The resultant baseline model fit demonstrated configural measurement invariance, $\chi^2(4) = 2.67$, $p = .62$; $CFI = 0.99$; $TLI = 0.99$; $RMSEA < 0.001$, 90 % CI [0.000, 0.043]; $SRMR = 0.007$, which implied the same factorial structure held across test and retest administrations. Metric-level measurement invariance was established, implying equal factor loadings across test iterations, $\chi^2(7) = 4.74$, $p = .69$; $CFI = 0.99$; $TLI = 0.99$; $RMSEA < 0.001$, 90 % CI [0.000, 0.033]; $SRMR = 0.016$; $\Delta\chi^2(3) = 2.07$, $p = .55$.

Next, scalar invariance was demonstrated, $\chi^2(10) = 10.35$, $p = .41$; $CFI = 0.999$; $TLI = 0.999$; $RMSEA = 0.006$, 90 % CI [0.000, 0.038]; $SRMR = 0.016$; $\Delta\chi^2(3) = 5.61$, $p = .13$, implying equivalence of factor loadings and intercepts. Finally, residual-level invariance did not obtain, $\chi^2(14) = 21.36$, $p = .09$; $CFI = 0.99$; $TLI = 0.99$; $RMSEA = 0.025$, 90 % CI [0.005, 0.045]; $SRMR = 0.022$; $\Delta\chi^2(4) = 11.01$, $p = .03$, which implied residual variances were not equivalent across test administrations.

3.5. Structural equation modeling

The hierarchical model was tested on test-retest associations (see Fig. 1). This model included correlations between the four component factor residuals to account for correlations endemic to each factor that were not included within the apex factor. The fit indices of this model indicated good fit, $\chi^2(15) = 32.29$, $p < .01$; $CFI = 0.99$; $TLI = 0.99$; $RMSEA = 0.037$, 90 % CI [0.019, 0.055]; $SRMR = 0.041$.

The correlations of the factors across test attempts were moderate-to-strong, ranging between 0.67 and 0.80 ($p < .001$). In addition, the loadings of global trait EI onto its component factors were relatively consistent between test and retest. A correlational table is included in Table 2.

4. Discussion

Most studies examining the test-retest reliability of the TEIQue have utilized intervals of less than a month (Perazzo et al., 2021; Shahzad et al., 2014; Stassart et al., 2019). This therefore presents a gap in knowledge regarding the temporal stability of trait EI. The present study addressed this gap by investigating its temporal stability over a greater range of time periods than prior investigations. Moreover, the utilized participant-unique interval paradigm proposes a novel way to move beyond the non-discerning use of fixed test-retest intervals.

4.1. Temporal stability of the trait EI construct

As hypothesized, global trait EI ($r = 0.62$ to $r = 0.77$) and its four factors ($r = 0.59$ to $r = 0.82$) showed moderate-to-strong levels of rank-order stability across all temporal clusters. By definition, disattenuated correlations were higher for both global trait EI ($r = 0.70$ to $r = 0.85$) and its factors ($r = 0.73$ to $r = 1.00$). These results align with prior studies on the temporal stability of EI and personality traits (Petrides, 2009; Robins et al., 2001). Although zero-order Pearson correlations underestimate true relationships, disattenuated correlations should be interpreted cautiously according to Chmielewski and Watson (2009), and thus our interpretation focuses on the former.

Despite the significantly greater temporal intervals between administrations utilized in the present study, the obtained correlations were comparable to or even stronger than those found in prior studies (see Table S1), which supports the major theoretical premise of trait EI theory. A prior investigation of rank-order stability (Robins et al., 2001) reported that test-retest reliability estimates for personality measures can be expected to range from about $r = 0.50$ to $r = 0.59$ over about four years, while the presently obtained correlation for global trait EI for the same time period (Cluster 8), was significantly higher, $r = 0.77$.

The obtained Cronbach's alphas within clusters, per Elfenbein et al. (2017), indicated a high level of internal consistency for global trait EI ($\alpha > 0.80$), above the cut-off of 0.75 generally accepted for instruments in the health sciences (Aritzeta et al., 2016). There was greater variance

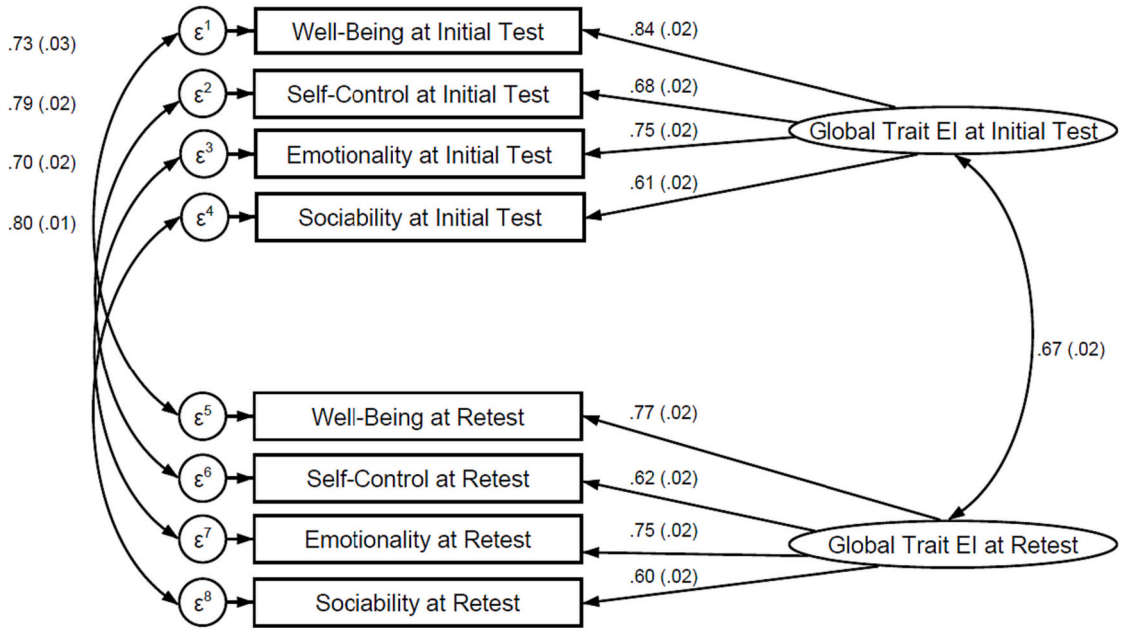

Fig. 1. Hierarchical model, correlations, and regression coefficients.

Table 2
Standardized correlations for SEM model variables.

Measure	Test-Retest Correlation
Well-being	0.73*
Self-control	0.79*
Emotionality	0.70*
Sociability	0.80*
Global trait EI	0.67*

* $p < .01$.

at the factor level, with alphas ranging between 0.61 and 0.94, although averages exceeded 0.75 in all cases.

These results demonstrate the temporal stability of trait EI at the global, factor, and facet levels for variable test-retest periods ranging from one month up to four years, lending further support to the notion of trait EI constituting a stable personality dimension. This was further substantiated by the Intraclass Correlation results which indicated good to excellent reliability for global trait EI and Emotionality as well as moderate to excellent reliability for Well-being and Sociability. The results were less consistent for Self-control, although the overall average ICC obtained for said factor was nevertheless indicative of good reliability at ICC = 0.83, $p < .001$.

With that said, it ought to be noted that the stability of scores on a measure does not necessarily imply that the underlying trait being measured is itself stable, due to a host of parameters that may cause the two to diverge, including measurement error, bias, and generalizability issues.

4.2. Impact of variable test-retest intervals

No significant moderating effect of time intervals on test-retest score deltas was found on the global, factor, and facet levels, indicating that the duration elapsed between administrations had no significant impact on the difference in trait EI profile scores, in accordance with hypothesis.

Thus, the difference in scores between administrations is attributable to other factors, such as inconsistent or random fluctuations within response patterns. It could also be due to mean-level change, wherein average global trait EI scores may change over time within an entire population (Keefer et al., 2013). This could be caused by genuine personal development or by factors not reflective of real personality change, such as changing societal values and economic trends.

4.3. Investigation of measurement invariance and factor-level stability

In the present study, configural, metric, and scalar invariance were observed, which exceeded expectations. This demonstrated that when retested following different intervals, the TEIQue maintains its factor structure across time-delayed administrations, with equivalent loadings and strong measurement equivalence.

The hierarchical SEM model was selected for analysis due to its superior explanatory power and the obtained result suggested a reasonably high test-retest consistency ($r = 0.67$ to $r = 0.80$), which is impressive considering the significant within-cluster variability as well as the range and length of the test-retest intervals. Additionally, the comparison of regression coefficients across test attempts afforded further convergent evidence for the temporal stability of the TEIQue at the factorial level.

5. Conclusion

The results of our investigation corroborate the substantial stability of trait EI scores over varying time periods, including over long-term and divergent intervals. This is consistent with the results of prior studies that investigated the affective aspects of personality and consistently found high levels of trait stability throughout adulthood (Allemand et al., 2013). Thusly, our results provide major and direct support for the leading operationalization of emotional intelligence as a personality trait and thence for its continuing measurement through questionnaires and rating scales.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Informed consent statement

Informed consent was obtained from all subjects involved in the study.

CRediT authorship contribution statement

Bogdan S. Zadorozhny: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Visualization. **K.V. Petrides:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision. **Joran Jongerling:** Methodology, Formal analysis. **Stephen Cuppello:** Investigation, Data curation. **Dimitri van der Linden:** Methodology, Validation, Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Study materials are available, free of charge for scientific research purposes only, from www.psychometriclab.com. The dataset is proprietary and not publicly available.

Appendix A

Table A1
Test-retest correlations and descriptive statistics for global trait EI.

Cluster No.	N	M	SD	Range	Alpha	r	p	Disattenuated r
1								
Test	69	4.93	0.66	3.33–6.44	0.92	–	–	–
Retest	69	5.26	0.63	3.50–6.52	0.91	**0.66**	<0.001	0.73
2								
Test	89	5.10	0.64	3.30–6.36	0.91	–	–	–
Retest	89	5.25	0.61	2.77–6.35	0.91	**0.64**	<0.001	0.70
3								
Test	92	5.08	0.63	3.51–6.65	0.90	–	–	–
Retest	92	5.23	0.60	3.51–6.41	0.90	**0.66**	<0.001	0.73
4								
Test	114	5.27	0.61	3.59–6.33	0.92	–	–	–
Retest	114	5.32	0.62	3.43–6.41	0.91	**0.75**	<0.001	0.82
5								
Test	122	5.30	0.60	3.29–6.43	0.90	–	–	–
Retest	122	5.41	0.54	3.33–6.49	0.89	**0.69**	<0.001	0.77
6								
Test	157	5.27	0.54	3.85–6.71	0.89	–	–	–
Retest	157	5.39	0.52	2.90–6.68	0.87	**0.62**	<0.001	0.70
7								
Test	144	5.31	0.56	3.29–6.47	0.90	–	–	–
Retest	144	5.38	0.54	3.88–6.54	0.89	**0.64**	<0.001	0.72
8								
Test	56	5.13	0.67	3.39–6.41	0.93	–	–	–
Retest	56	5.26	0.53	4.14–6.60	0.88	**0.77**	<0.001	0.85

Note. $n = 843$.

Table A2
Intraclass correlations for trait EI measures across administrations.

Measure	Intraclass Correlation	95 % CI		*F* test with True Value 0			
		Lower Bound	*Upper Bound*	*Value*	*df1*	*df2*	*Sig*
Global trait EI							
Individual	0.022	0.004	0.956	59.06	1	842	<0.001
Average	0.950	0.778	0.999				
Well-being							
Individual	0.013	0.002	0.932	38.26	1	842	<0.001
Average	0.917	0.661	0.999				
Self-control							
Individual	0.006	0.001	0.864	18.94	1	842	<0.001
Average	0.833	0.434	0.999				
Emotionality							
Individual	0.023	0.004	0.961	58.26	1	842	<0.001
Average	0.953	0.788	0.999				
Sociability							
Individual	0.017	0.003	0.948	65.44	1	842	<0.001
Average	0.937	0.736	0.999				

Note. $n = 843$.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.paid.2023.112467.

References

Allemand, M., Steiger, A. E., & Hill, P. L. (2013). Stability of personality traits in adulthood: Mechanisms and implications. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry, 26*(1), 5–13. https://doi.org/10.1024/1662-9647/a000080

Aritzeta, A., Balluerka, N., Gorostiaga, A., Alonso-Arbiol, I., Haranburu, M., & Gartzia, L. (2016). Classroom emotional intelligence and its relationship with school performance. *European Journal of Education and Psychology, 9*(1), 1–8. https://doi.org/10.1016/j.ejeps.2015.11.001

Assaad, L., Griffin, S. A., Derefinko, K. J., Lynam, D. R., & Samuel, D. B. (2022). Twelve-month stability of a brief FFM measure and validity of its impulsigenic facets. *Journal of Psychopathology and Behavioral Assessment, 44*(1), 115–124. https://doi.org/10.1007/s10862-021-09933-9

Austin, E. J., Parker, J. D. A., Petrides, K. V., & Saklofske, D. H. (2008). Emotional intelligence. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment* (pp. 576–596). SAGE Publications.

Baumgartner, T. A. (2000). Estimating the stability reliability of a score. *Measurement in Physical Education and Exercise Science, 4*(3), 175–178. https://doi.org/10.1207/s15327841mpee0403_3

Chmielewski, M., & Watson, D. (2009). What is being assessed and why it matters: The impact of transient error on trait research. *Journal of Personality and Social Psychology, 97*(1), 186–202. https://doi.org/10.1037/a0015618

Costa, P. T., & McCrae, R. R. (1998). Trait theories of personality. In D. F. Barone, M. Hersen, & V. B. Van Hasselt (Eds.), *Advanced personality* (pp. 103–121). Springer US.

Dave, H. P., Keefer, K. V., Snetsinger, S. W., Holden, R. R., & Parker, J. D. A. (2021). Stability and change in trait emotional intelligence in emerging adulthood: A four-year population-based study. *Journal of Personality Assessment, 103*(1), 57–66. https://doi.org/10.1080/00223891.2019.1693386

Davies, K. A., Lane, A. M., Devonport, T. J., & Scott, J. A. (2010). Validity and reliability of a brief emotional intelligence scale (BEIS-10). *Journal of Individual Differences, 31*(4), 198–208. https://doi.org/10.1027/1614-0001/a000028

Davis, S. K., & Wigelsworth, M. (2018). Structural and predictive properties of the emotional quotient inventory youth version-short form (EQ-i:YV S). *Journal of Personality Assessment, 100*(2), 197–206. https://doi.org/10.1080/00223891.2017.1280502

Elfenbein, H. A., Jang, D., Sharma, S., & Sanchez-Burks, J. (2017). Validating emotional attention regulation as a component of emotional intelligence: A Stroop approach to individual differences in tuning in to and out of nonverbal cues. *Emotion, 17*(2), 348–358. https://doi.org/10.1037/emo0000145

Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 1–15). Sage Publications.

Keefer, K. V., Holden, R. R., & Parker, J. D. A. (2013). Longitudinal assessment of trait emotional intelligence: Measurement invariance and construct continuity from late childhood to adolescence. *Psychological Assessment, 25*(4), 1255–1272. https://doi.org/10.1037/a0033903

Koo, T. K., & Li, M. Y. (2015). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lievens, F., Buyse, T., & Sackett, P. R. (2005). Retest effects in operational selection settings: Development and test of a framework. *Personnel Psychology, 58*(4), 981–1007. https://doi.org/10.1111/j.1744-6570.2005.00713.x

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review, 19*(2), 97–112. https://doi.org/10.1177/1088868314541857

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15*(1), 28–50. https://doi.org/10.1177/1088868310366253

Murray, G., Rawlings, D., Allen, N. B., & Trinder, J. (2003). NEO five-factor inventory scores: Psychometric properties in a community sample. *Measurement and Evaluation in Counseling and Development, 36*(3), 140–149. https://doi.org/10.1080/07481756.2003.11909738

Oostrom, J. K., De Vries, R. E., & De Wit, M. (2019). Development and validation of a HEXACO situational judgment test. *Human Performance, 32*(1), 1–29. https://doi.org/10.1080/08959285.2018.1539856

Parker, J. D. A., Summerfeldt, L. J., Walmsley, C., O'Byrne, R., Dave, H. P., & Crane, A. G. (2021). Trait emotional intelligence and interpersonal relationships: Results from a 15-year longitudinal study. *Personality and Individual Differences, 169*, Article 110013. https://doi.org/10.1016/j.paid.2020.110013

Perazzo, M. F., Abreu, L. G., Pérez-Díaz, P. A., Petrides, K. V., Granville-Garcia, A. F., & Paiva, S. M. (2021). Trait emotional intelligence questionnaire-short form: Brazilian validation and measurement invariance between the United Kingdom and Latin-American datasets. *Journal of Personality Assessment, 103*(3), 342–351. https://doi.org/10.1080/00223891.2020.1758118

Petrides, K. V. (2009). Psychometric properties of the trait emotional intelligence questionnaire (TEIQue). In C. Stough, D. H. Saklofske, & J. D. A. Parker (Eds.), *Assessing emotional intelligence* (pp. 85–101). Springer. https://doi.org/10.1007/978-0-387-88370-0_5.

Petrides, K. V., Furnham, A., & Mavroveli, S. (2007). Trait emotional intelligence: Moving forward in the field of EI. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns* (pp. 151–166). Oxford University Press.

Petrides, K. V., Mikolajczak, M., Mavroveli, S., Sánchez-Ruiz, M.-J., Furnham, A., & Pérez-González, J.-C. (2016). Developments in trait emotional intelligence research. *Emotion Review, 8*(4), 335–341. https://doi.org/10.1177/1754073916650493

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology, 98*, 273–289.

Qin, S. S., Nelson, L., McLeod, L., Eremenco, S., & Coons, S. J. (2019). Assessing test-retest reliability of patient-reported outcome measures using intraclass correlation coefficients: Recommendations for selecting and documenting the analytical formula. *Quality of Life Research, 28*(4), 1029–1033. https://doi.org/10.1007/s11136-018-2076-0

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69*(4), 617–640. https://doi.org/10.1111/1467-6494.694157

Russo, C. R., & Lathan, C. E. (2015). An evaluation of the consistency and reliability of the defense automated neurocognitive assessment tool. *Applied Psychological Measurement, 39*(7), 566–572. https://doi.org/10.1177/0146621615577361

Schuerger, J. M., Zarrella, K. L., & Hotz, A. S. (1989). Factors that influence the temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 56*(5), 777–783. https://doi.org/10.1037/0022-3514.56.5.777

Shahzad, S., Riaz, Z., Begum, N., & Khanum, S. J. (2014). Urdu translation and psychometric properties of trait emotional intelligence questionnaire short form (TEIQue-SF). *Asian Journal of Management Sciences & Education, 3*(1), 1–11.

Siegling, A. B., Saklofske, D. H., & Petrides, K. V. (2015). Measures of ability and trait emotional intelligence. In *Measures of personality and social psychological constructs* (pp. 381–414). Elsevier.

Sovet, L., Park, M. S.-A., & Jung, S. (2014). Validation and psychometric properties of academic major satisfaction scale among Korean college students. *Social Indicators Research, 119*(2), 1121–1131. https://doi.org/10.1007/s11205-013-0537-y

Stassart, C., Etienne, A.-M., Luminet, O., Kaïdi, I., & Lahaye, M. (2019). The psychometric properties of the French version of the trait emotional intelligence questionnaire–child short form. *Journal of Psychoeducational Assessment, 37*(3), 293–306. https://doi.org/10.1177/0734282917740499

Wood, J. K., Anglim, J., & Horwood, S. (2022). A less evaluative measure of Big Five personality: Comparison of structure and criterion validity. *European Journal of Personality, 36*(5), 809–824. https://doi.org/10.1177/08902070211012920