Psychometric validation of the polish version of the emotional style questionnaire

Agata Gasiorowska[1] · Kaja Glomb[2] · Pelin Kesebir[3]

Abstract

In three studies, we document the validity of the Polish version of the Emotional Style Questionnaire (ESQ)—a 24-item self-report measure that captures how people vary across six dimensions that make up a healthy emotional life. Study 1 provides support for the proposed factorial structure of the Polish version of the scale and gender measurement invariance, demonstrates the similarity of the Polish and English versions of the scale, and reveals the excellent test-retest reliability of the scale across an interval of two months. Study 2 confirms the adequacy of the factorial structure in another sample, shows that the construct validity for each of the six subscales is similar to that of the original scale, and documents the age measurement invariance of the scale. Finally, in Study 3, we investigate measurement invariance across Polish and American samples and demonstrate that Poles are characterized by lower scores on healthy emotionality and its dimensions compared to Americans. We conclude that the Polish version of the ESQ is a psychometrically valid and easily implementable measurement tool that can be applied in research settings to measure healthy emotionality and its components.

Keywords Emotional style · Validation · Affective neuroscience · Well-being

Emotions and affect play a crucial role in diverse areas of the human experience, including interpersonal processes, decision-making, psychopathology, and well-being (Fox et al., 2018). Since antiquity, it has been believed that emotionality is one of the building blocks of personality, and that tendencies in affective response and affect regulation are what make individuals unique (e.g., Reisenzein & Weber, 2009). Trait differences in emotionality are also seen as a reflection of basic neural processes (Jackson et al., 2003; Panksepp, 2004; Sander & Scherer, 2009). In line with that assumption, Davidson (2001, 2004; Davidson & Irwin, 1999) proposed a theoretical model rooted in neuroscientific studies of emotion describing trait differences in emotionality via six dimensions governed by specific brain circuits (Davidson & Begley, 2012). These dimensions extend on a continuum between two extremes that reflect amplified or reduced activity in the particular brain circuits that underlie them. Consequently, emotional style is a function of where an individual falls along these six dimensions and is linked to the specific emotional states that people experience, their duration, and intensity. Davidson suggested that Emotional Style is a complex behavioral trait, with variations partially explained by genetic, heritable effects and partially by experiential factors. It has been theorized that individual differences in emotional style are associated with personality and temperament, but are not reducible to them (Davidson & Begley, 2012; Kesebir et al., 2019). To the best of our knowledge, no research to date has investigated cultural differences in emotional style and healthy emotionality.

The Emotional Style paradigm differs from the majority of approaches to individual differences in emotional dispositions, because unlike popular personality inventories that focus mainly on pathological or problematic aspects of emotionality (e.g., anxiety, fearfulness, difficulties in regulating emotion), it focuses on the positive side

✉ Agata Gasiorowska
agasiorowska@swps.edu.pl

Kaja Glomb
kaja.glomb@uj.edu.pl

Pelin Kesebir
kesebir@gmail.com

[1] Faculty of Psychology in Wroclaw, SWPS University of Social Sciences and Humanities, Warsaw, Poland

[2] Faculty of Management and Social Communication, Jagiellonian University, Krakow, Poland

[3] Center for Healthy Minds, University of Wisconsin, Madison, WI, USA

of emotionality, namely emotional health (Kesebir et al., 2019). This approach to emotional health aligns with World Health Organization's classical definition of health, which emphasizes the ability to achieve optimal bio-psycho-social functioning rather than conceiving of health as a lack of pathology (Callahan, 1973). Accordingly, the ESQ captures more than the absence of maladaptive emotional patterns. This understanding of healthy emotionality is also broader than just the intensity and frequency of emotional experiences. Therefore, the ESQ differs from other tools exploring individual differences in currently experienced states or predispositions to experience particular emotions or moods.

The theory underlying the ESQ is also different from other concepts that describe mental and emotional health, such as psychological well-being or emotional intelligence. These concepts, theoretical in nature, emphasize other aspects of functioning than healthy emotionality does. EI is often understood as a combination of cognitive predispositions and personality traits (e.g., Schulte et al., 2004) that do not refer to affective predispositions, but rather to the ability to understand and perceive emotions or to use emotions to achieve one's goals (Mayer et al., 2003). Subjective well-being, in turn, refers to a global judgment of one's affective experiences and one's satisfaction with life and does not offer direct information about a person's functioning in different domains of emotional functioning. Hence, such measures of emotional experiences, personality traits, EI abilities, or psychological well-being are not well-positioned to assess a person's emotional make-up or to study healthy emotionality. The Emotional Style Questionnaire (ESQ) developed by Kesebir et al. (2019) is the first tool appropriate for this purpose. The ESQ is rooted in Davidson's original concept describing the connection between neural asymmetry and the main dimensions of emotional life relevant to psychological well-being (Davidson, 2001, 2004; Davidson & Irwin, 1999). Accordingly, an additional strength of the ESQ framework is its firm grounding in affective neuroscience research, which lends it a different status from purely theoretical constructs.

To sum up, the ESQ is a fully validated tool based on a neuroscientifically supported framework that allows for studying healthy emotionality, understood as a global predisposition to adaptive emotional responses and regulation, together with its various aspects. Its psychometric properties make the ESQ a sound research tool to study several aspects of emotional functioning and overall healthy emotionality, and a valid and reliable measure for use in clinical settings. However, the questionnaire has only been available in the original English version, in Persian (Nazari & Griffiths, 2020), and in German (Jekauc et al., 2021), limiting its potential for language-specific regions. We therefore propose its adaptation and validation for the Polish population, which will enable broader, cross-cultural research on individual differences in emotionality. It has, for example, been demonstrated that when asked about their mood, Polish people frequently describe it "as worse than usual" (Dolinski, 1996). It is of interest to determine whether this tendency also manifests itself as lower levels of healthy emotionality in comparison to other nations, especially those who, such as Americans, typically define their mood as "better than usual" (Szarota, 2010). On the other hand, as indicated by Szarota (2011), it is possible that at least some Poles might not feel like masking their everyday negative feelings to the same extent as Americans do, which might actually be a sign of better emotional health. This issue is investigated further in this manuscript.

Emotional style questionnaire and its dimensions

The ESQ aims to capture the six dimensions of emotional style: Outlook, Resilience, Social Intuition, Self-Awareness, Sensitivity to Context, and Attention. Additionally, as a whole, the questionnaire also assesses a person's overall level of emotional health, labeled Healthy Emotionality (Kesebir et al., 2019).

Outlook refers to an individual's ability to sustain the positive emotions that emerge in the moment, and a general disposition toward positivity over time. Importantly, this dimension should not be equated with the ability to experience positive emotions at all or with intensity. Rather, it is about the temporal course of positive emotions. People who score highly on the Outlook dimension tend to have enduring positive emotions, which have a strong carryover effect and translate into a generally positive, optimistic outlook on life. Indeed, the Outlook dimension is associated with various indicators of well-being (Kesebir et al., 2019; Nazari & Griffiths, 2020). Conversely, individuals with low levels of Outlook have shorter-lived reactions to positive stimuli. Their inability to keep positive emotions alive and vivid for more extended periods results in a negative, pessimistic outlook and might be a risk factor for depression (Kesebir et al., 2019).

Resilience, like Outlook, is a dimension related to the temporal dynamics of an emotional response. Whereas Outlook is about the temporal course of positive emotional stimuli, Resilience, however, is about the temporal course of negative emotional stimuli: it is the capacity to bounce back from negative emotions. People high in Resilience can promptly recover from negative emotions such as sadness, anger, and fear, and they can quickly regain their emotional equanimity after both a minor everyday hassle and major life challenges. Kesebir et al. (2019) found that, although based on distinct brain mechanisms, the Resilience and Outlook dimensions are very closely related. This is not unexpected,

as the ability to bounce back quickly from adverse events would inevitably help the preservation of positive affect, and the ability to sustain positive affect would inevitably help with bouncing back from negative events. Hence, like Outlook, the Resilience dimension is also very closely linked to psychological well-being.

Social Intuition indicates a person's degree of attunement to non-verbal social cues (such as body language, gestures, facial expressions, or vocal intonation), and the ability to infer social information from the emotional states of others. Since noticing and decoding emotional cues is a prerequisite for responding to them, people with high Social Intuition are empathetic and compassionate. On the other hand, extreme insensitivity to the signals mentioned above is characteristic for people on the autism spectrum and may have critical consequences for personal and professional relationships and for psychological well-being (Kesebir et al., 2019).

Self-Awareness refers to the ability to perceive one's bodily signals reflecting emotions, sensitivity to the physiological and emotional cues that occur within one's body, and the ability to recognize and interpret these cues. Low Self-Awareness is associated with limited insight into one's emotional life and into the causes why one acts and reacts in the way one does. Self-Awareness also plays an essential role in empathy and should accordingly lead to better interpersonal relationships and psychological well-being. That said, extreme levels of Self-Awareness may underlie conditions such as panic disorder and hypochondriasis, or be a likely factor in the burnout experienced by people engaged in helping professions (Kesebir et al., 2019).

Sensitivity to Context refers to whether one's emotional and behavioral responses are appropriate given the specific social context and available social cues. In a sense, it is an outer-directed version of Self-Awareness; Self-Awareness reflects attunement to one's own physiological and emotional cues, while Sensitivity to Context reflects attunement to the social environment. People low in Sensitivity to Context might overlook the tacit rules and expectations governing different social situations and fail to modify their behavior accordingly. Therefore, their behavior, such as talking loudly on the phone on crowded public transportation, dressing inappropriately for the workplace, or telling dirty jokes to one's boss, can be seen as insensitive or even improper (Kesebir et al., 2019).

The last dimension of Emotional Style, Attention, pertains to the ability to eliminate distractions and stay focused instead of being caught by the most attention-grabbing stimuli in the environment. In that sense, out of the different forms of attention, it is most similar to selective attention. Although Attention is typically seen as a component of cognitive ability, Kesebir et al. (2019), following Davidson's theory (Davidson & Begley, 2012), proposed that it is also an essential part of the emotional style. This is because

emotional stimuli control a significant share of one's attention, and the capacity to filter out emotional distractions is closely associated with psychological well-being.

In summary, the Emotional Style Questionnaire measures individual differences along six dimensions. It consists of 24 items (four items per dimension), and participants indicate their agreement with these items on a Likert scale ranging from 1 ("Strongly Disagree") to 7 ("Strongly Agree"). Dimensions of Emotional Style are computed as the mean of four items, and an overall Healthy Emotionality score is calculated as the average of all 24 items. Kesebir et al. (2019) demonstrated the convergent and discriminant validity of the individual dimensions and the validity of the overall Healthy Emotionality scale.

Overview of studies

This article presents studies to validate the Polish version of the Emotional Style Questionnaire (ESQ-PL), capturing six dimensions of healthy emotionality as introduced by Kesebir et al. (2019) (*Outlook*, *Resilience*, *Social Intuition*, *Self-Awareness*, *Sensitivity to Context*, and *Attention*). We conducted three studies in which we tested the internal and external validity and reliability of the ESQ-PL and demonstrated that this measure is equivalent to its original English version. Study 1 confirmed the adequacy of the factorial structure of the Polish version of the scale, showed the equivalence of the Polish and English versions, established gender measurement invariance, and found excellent test-retest reliability across an interval of two months. Study 2 confirmed the adequacy of the factorial structure in another sample, demonstrated age measurement invariance, and revealed that the construct validity of each of the six subscales is comparable to that of the original scale. Finally, Study 3 tested cross-cultural measurement invariance and demonstrated that Poles are characterized by a lower level of Healthy Emotionality than Americans. We conclude that the Polish version of the ESQ is a psychometrically valid instrument that can easily be implemented to measure healthy emotionality and its components in research settings.

Study 1

The aim of Study 1 was to provide initial evidence for the internal validity, factor structure, and reliability of the ESQ-PL. Three people fluent in English and Polish (one knowledgeable about the scale content; two Polish native speakers, one English native speaker) independently translated the ESQ to Polish. Polish-speaking members of the research team then reviewed the translations and synthesized a consensual version. Another native English speaker then

translated the consensual version back to English. Finally, all authors reviewed the back translation compared to the original version and provided necessary corrections to create the final scale. In the process, we decided to change the tense of one item from past to present, as that was the only item in the original version that was formulated in the past tense. In addition, since the Polish language distinguishes feminine and masculine grammatical forms, we prepared separate versions of ESQ-PL for women and men.

The readability of the final version of ESQ-PL was good, with a Gunning Fog index of 7, meaning that it would be easily understood by a person with seven years of formal education (i.e., by a student in the last year of primary school).

In Study 1, we invited a sample of Polish participants from Prolific Academic and tested the original structure of the scale, as proposed by Kesebir et al. (2019). We also tested the equivalence of the two language versions and test-retest reliability of the ESQ-PL. After all participants completed the Polish version of the ESQ in the first part of the study, we invited those participants who had declared high English proficiency to fill in the English version, and tested the relationship between their scores in the two language versions. Finally, participants who had completed the ESQ-PL in the first part of the study were contacted after eight weeks and asked to complete the questionnaire again. We chose a time interval longer than in the original ESQ to provide an even more conservative test for the scale score reliability.

Method

Participants and recruitment

Six hundred Polish participants were recruited on Prolific Academic (Palan & Schitter, 2018) to complete an online survey in exchange for 0.40 GBP. Forty-eight participants did not provide valid answers to the attention checks and were excluded from data analysis, leaving a final sample of $N = 552$, comprising 264 women, 282 men, and six people who indicated their gender as "other"; age 17–57 years, $M = 23.51$ ($SD = 6.33$). Out of this sample, 232 participants who had declared an English proficiency level of either C1 or C2 were contacted within 48 h of the first study and invited to complete an online survey in exchange for 0.40 GBP. One hundred and eighty-seven responded within the next 76 h; 104 women, 80 men, three people indicated their gender as "other"; age 18–46 years, $M = 23.06$ ($SD = 5.26$).

Finally, after eight weeks, we invited all participants who had passed attention checks in the first part of the study to participate in an additional study in exchange for 0.40 GBP. Two-hundred-and-thirty of the original 552 participants

(41.67%) accepted this invitation and completed ESQ-PL within the next 24 h. This subsample consisted of 111 women, 116 men, and 3 "other"; age 18–56 years, $M = 24.59$ ($SD = 7.30$).

Procedure and materials

In the first part of the study, after giving their informed consent, the participants provided information on their gender, age, and proficiency in English from 0 ("none") to C2 ("proficient"). Next, they responded to the 24 items of ESQ-PL. They indicated their agreement for each item on a Likert scale ranging from 1 ("Strongly Disagree") to 7 ("Strongly Agree"). All study materials were in Polish. The same procedure was used in the retest part of the study conducted eight weeks after the first part. In the language equivalence part of the study, conducted with those participants proficient in English, after giving informed consent, participants filled out the original English version of ESQ (Kesebir et al., 2019).

Statistical analyses

Using data from the first part of the study, we conducted a confirmatory factor analysis with JASP 0.16. We tested the originally proposed structure of the ESQ scale, which is a hierarchical model grouping the 24 items into six first-order factors representing each dimension of Emotional Style, and grouping the six dimensions into one second-order factor indicating Healthy Emotionality. Additionally, we tested a model in which all 24 items were grouped in one general dimension indicating Healthy Emotionality. We applied the generalized least squares estimation method, i.e. the same one that was used by Kesebir et al. (2019) in their work on the original English version of the scale. We examined goodness of fit using multiple indices: χ^2/df, the root mean square error of approximation (RMSEA) and its 90% confidence interval (90% CI), as well as the goodness-of-fit index (GFI), Comparative-Fit Index (CFI) and Tucker-Lewis Index (TLI). We used multiple fit indices to assess different types of model fit (e.g., model parsimony, absolute fit) and to provide a more reliable, conservative evaluation (Brown, 2015). Following Hu and Bentler (1999), we assumed that the upper boundary for an acceptable fit χ^2/df is 3, the lower boundary for an acceptable fit for GFI, TLI and CFI is 0.90, and the upper boundary for an acceptable fit for RMSEA is 0.06. We also used the expected cross-validation index (ECVI) for model comparisons, with lower values of the index indicating better model fit (Browne & Cudeck, 1992).

After establishing initial validity and measurement invariance for the ESQ-PL (see Results), we calculated the Cronbach's alphas, composite reliability, and average variance extracted (AVE). We used the rule proposed by Fornell and

Larcker ([1981]) to evaluate AVE: the positive square root of the AVE for each latent variable should be higher than the highest correlation with any other latent variable.

In the next step, we conducted a measurement invariance analysis using multi-group CFA (Byrne, [2012]; Kline, [2016]) to assess the psychometric equivalence of our construct across gender groups. First, we evaluated the six-factor model separately for men and women. Then, we examined the psychometric equivalence of the ESQ across the two gender groups in the following phases: (1) configural invariance, assuming the same factor structure in the two groups; (2) metric invariance—additionally, assuming equal factor loadings from items to first-order factors and from first- to second-order factors); (3) scalar invariance, assuming equal intercepts for items; and (4) scalar invariance, with additional constraints of equal intercepts for first-order dimensions. We tested invariance by assessing the changes in model fit and fit indices (i.e., ΔRMSEA, ΔCFI, and ΔTLI). Following Cheung and Rensvold ([2002]) and Vandenberg and Lance ([2000]), we assumed that a change in RMSEA equal to 0.015 and a change in CFI and TLI equal to 0.01 or less would mean that the two models do not differ from each other; a change between 0.01 and 0.02 would mean that the two models may possibly differ, and a change greater than 0.02 would mean that the two models definitely differ from each other. Having established gender measurement invariance (see the Results), we assessed gender differences in Healthy Emotionality and its dimensions using analysis of variance.

In the last steps, we performed attrition analyses to test whether the participants who accepted our invitation to complete the English version of ESQ or the second invitation to complete ESQ-PL differed from those who did not respond to these invitations. Finally, we investigated the similarities and differences between the ESQ and ESQ-PL scores (equivalence of the two language versions) and the test-retest correlations.

Results and discussion

Confirmatory factor analysis

A generalized least squares confirmatory factor analysis (CFA) yielded an acceptable fit for this model in light of most indices, $\chi^2/\text{df} = 2.25$, RMSEA = 0.048, 90% CI [0.042, 0.053], GFI = 0.982, ECVI = 1.290, CFI = 0.683, TLI = 0.644[1]. The fit for an alternative CFA model that grouped all 24 items in a single factor representing Healthy Emotionality was worse than the fit for the original, theoretically derived scale structure, $\chi^2/\text{df} = 4.02$, RMSEA = 0.074, 90% CI [0.069, 0.079], GFI = 0.967, ECVI = 2.103, CFI = 0.217, TLI = 0.142, $\Delta\chi^2/\Delta\text{df} = 76.68$, $p < .001$. Ultimately, although the model fit for the ESQ-PL was worse than for the original version of the scale (Kesebir et al., [2019]), we decided not to reject the solution with the six first-order factors and one second-order factor, concluding that this theoretical model characterized the data collected in the Polish sample in an acceptable way.

Table [1] shows the factor loadings of items to their dimensions and corrected item-total correlations within the dimensions. All standardized regression weights between the first- and second-order latent factor in this model were significant (βs > 0.15, ps < 0.012). Most standardized regression weights between items and first-order latent factors indicated a moderate or strong relationship between the items and the first-order latent factors (βs > 0.24, ps < 0.001), with the only exceptions being item 22 from the Self-Awareness dimension ("Usually, I am not attentive to what is going on in my body" in the English version) and item 23 from the Sensitivity to Context dimension ("Oftentimes, when other people think something is inappropriate, I disagree" in the English version). These two items had a factor loading lower than in the original version of the ESQ scale. Although the discrimination indices for these items were satisfactory, we decided to test whether these two items should potentially be included in different scales. We conducted additional CFA, this time assuming that item 22 and item 23 can load on any of the six dimensions of ESQ-PL. The model fit was similar to the original one, $\chi^2/\text{df} = 2.29$, $\Delta\chi^2/\Delta\text{df} = 1.41$, $p = .170$, RMSEA = 0.048, 90% CI [0.043, 0.054], GFI = 0.982, ECVI = 1.301, CFI = 0.687, TLI = 0.634, and the two focal items clearly did not load to different latent variables to a higher extent than to their original dimensions (see Table S1 in Supplementary materials). Furthermore, we tested three alternative models, (1) excluding item 22 from the Self-Awareness dimension, (2) excluding item 23 from the Sensitivity to Context dimension, and (3) excluding both items 22 and 23 from the scale. The fit for these models is presented in Table S2 in the Supplementary materials. None of the three alternative models was significantly better than the default model with all 24 items. Therefore, we decided to keep items 22 and 23 in their original dimensions.

---

[1] The CFI and the TLI are incremental fit indices that compare the theoretical model with a null (independence) model, while RMSEA is an absolute fit index. In our case, RMSEA and GFI suggest adequate fit, while CFI and TFI do not, which might imply that the fit for a null model is relatively good. Hence, we investigated the RMSEA for a model that assumes no relation between 24 items in the ESQ scale and found that it was relatively low, for example in study 1:

Footnote 1 (continued)

RMSEA = 0.08, 90% CI [0.075, 0.084]. As suggested by Kenny ([2015]) and van Laar et al. ([2021]), when the RMSEA of the null model is below 0.158, using incremental fit indices for the evaluation of the proposed model may not be informative.

Table 1 Psychometric properties of the Polish version of the Emotional Style Questionnaire

Item number	Dimension	Study 1 (N = 552)		Study 2 (N = 418)	
		β	r	β	r
1 (r)	Outlook	0.498	0.418	0.527	0.700
7	Outlook	0.741	0.622	0.786	0.632
13	Outlook	0.745	0.587	0.728	0.622
19 (r)	Outlook	0.746	0.592	0.694	0.639
2 (r)	Resilience	0.725	0.613	0.740	0.710
8	Resilience	0.614	0.521	0.748	0.716
14 (r)	Resilience	0.655	0.571	0.668	0.758
20	Resilience	0.764	0.630	0.777	0.697
3	Social Intuition	0.785	0.697	0.794	0.670
9 (r)	Social Intuition	0.589	0.550	0.509	0.756
15	Social Intuition	0.666	0.617	0.671	0.686
21	Social Intuition	0.640	0.629	0.521	0.716
4 (r)	Self-Awareness	0.630	0.520	0.625	0.615
10	Self-Awareness	0.802	0.586	0.694	0.679
16 (r)	Self-Awareness	0.743	0.655	0.599	0.584
22 (r)	Self-Awareness	0.277	0.415	0.114	0.709
5 (r)	Sensitivity to Context	0.818	0.659	0.762	0.628
11 (r)	Sensitivity to Context	0.534	0.518	0.548	0.706
17 (r)	Sensitivity to Context	0.694	0.598	0.743	0.635
23 (r)	Sensitivity to Context	0.245	0.296	0.216	0.789
6	Attention	0.773	0.633	0.775	0.508
12	Attention	0.760	0.604	0.787	0.580
18 (r)	Attention	0.451	0.444	0.296	0.690
24 (r)	Attention	0.722	0.593	0.677	0.545
Outlook	Healthy emotionality	0.879	0.568	0.901	0.575
Resilience	Healthy emotionality	0.958	0.426	0.930	0.475
Social Intuition	Healthy emotionality	0.226	0.210	0.362	0.293
Self-Awareness	Healthy emotionality	0.588	0.534	0.740	0.569
Sensitivity to Context	Healthy emotionality	0.148	0.272	0.258	0.303
Attention	Healthy emotionality	0.638	0.439	0.725	0.520

(r) – reverse coded item; β – standardized factor loadings; r – corrected item-total correlations

Measurement invariance

Next, we assessed measurement invariance across gender groups using multi-group CFA (Byrne, 2012; Kline, 2016). We evaluated the six-factor model separately for men

($n = 282$) and women ($n = 264$), excluding six participants who had not reported their gender. Generalized least squared CFA for the men yielded an acceptable model fit for most but not all indices, $\chi^2/df = 1.61$, RMSEA = 0.047, 90%CI [0.038, 0.055], GFI = 0.976, ECVI = 1.964, CFI = 0.680, TLI = 0.641, similarly to the CFA for the women, $\chi^2/df = 1.49$, RMSEA = 0.048, 90%CI [0.039, 0.056], GFI = 0.975, ECVI = 2.085, CFI = 0.674, TLI = 0.634. Overall, we conclude that these results provide initial support for the configural invariance. Hence, we proceeded to perform a test of measurement invariance across gender groups regarding metric, scalar, and residual invariance (Table 2).

We examined invariance by assessing the changes in model fit and fit indices (i.e., ΔRMSEA, ΔCFI, and ΔTLI). Concerning metric and scalar invariance, the ΔRMSEA and ΔTLI were lower than 0.01, indicating that factor loadings from items to first-order factors and from first-order factors to Healthy Emotionality, as well as intercepts for items, did not differ across gender, hence allowing for between-gender comparisons. However, this conclusion should be treated with caution since ΔCFI was larger than 0.01, thus not providing support for the existence of invariance. Finally, we did not find support for scalar invariance when we fixed the intercepts for first-order factors across gender, indicating gender differences concerning the dimensions of healthy emotionality.

Gender differences

In its original version, the ESQ yielded gender differences with regard to the overall Healthy Emotionality score, as well as the Social Intuition and Sensitivity to Context dimensions. Although we did not find gender differences in Healthy Emotionality, $t(544) = 0.51$, $p = .613$, Cohen's $d = 0.04$, we replicated the differences with regard to Social Intuition, $t(544) = -3.78$, $p < .001$, $d = 0.32$ (women: $M = 5.11$, $SD = 1.14$, men: $M = 4.74$, $SD = 1.14$) and Sensitivity to Context, $t(544) = -5.38$, $p < .001$, $d = 0.46$ (women: $M = 4.68$, $SD = 1.14$, men: $M = 4.12$, $SD = 1.28$). Additionally, we found that, in our sample, men scored higher on Outlook, $t(544) = 3.14$, $p = .002$, $d = 0.27$ (women: $M = 4.06$, $SD = 1.23$, men: $M = 4.38$, $SD = 1.19$), Resilience, $t(544) = 5.23$, $p = .002$, $d = 0.27$ (women: $M = 3.47$, $SD = 1.20$, men: $M = 4.01$, $SD = 1.23$), and Attention, $t(544) = 3.47$, $p < .001$, $d = 0.30$ (women: $M = 3.60$, $SD = 1.23$, men: $M = 3.95$, $SD = 1.13$). We did not find gender differences in Self-Awareness, $t(544) = -0.98$, $p = .327$, $d = 0.08$.

Internal consistency and composite reliability

Table 3 presents the descriptive statistics for the six dimensions of the ESQ-PL, their factor loadings to, and intercorrelations with, the overall score, Cronbach's alphas, composite

Table 2 Measurement invariance across gender groups in Study 1 ($n = 546$)

Type of invariance	χ^2/df	$\Delta\chi^2$/Δdf	RMSEA	GFI	CFI	TLI	ΔRMSEA	ΔCFI	ΔTLI
Configural	1.60	-	0.047	0.975	0.677	0.637	-	-	-
Metric	1.60	1.68*	0.047	0.974	0.660	0.635	< 0.001	− 0.017	− 0.002
Scalar Intercepts for items only	1.62	2.92*	0.048	0.973	0.642	0.629	0.001	− 0.018	− 0.006
Scalar Intercepts for items and first-order factors	1.67	6.98**	0.050	0.972	0.609	0.609	0.002	− 0.033	− 0.030

* p < .05, ** p < .01

reliability, and average variance extracted. Although the internal consistency indicators for the 24-item Polish version of ESQ and each of its dimensions were slightly lower than those for original and Persian versions of the scale (Kesebir et al., 2019; Nazari & Griffiths, 2020), the Cronbach's alphas and composite reliability reached acceptable-to-good values (see Table 3). Concerning the correlations between latent variables representing the ESQ-PL dimensions, the square root of the AVE for four out of six dimensions was higher than their correlations with other dimensions, asserting their good discriminant validity. However, the square root of the AVE for Outlook and Resilience was lower than the intercorrelation between these two dimensions, suggesting potential overlap between their content. This result was expected since the two dimensions were highly correlated in the English (Kesebir et al., 2019) and Persian (Nazari & Griffiths, 2020) versions of the scale as well. This correlation makes theoretical sense: Outlook pertains to the ability to maintain positive emotion, while Resilience refers to the ability to recuperate from negative emotion. These two abilities doubtlessly facilitate each other and are closely intertwined.

Unlike in the original English version, we did not find significant correlations between all the dimensions of Emotional Style. Specifically, Resilience did not correlate with Social Intuition and Sensitivity to Context, and Social Intuition did not correlate with Attention ($ps > 0.30$). However, as in the original ESQ, the strongest correlation was between Outlook and Resilience, highlighting the close association between the ability to recover from negative emotion and to sustain positive emotion.

Attrition analysis for test-retest

We performed an attrition analysis to test whether the 230 individuals who accepted our invitation to complete the ESQ-PL at Time 2 differed from the 322 who did not respond to the invitation. The only dimension for which we found a significant difference was Attention, $F(1, 550) = 4.99$, $p = .026$, $\eta^2 = 0.009$, with those who saw and accepted our second invitation scoring higher ($M = 3.90$, $SD = 1.19$) than those who did not ($M = 3.66$, $SD = 1.19$). However, the effect size was less than 1% of the explained

variance. We did not observe any differences in average scores for Healthy Emotionality or the remaining five dimensions between these two groups, $Fs < 1.5$, $ps > 0.230$, $\eta^2s < 0.003$. We also observed some differences between the two groups in their age, $F(1, 550) = 11.68$, $p = .001$, $\eta^2 = 0.010$, with people who responded to our second invitation being older and more diverse in age ($M = 24.59$, $SD = 7.30$) than those who did not ($M = 22.74$, $SD = 5.41$). Again, the effect size was relatively small, not exceeding 1% of the variance. We did not detect any differences in drop-out depending on gender, $\chi^2(2, N = 552) = 0.22$, $p = .900$. These results suggest that although the sample did not experience substantial differential attrition for the variable of primary interest to this study, it did become slightly older and more attentive.

Test-retest reliability

We calculated Pearson correlations between the Time 1 and Time 2 scores for overall Healthy Emotionality and the six subscales of the ESQ. Across eight weeks, the test-retest reliability coefficient for Healthy Emotionality was 0.81, signifying excellent reliability. The coefficients for the individual subscales ranged between acceptable and very good ($rs > 0.70$, $ps < 0.001$; see Table 3).

Attrition analysis for equivalence of language versions

We performed an attrition analysis to test whether the 187 participants who accepted our invitation to complete the English version of ESQ differed from the 277 participants who completed the Polish version of the scale only. We did not observe any differences between these two groups in overall Healthy Emotionality or its dimensions, $F(6, 545) = 0.89$, $p = .499$, $\eta^2 = 0.010$, Wilks $\lambda = 0.029$. No significant age differences were found between the two groups either, $F(1, 550) = 1.46$, $p = .228$, $\eta^2 = 0.003$. Men were slightly more likely to drop out of the study than women, $\chi^2(2, N = 552) = 8.10$, $p = .017$; however, overall, these results suggest that the sample in this study did not experience markedly differential attrition.

Table 3 Descriptive statistics in Study 1 ($N = 552$)

Dimension	M	SD	β	α	CR	AVE	Correlations between dimensions and total score						Test-retest correlations ($n = 230$)	ESQ - ESQ-PL correlations ($n = 187$)
							1	2	3	4	5	6		
1. Outlook	4.22	1.22	0.88	0.75 [0.72, 0.79]	0.78	0.48							0.79***	0.84***
2. Resilience	3.74	1.25	0.96	0.78 [0.75, 0.81]	0.78	0.48	0.68***						0.71***	0.79***
3. Social Intuition	4.92	1.16	0.23	0.81 [0.78, 0.83]	0.77	0.45	0.11**	−0.04					0.77***	0.84***
4. Self-Awareness	4.87	1.17	0.58	0.74 [0.71, 0.78]	0.72	0.42	0.37***	0.24***	0.36***				0.72***	0.71***
5. Sensitivity to Context	4.39	1.24	0.15	0.72 [0.69, 0.76]	0.68	0.37	0.15***	0.04	0.24***	0.33***			0.70***	0.71***
6. Attention	3.76	1.20	0.64	0.77 [0.73, 0.80]	0.78	0.47	0.40***	0.42***	0.05	0.34***	0.16***		0.73***	0.79***
7. Healthy Emotionality	4.32	0.75	-	0.85 [0.83, 0.86]	0.91	0.31	0.73***	0.64***	0.45***	0.71***	0.52***	0.65***	0.81***	0.86***

*** $p < .001$, ** $p < .01$, * $p < .05$; β - factor loadings of dimensions to the overall score, α - internal consistency (Cronbach's α) with 95% CI, CR - composite reliability, AVE - average variance extracted

Equivalence of the two language versions

As expected, we observed high correlations between participant scores in the Polish and English versions of the ESQ (rs < 0.70, ps < 0.011, see Table 2). The scores in the Polish version were slightly higher than those in the English version for the overall score and for two out of the six dimensions: Healthy Emotionality: $F(1, 186) = 14.12$, $p < .001$, $\eta^2 = 0.071$ (PL: $M = 4.29$, $SD = 0.80$, EN: $M = 4.18$, $SD = 0.79$), Self-Awareness: $F(1, 186) = 20.04$, $p < .001$, $\eta^2 = 0.097$ (PL: $M = 4.93$, $SD = 1.24$, EN: $M = 4.62$, $SD = 4.62$), and Sensitivity to Context: $F(1, 186) = 7.14$, $p < .001$, $\eta^2 = 0.037$ (PL: $M = 4.34$, $SD = 1.24$, EN: $M = 4.16$, $SD = 1.20$). No differences were obtained for the remaining four dimensions, i.e. Outlook: $F(1, 186) = 2.03$, $p = .156$, $\eta^2 = 0.011$ (PL: $M = 4.13$, $SD = 1.30$, EN: $M = 4.05$, $SD = 1.27$), Resilience: $F(1, 186) = 1.03$, $p = .311$, $\eta^2 = 0.006$ (PL: $M = 3.71$, $SD = 1.31$, EN: $M = 3.65$, $SD = 1.33$), Social Intuition: $F(1, 186) = 1.00$, $p = .200$, $\eta^2 = 0.009$ (PL: $M = 4.95$, $SD = 1.20$, EN: $M = 4.88$, $SD = 1.19$), and Attention: $F(1, 186) = 0.004$, $p = .950$, $\eta^2 < 0.001$ (PL: $M = 3.70$, $SD = 1.27$, EN: $M = 3.69$, $SD = 1.39$).

Conclusions

Overall, this study provided initial support for the six distinct dimensions constituting Healthy Emotionality in a Polish sample. It also demonstrated that measurement with the Polish version of the ESQ could be considered similar to measurement with the English version in a Polish sample. Furthermore, as in the Persian version of the ESQ (Nazari & Griffiths, 2020), we verified configural, metric, and scalar invariance in gender groups and tested gender differences in Healthy Emotionality and its dimensions. Finally, we found high test-retest reliability for the ESQ-PL and its subscales over eight weeks, further confirming the scale's psychometric adequacy. However, two of the scale items had relatively low factor loadings. Although we did not find a measurement model that would fit the data better than the original theoretical model, the overall fit of the six-factor model was not perfect.

Study 2

Having gathered support for the proposed structure and reliability of the ESQ-PL and its equivalence with the English version in Study 1, Study 2 had three aims: (1) to validate the adequacy of ESQ-PL's factor structure in a different sample, more diverse in terms of age, (2) to determine whether the scale and its subscales demonstrate any age or gender differences, and (3) most importantly, to establish the convergent and divergent validity of the subscales that make up

the ESQ-PL (Campbell & Fiske, 1959). With these aims, we asked a large sample of participants to complete the ESQ-PL and several other questionnaires measuring constructs that had been demonstrated to be related to the ESQ dimensions by Kesebir et al. (2019).

We expected the structure of correlations to be similar to that gathered for the original ESQ. Specifically, we expected that: (1) Outlook will be strongly related to optimism, as the capacity to sustain positive expectancies for the future should correlate with the ability to maintain positive emotions; (2) Resilience will be strongly related to resilience measured with the Brief Resilience Coping Scale; (3) Social Intuition will reveal a strong negative relationship with the Autism Spectrum Quotient; (4) Self-Awareness will be strongly related to mindful attention awareness; (5) Sensitivity to Social Context will correlate with impression management, which is an indicator of being aware of the implicit rules and expectations that rule social situations and being willing and able to adjust one's responses to these; and (6) Attention will correlate with another attention scale measuring individual differences in focusing attention and in shifting attention between tasks, and with mindful attention awareness, which refers to being attentive to what is happening in the present.

Method

Participants and recruitment

Four hundred and fifty-seven Polish participants were recruited on Prolific Academic (Palan & Schitter, 2018) to complete an online survey in exchange for 1.90 GBP. We excluded 39 participants who failed to respond correctly to attention checks, resulting in a sample of 418 participants (226 women, 188 men, 4 "Other"; age 18–75, $M = 31.88$, $SD = 14.37$). Since the participants could quit the study at any point and some did, sample sizes for individual analyses vary.

Procedure and materials

After giving their informed consent, participants completed the questionnaires described below. The ESQ was always presented first, and then the subsequent measures were presented in random order.

Emotional style questionnaire Participants filled in the 24-item ESQ-PL. For each item, they answered using a Likert scale from 1 ("Strongly Disagree") to 7 ("Strongly Agree"). As in Study 1, scores for the dimensions of Emotional Style were computed as the mean of four items, and

an overall Healthy Emotionality score was calculated as the average of all 24 items.

Life orientation test-revised We measured optimism with the Life Orientation Test-Revised (LOT-R; Scheier et al., 1994; Polish adaptation by Poprawa and Juczynski in Juczynski, 2001), which assesses people's expectations concerning the favorability of future events and outcomes. On a seven-point scale ranging from 1 ("Strongly Disagree") to 7 ("Strongly Agree"), participants indicated their agreement with six statements such as "I am always optimistic about my future" in addition to four filler statements.

Brief Resilience Coping Scale (BRCS) As the Brief Resilience Scale (Smith et al., 2008) used for validating the original ESQ has not been adapted to Polish, we used the Polish version of the Brief Resilience Coping Scale (Sinclair & Wallston, 2004; Polish adaptation: Piórowska et al., 2017). It consists of four statements such as "I actively look for ways to replace the losses I encounter in life." Participants responded on a five-point Likert scale from 1 ("Strongly Disagree") to 5 ("Strongly Agree").

Autism Spectrum Quotient (AQ-10) To measure autism, we used the 10-item Autism Spectrum Quotient (Allison et al., 2012; Polish adaptation: Pisula et al., 2013). On a four-point Likert scale ranging from 1 ("Strongly Agree") to 4 ("Strongly Disagree"), participants answered to items such as "When I'm reading a story, I find it difficult to work out the characters' intentions" and "I often notice small sounds when others do not."

Mindful Attention Awareness Scale (MAAS) Participants responded to 15 items, such as "I snack without being aware that I'm eating" and "I find myself doing things without paying attention," using a scale from 1 ("Almost always") to 6 ("Almost Never") (Brown & Ryan, 2003; Polish adaptation: Radon, 2014).

Multidimensional Assessment of Interoceptive Awareness (MAIA) Participants responded to 32 items constituting a multidimensional measure of interoceptive body awareness (Mehling et al., 2012; Polish adaptation: Brytek-Matera & Kozieł, 2015). MAIA measures dimensions such as awareness of body sensations and awareness of the connection between the body, sensations, and emotional states. Sample items include "I trust my body sensations," and "I notice when I am uncomfortable in my body." Participants responded to this measure using a six-point Likert scale from 0 ("it does not apply to me at all") to 6 ("it applies to me very much").

Table 4 Measurement invariance across age groups in Study 2 (n = 418)

Type of invariance	χ^2/df	$\Delta\chi^2$/Δdf	RMSEA	GFI	CFI	TLI	ΔRMSEA	ΔCFI	ΔTLI
Configural	1.60	-	0.054	0.968	0.549	0.494	-	-	-
Metric (weak factorial)	1.57	0.84	0.052	0.967	0.556	0.524	< 0.002	0.006	0.028
Scalar (strong factorial) Intercepts for items only	1.56	1.49†	0.052	0.966	0.541	0.524	< 0.001	− 0.014	0.003
Scalar (strong factorial) Intercepts for items and first order factors	1.61	7.06***	0.055	0.965	0.493	0.480	− 0.003	− 0.047	− 0.044

† $p < .1$, ***$p < .001$

Balanced inventory of desirable responding (Impression management subscale) In the original ESQ validation, Kesebir et al. (2019) used the Impression Management subscale of the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1991). We employed the same scale adapted to Polish by Fronczak (unpublished), with items measuring impression management such as "I sometimes tell lies if I have to" or "I never swear." Participants indicated their agreement with these statements on a seven-point Likert scale from 1 ("Completely disagree") to 7 ("Completely agree"). Following scale instructions, the score was determined by counting the number of items for which participants responded "6" and "7".

Attentional control scale Participants responded to the 20 items constituting the Attentional Control Scale (Derryberry & Reed, 2002; Polish adaptation: Fajkowska & Derryberry, 2010), using a scale from 1 ("Never") to 4 ("Always"). Sample scale items include "My concentration is good even if there is music in the room around me" and "I can quickly switch from one task to another."

Statistical analyses

As in Study 1, we conducted confirmatory factor analysis with JASP 0.16. We tested the ESQ's original hierarchical structure and a model in which all 24 items were grouped in one general dimension indicating Healthy Emotionality (see Fig. 1 in the Supplemental materials) using the generalized least squares estimation method. We examined the goodness of fit using the same multiple indices as in Study 1. Furthermore, we conducted measurement invariance analysis using multi-group CFA (Byrne, 2012; Kline, 2016) to assess the psychometric equivalence of our construct across age groups using the same procedure as in Study 1. Then, we analyzed the age and gender differences for the ESQ-PL scores. Finally, we tested the convergent and divergent validity of ESQ-PL using Pearson correlations between Healthy Emotionality and its dimensions and the other questionnaires in the survey.

Results and discussion

Confirmatory factor analysis and scale psychometrics

As a first step, we retested the proposed model of the scale, grouping the 24 items into six first-order factors representing each dimension of Emotional Style, and the six dimensions into one second-order factor representing Healthy Emotionality. A generalized least squares CFA again demonstrated good fit for this model, χ^2/df = 2.31, RMSEA = 0.056, 90% CI [0.050, 0.062], GFI = 0.975, ECVI = 1.753, CFI = 0.564, TLI = 0.511. An alternative CFA model grouping all 24 items in a single factor representing Healthy Emotionality again was worse in terms of model fit than the fit with the original scale structure, χ^2/df = 3.32, RMSEA = 0.075, 90% CI [0.069, 0.081], GFI = 0.83, ECVI = 2.374, CFI = 0.208, TLI = 0.133. As in Study 1, most standardized regression weights between items and first-order latent factors indicated a moderate or strong relationship between the items and the first-order latent factors, except for item 22 from the Self-Awareness dimension and item 23 from the Sensitivity to Context dimension (see Table 1). However, again, the discrimination indices for these items proved to be satisfactory. Therefore, we assumed that these results supported the six-factor structure of the ESQ-PL.

The scale demonstrated good internal consistency (Cronbach's α = 0.85, 95%CI [0.83, 0.87]). Table 4 presents the descriptive statistics, internal consistency information, and intercorrelations with other dimensions for each of the ESQ subscales.

Measurement invariance

We assessed measurement invariance across age groups using multi-group CFA (Byrne, 2012; Kline, 2016). We evaluated the six-factor model separately for participants aged 25 and below (n = 206) and those aged 26 and above (n = 212). Generalized least squared CFA yielded a slightly better fit in the younger subsample, χ^2/df = 1.44, RMSEA = 0.046, 90%CI [0.035, 0.057], GFI = 0.974, ECVI = 2.491, CFI = 0.654, TLI = 0.612, than in the older subsample, χ^2/df = 1.75, RMSEA = 0.060, 90%CI [0.050, 0.069], GFI = 0.962, ECVI = 2.798, CFI = 0.452,

TLI = 0.385. However, since the model fit was acceptable in light of most indices in both subgroups, we concluded that it supports configural invariance. Hence, we proceeded to perform a test of measurement invariance across age groups regarding metric, scalar, and residual invariance (Table 4).

As in Study 1, we examined invariance by assessing the changes in model fit and fit indices (i.e., ΔRMSEA, ΔCFI, and ΔTLI). Concerning metric invariance, the ΔRMSEA and ΔCFI were lower than 0.01, and all the index changes were positive, so the fit for the model with fixed factor loading was better than for the model with no such constraints, indicating that factor loadings from items to first-order factors and from first-order factors to Healthy Emotionality were equal across age groups. We also found support for scalar invariance when we imposed constraints on item intercepts, meaning that they did not differ across age groups, allowing for age comparisons. However, this conclusion should be treated with caution since ΔCFI was higher than 0.01, therefore not providing support for the existence of invariance. Finally, we did not find support for scalar invariance when we fixed the intercepts for first-order factors across age groups, indicating age differences concerning the dimensions of Healthy Emotionality.

Age and gender differences

As in the original version of the ESQ, we found a significant positive correlation between Healthy Emotionality and age, $r = .172$, $p < .001$. In the Polish sample, like in the American sample, age was significantly correlated with Resilience ($r = .22$, $p = .001$) and Attention ($r = .24$, $p < .001$). However, unlike in the American sample, no correlations were observed with Outlook ($r = .08$, $p = .091$), Self-Awareness ($r = .07$, $p = .129$), Sensitivity to Context ($r = .07$, $p = .175$), or Social Intuition ($r = -.02$, $p = .668$).

As in Study 1, we did not find any gender differences in overall Healthy Emotionality, $t(412) = -0.55$, $p = .582$, Cohen's $d = 0.05$, but we replicated the differences in Social Intuition, $t(412) = -3.14$, $p = .002$, $d = 0.31$ (women: $M = 5.05$, $SD = 1.12$, men: $M = 4.50$, $SD = 1.10$) and Sensitivity to Context, $t(412) = -3.30$, $p = .012$, $d = 0.25$ (women: $M = 4.56$, $SD = 1.31$, men: $M = 4.24$, $SD = 1.24$). Again, as in Study 1, we found that men scored higher on Resilience, $t(412) = -3.70$, $p < .001$, $d = 0.36$ (women: $M = 3.58$, $SD = 1.27$, men: $M = 4.03$, $SD = 1.18$) and Attention, $t(412) = -2.12$, $p = .035$, $d = 0.21$ (women: $M = 3.97$, $SD = 1.15$, men: $M = 4.19$, $SD = 1.02$). We did not replicate the gender differences in Outlook that we found in Study 1, $t(412) = -1.60$, $p = .110$, $d = 0.16$. Yet, as the effect size was comparable to the one observed in Study 1, we suspect that the lack of significance might be due to the smaller sample size in this study. Finally, as in Study 1, we did not find significant gender differences in Self-Awareness, $t(412) = -0.40$,

$p = .689$, $d = 0.04$. In sum, the structure and effect sizes for gender differences in this study were similar to those found in Study 1.

Convergent validity analyses

The descriptive statistics and Cronbach's alphas for the ESQ dimensions and each convergent validity measure, together with the correlations between these variables, are presented in Table 5. Our discussion on convergent validity is organized around each instrument that we used for testing the validity of ESQ-PL.

Outlook The validity of the Outlook dimension was confirmed in the original version of the ESQ by its strong correlation with optimism, as assessed with the Revised Life Orientation Test (Scheier et al., 1994). In the Polish version, the correlation between Outlook and optimism was weaker than in its English version. However, it was still a strong correlation and the strongest correlation observed among the six dimensions of Healthy Emotionality. We also noted significant correlations of Outlook with some other measures in theoretically meaningful ways (e.g., Brief Resilience Coping Scale), although these correlations were lower than those found in the original version. Yet, no other measure had a higher correlation with Outlook than LOT-R in our version of the scale. Altogether, this pattern of results confirms the convergent and divergent validity of the Outlook subscale of the ESQ-PL.

Resilience In the English version of the ESQ, Resilience correlated strongly with the Brief Resilience Scale (Smith et al., 2008), followed by the LOT-R and Attentional Control Scale. For the ESQ-PL, we found a similar pattern of results, with the difference that Resilience correlated with the LOT-R slightly more strongly than it correlated with the Brief Resilience Coping Scale (albeit the difference between these two correlations was not significant). This likely stems from using a different instrument to measure resilience than the one used in the original study. However, this pattern of results still is adequate to establish the validity of the Resilience dimension of the ESQ-PL.

Social intuition In the original version of the ESQ, Social Intuition showed a strong negative correlation with autistic tendencies and a strong positive correlation with MAIA; the correlations with the other scales were much lower. We observed a similar pattern of results with the ESQ-PL, although the correlations with AQ-10 and MAIA were slightly weaker. In addition, we also found a correlation between Social Intuition and the Brief Resilience Coping

Table 5 Descriptive statistics, internal consistencies and correlations for the Emotional Style Questionnaire dimension and external validation measures ($N = 418$)

Dimension	M	SD	Cronbach's α	1	2	3	4	5	6	7
1. Outlook	4.23	1.18	0.71 [0.66, 0.75]	-						
2. Resilience	3.78	1.25	0.77 [0.74, 0.81]	0.65***	-					
3. Social Intuition	4.89	1.13	0.76 [0.72, 0.80]	0.20***	0.05	-				
4. Self-Awareness	4.80	1.21	0.71 [0.67, 0.75]	0.38***	0.24***	0.41***	-			
5. Sensitivity to Context	4.40	1.28	0.75 [0.71, 0.79]	0.15**	0.10*	0.19***	0.44***	-		
6. Attention	4.06	1.10	0.65 [0.59, 0.70]	0.47***	0.52***	0.16**	0.37***	0.19***	-	
7. Healthy Emotionality	4.36	0.77	0.85 [0.83, 0.87]	0.73***	0.67***	0.51***	0.73***	0.55***	0.69***	-
8. LOT-R ($n = 403$)	3.18	0.76	0.76 [0.72, 0.79]	0.71***	0.50***	0.20***	0.34***	0.10	0.34***	0.56***
9. Brief Resilience Coping Scale ($n = 410$)	3.48	0.76	0.73 [0.69, 0.77]	0.49***	0.45***	0.32***	0.29***	0.07***	0.38***	0.51***
10. Autism Spectrum Quotient ($n = 405$)	2.22	0.40	0.64 [59, 0.69]	−0.15**	−0.07	−0.49***	−0.40***	−0.17***	−0.14**	−0.36***
11. MAAS ($n = 405$)	3.96	0.76	0.86 [0.84, 0.87]	0.27***	0.24***	0.12*	0.41***	0.32***	0.39***	0.45***
12. MAIA ($n = 405$)	3.70	0.68	0.90 [0.89, 0.91]	0.31***	0.28***	0.32***	0.29***	0.01	0.31***	0.39***
13. Impression Management (BIDR) ($n = 405$)	3.97	0.87	0.80 [0.77, 0.83]	0.24***	0.27***	0.25***	0.33***	0.29***	0.28***	0.43***
14. Attentional Control Scale ($n = 407$)	2.64	0.41	0.84 [0.82, 0.86]	0.45***	0.47***	0.23***	0.36***	0.19***	0.62***	0.59***

*** $p < .001$, ** $p < .01$, * $p < .05$

LOT-R = Life Orientation Test-Revised; MAAS = Mindful Attention Awareness Scale; MAIA = Multidimensional Assessment of Interoceptive Awareness; BIDR = Balanced Inventory of Desirable Responding

Scale, similar in size to the correlation with MAIA. This correlation was much weaker in the original version of the ESQ, and has likely at least partially to do with the different resilience measure we used in our study. The overall pattern of results confirmed the validity of the Social Intuition subscale.

Self-awareness Kesebir et al. (2019) found that Self-Awareness correlated strongly and positively with the MAAS and the MAIA, while it correlated negatively with AQ-10. All other correlations, although significant, were much weaker. We found a slightly different pattern of correlations for Self-Awareness in our study, ranging only from 0.29 to 0.41, with the strongest correlations for MAAS and AQ-10 (negative). Although this pattern of results confirms the convergent validity of the Self-Awareness dimension, it does not provide enough support for its discriminant validity.

Sensitivity to context Kesebir et al. (2019) demonstrated a close association between Sensitivity to Context and two external criteria: the desire to manage impressions about oneself and Mindful Attention Awareness. Our study found that the correlations of the Sensitivity to Context with the IM subscale of the BIDR (Paulhus, 1991) and MAAS were modest in size. The lack of a stronger correlation between impression management and the Sensitivity to Context dimension, as in the American sample, might potentially be explained by cultural differences in impression management concerns in the two countries, or the BIDR scale having worse psychometric properties in a Polish sample (Fronczak, unpublished). However, it might also be attributed to the generally lower level of correlations between dimensions of ESQ-PL and the other measures we used. Moreover, since the correlations between Sensitivity to Context from the ESQ-PL and the remaining criteria were much weaker than the correlations with IM and MAAS, we believe this attests to both the convergent and divergent validity of this subscale.

Attention The pattern of correlations for the Attention subscale of the ESQ-PL was the same as in the original version of the scale. As expected, we found the highest correlation between the Attention subscale and the Attentional Control Scale, followed by medium-sized correlations with the Brief Resilience Coping Scale and MAAS. This result supports the convergent and discriminant validity of the Attention subscale of the ESQ-PL.

In summary, Study 2 confirmed the proposed structure of the ESQ-PL and its psychometric adequacy in a second sample and replicated the small but significant age and gender effects on the dimensions of Healthy Emotionality. Most importantly, we demonstrated that the subscales of the

ESQ-PL correlate with theoretically related constructs, for the most part comparable to the relevant subscales in the English version, with somewhat less consistent results obtained only for the Self-Awareness dimension. Overall, these results gave us confidence that the Polish version of the ESQ has a construct validity comparable to that of the English version.

Study 3

The previous studies showed that the Polish version of the ESQ demonstrates good reliability and validity, notwithstanding slight differences from the original English version. In the final study, we aimed to test measurement invariance across countries and compare participants from Poland and the U.S.A. with regards to Healthy Emotionality and its dimensions. However, we decided not to utilize the data for the original scale collected by Kesebir et al. (2019). These data were collected at the beginning of 2018, while the data for the Polish version were collected during and right after the third wave of the COVID-19 outbreak (June-July 2021). If we observed any differences in ESQ scores between the two samples, it would be unclear whether this reflects a generally less or more adaptive emotional style or whether it is an artifact of the timing of the study, capturing the temporarily lowered emotional health in response to the pandemic. We thus collected new data in the U.S.A. to test measurement invariance across the two countries and to contribute to the literature on cross-cultural comparisons of well-being.

Method

Participants and recruitment

In September 2021, we recruited via Prolific Academic (Palan & Schitter, 2018) 976 American participants, constituting a representative sample with respect to age, gender, and ethnicity (487 women, 482 men, 7 "other"; age $M = 44.57$, $SD = 16.02$; ethnicity 7.17% Asian, 13.41% Black, 3.95% mixed, 3.64% other, 71.83% White). Participants (to be referred as the "American COVID sample") completed an online survey in exchange for 1.50 GBP. No data was discarded.

We analyzed this data together with the combined ESQ data from our Studies 1 and 2 (to be referred as the "Polish COVID sample"). The Polish COVID sample consisted of 970 participants (490 women, 470 men, 10 "other", age $M = 27.12$, $SD = 11.35$).

Procedure and materials

After giving their informed consent and answering demographic questions, participants completed several questionnaires, including the English version of the Emotional Style

Table 6 Measurement invariance across countries in Study 3 ($n=1946$)

Type of invariance	χ^2/df	$\Delta\chi^2$/Δdf	RMSEA	GFI	CFI	TLI	ΔRMSEA	ΔCFI	ΔTLI
Configural	3.66	-	0.052	0.983	0.622	0.576	-	-	-
Metric	3.71	4.71***	0.053	0.982	0.597	0.568	−0.001	−0.025	−0.008
Scalar Intercepts for items only	4.27	20.21***	0.058	0.979	0.497	0.480	−0.005	−0.100	−0.088
Partial Scalar Intercepts for items only (excluding for items 11, 17, and 23)	3.75	5.46***	0.053	0.982	0.578	0.560	<0.001	−0.019	−0.008
Partial Scalar Intercepts for items and first-order factors (excluding items 11, 17, and 23)	3.81	10.89***	0.054	0.981	0.567	0.552	−0.001	−0.011	−0.008

*** $p<.001$

Questionnaire on a seven-point Likert scale ranging from 1 ("Strongly Disagree") to 7 ("Strongly Agree").

Statistical analyses

We examined invariance by assessing the changes in model fit and fit indices (i.e., the significance of $\Delta\chi^2$/Δdf and the size of ΔRMSEA, ΔCFI, and ΔTLI) using the same criteria as in Study 1. First, we tested our model of the ESQ, grouping the 24 items into six dimensions of Emotional Style and grouping these six dimensions into Healthy Emotionality in the merged sample ($N=1,946$) using generalized least squares CFA (baseline model). Next, we tested this model separately in the two samples, assuming the factor structure was the same across samples (configural invariance). We fixed the factor loadings across the samples in the following step, testing the metric invariance. Finally, we fixed the means and intercepts across the samples, testing the scalar invariance, and residuals, testing strict invariance. We examined invariance by assessing the changes in model fit and fit indices in the same way as in Studies 1 and 2.

Results and discussion

Our model of the ESQ, grouping the 24 items into six dimensions of Emotional Style and grouping these six dimensions into Healthy Emotionality in the merged sample ($N=1,946$) yielded an acceptable or very good fit for this model for most but not all indices, χ^2/df $=6.17$, RMSEA $=0.052$, 90% CI [0.049, 0.054], GFI $=0.985$, ECVI $=0.836$, CFI $=0.631$, TLI $=0.583$. We observed a similar fit in the American COVID sample, χ^2/df $=3.80$, RMSEA $=0.054$, 90% CI [0.050, 0.057], GFI $=0.983$, ECVI $=1.118$, CFI $=0.595$, TLI $=0.545$, and in the Polish COVID sample, χ^2/df $=3.52$, RMSEA $=0.051$, 90% CI [0.047, 0.055], GFI $=0.984$, ECVI $=1.056$, CFI $=0.648$, TLI $=0.605$, providing initial support for configural invariance. Hence, we performed a test of metric and scalar measurement invariance across the countries (Table 6).

Concerning metric invariance, the ΔRMSEA and ΔTLI were lower than 0.01, indicating that factor loadings from items to first-order factors and from first-order factors to Healthy Emotionality were equal across the country samples. However, this conclusion should be treated with caution since ΔCFI was between 0.01 and 0.02, hence not providing support either for the existence or non-existence of invariance. Furthermore, we did not find support for scalar invariance when we imposed constraints on item intercepts, suggesting that at least some item intercepts differ between the countries. We further investigated whether partial scalar invariance can be reached using a backward approach, by releasing constraints on successive items. We found that releasing constraints on items 11, 17, and 23 (three out of the four items from the Sensitivity to Context dimension) resulted in reaching partial scalar invariance in terms of ΔRMSEA and ΔTLI. Again, ΔCFI did not provide conclusive results (see Table 6).

In the last step, we tested partial scalar invariance, also fixing the intercepts for first-order factors. The items with unequal intercepts in the scalar invariance model were also allowed to vary in this second scalar invariance model (Putnick & Bornstein, 2016). For the comparison of this model vs. the partial scalar invariance model, the ΔRMSEA and ΔTLI were lower than 0.01, indicating that the intercepts for five first-order factors did not differ across countries. The value of ΔCFI was again higher than 0.01, indicating inconclusive results. In sum, we found that the measurement model was invariant across country groups for five of the six Healthy Emotionality dimensions. Hence, we decided to conduct between-country comparisons, although the results for the Sensitivity to Context dimension need to be treated with caution.

Country differences

In the second step, we conducted a series of analyses of variance comparing average scores for Healthy Emotionality and its six dimensions in the Polish and American

Table 7 Healthy emotionality and its dimensions in polish and American samples

Dimension	PL COVID ($n = 970$)		U.S.A. COVID ($n = 976$)		Without covariates		Controlling for age and gender	
	M	SD	M	SD	$F(1, 1944)$	η_p^2	$F(1, 1942)$	η_p^2
Outlook	4.22	1.20	4.72	1.30	76.43***	0.038	17.36***	0.009
Resilience	3.76	1.25	4.39	1.29	122.35***	0.059	21.52***	0.011
Social Intuition	4.90	1.14	5.00	1.09	3.24†	0.002	8.91**	0.005
Self-Awareness	4.84	1.19	5.11	1.22	23.34***	0.012	0.01	0.001
Sensitivity to Context	4.39	1.26	4.93	1.38	81.33***	0.040	25.67***	0.013
Attention	3.89	1.17	4.62	1.27	174.41***	0.082	22.43***	0.011
Healthy Emotionality	4.33	0.76	4.79	0.85	159.78***	0.076	31.05***	0.016

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Results of ANOVA

samples (see Table 7). We found significant differences for the five dimensions and the general score, while the difference for the Social Intuition dimension was marginally significant. On all dimensions and the total score, participants from Poland scored lower than participants from the U.S.A. When controlling for gender and age as covariates, these differences, although smaller, remained significant in all cases except Self-Awareness. These results suggest at least some cross-cultural variability in Healthy Emotionality and its dimensions, especially for Resilience, Sensitivity to Context, and Attention.

General discussion

The current work tested the psychometric properties of the Polish version of the Emotional Style Questionnaire, a self-report measure to evaluate how people vary in the six dimensions making up healthy emotionality. ESQ-PL was found to be a psychometrically sound measure for use in academic research. We have demonstrated in three studies (total $N = 1,180$) that the ESQ-PL shows robust psychometric properties comparable to the original version of the scale and is a reliable and valid measure of the six dimensions making up Healthy Emotionality: Outlook, Resilience, Social Intuition, Self-Awareness, Sensitivity to Context, and Attention.

We concentrated on the six dimensions in the validation process rather than on the single Healthy Emotionality score since the dimensions evaluate more fine-grained domains of emotional life and may be used to identify the strengths and weaknesses in individuals' emotional functioning and allow for specifically targeted interventions. Firstly, we provided evidence for the internal validity of the scale by demonstrating in two different samples (Study 1 and Study 2) that the structure of the scale in its Polish version replicates the original hierarchical structure: the scale consists of 24 items grouped in six first-order factors that are then grouped in one second-order factor, representing Healthy Emotionality. Although two items seemed to be weaker in item loadings, deleting them from the scale did not improve the measurement model. Hence, we decided to keep them to have the same item numbers and content as in the original ESQ. Furthermore, we observed that although the pattern of dimension intercorrelations was similar to the original, they were slightly weaker than those found by Kesebir et al. (2019), indicating greater discrimination between the dimensions in the Polish version of the scale. We also found that participants' scores for the six respective dimensions in the Polish and English versions were highly correlated, establishing similarities between the two language versions (Study 1). Secondly, we tested measurement reliability using the test-retest method and found that the correlations between test and retest measurements were only slightly lower than for the original scale. We believe that this indicates good reliability for our Polish version of the ESQ, given that the period between test and retest in our case was twice as long as it was for the original version. Thirdly, we provide evidence for the configural, metric, and strict measurement invariance of the ESQ-PL with regard to gender (Study 1) and age (Study 2).

To provide convergent and discriminant validity evidence for the ESQ-PL, we tested the relationship between the six dimensions of Emotional Style and the constructs/measures that matched those used by Kesebir et al. (2019). The pattern of correlations remained relatively similar to that found in the original scale. While the focal correlations were generally similar, the correlations indicating discriminant validity of the subscales were substantially weaker for most of the scales. For example, the Autism Spectrum Quotient correlated with Social Intuition and Self-Awareness much more strongly than with the other dimensions, while in the original version of the scale, these differences were less prominent. We found the weakest evidence for validity for the Sensitivity to Context dimension; it correlated with mindful awareness to a greater extent than it did with impression

management, the construct that was used to test its convergent validity. Mindful awareness was originally used to demonstrate the convergent validity of the Self-Awareness subscale. However, as noted by Kesebir et al. (2019), Sensitivity to Context can be understood as the outer-directed version of Self-Awareness, and people who are more attentive to, and more aware of, what is going on around them would also be expected to score higher on Sensitivity to Context.

In summary, our results indicate that the discriminant and convergent validity of the ESQ-PL subscales are comparable to that of the original version. We believe that the more diverse pattern of correlations might stem from the differences in data collection. Amazon MTurk has received some criticism for a decrease in data quality around the time Kesebir et al. (2019) collected their data (Chmielewski & Kucker, 2020), so we used Prolific Academic for data collection (Palan & Schitter, 2018). Although the authors of the original version of the scale used attention checks in each study, the relationships they found might still have been exaggerated because of a lack of attention to answers (Chmielewski & Kucker, 2020) or because of the common method variance (Podsakoff et al., 2003). Nevertheless, our results provide theoretical links between the ESQ and related methods, which confirm not only the accuracy of the Polish version but also the underlying theory.

Finally, we tested the measurement invariance across the Polish and American samples and found evidence for configural and metric invariance, but only partial scalar invariance. Specifically, we found that the intercepts for three of the four Sensitivity to Context items were non-invariant, meaning that participants in one culture declare higher levels of (in)appropriate, tactless, or embarrassing behavior than participants from the other culture, but the difference is not related to differences in the overall level of Sensitivity to Context. It is important to note that one of these items (Item 11) was formulated in the past tense in the original English version ("I have suffered setbacks at work or had falling outs with friends, because the way I acted was apparently not acceptable"), and we decided to rephrase it to present tense in the Polish version, which might have been a source of non-invariance. Furthermore, the Sensitivity to Context subscale had the lowest loading in confirmatory factor analyses for the latent second-order variable representing healthy emotionality in the Polish sample. That may indicate that Sensitivity to Context plays a relatively smaller role in the healthy emotionality of Poles when compared to the healthy emotionality of Americans. However, the reason for these differences remains unknown, and should be investigated in future research.

In the cross-cultural comparison, we found that Polish participants were characterized by lower levels of emotional health and its components than American participants. This result is consistent with the so-called "Polish culture of complaining" (Wojciszke, 2004), according to which it is normative among Poles to have negative emotional states, to view the social world in general negatively, and frequently engage in complaining behavior (rituals of expressing dissatisfaction). In turn, the U.S. culture is regarded as "a culture of affirmation," where one must be happy, or at least look so (Wojciszke, 2004). That said, our results need to be approached with caution since the Polish and the American samples were not balanced in terms of age, and since the effect sizes for the differences did not exceed 2% of the variance.

Limitations and future research

Our findings should be interpreted with some limitations and future research directions in mind. The first limitation of our work is that the validation of the ESQ and its dimensions relied on self-reported data collected from online panels and Prolific Academic samples. Although the quality of the data obtained from online labor markets has been questioned, research demonstrates that data collected on Prolific Academic is valid and equivalent to data collected via traditional methods (Newman et al., 2021; Palan & Schitter, 2018; Peer et al., 2021). Since our studies relied on self-reports, however, the extent to which they represent actual attitudes, judgments, or preferences is uncertain. Thus, a major research study for the future could involve collecting data from participants who are not members of online panels, and using behavioral, physiological, and neural measures to test the scale's construct validity.

Second, our data was collected during and just after the third wave of the COVID-19 outbreak, which might have negatively affected emotional well-being and other psychological study constructs. Therefore, it would be interesting to investigate whether the challenges characterizing the time frame of our study, such as experiencing uncertainty, disease, and stress might have temporarily hampered emotional health. Relatedly, further investigation of cross-cultural differences in emotionality appears to be a potentially valuable direction for future research, not only from a clinical but also from a social perspective. Uncovering how social groups or even nations differ in terms of their emotional climate (de Rivera, 1992) may help to separate an individual predisposition from the characteristics of a reference group.

Although we find some cross-cultural differences, both in terms of the structure of our focal construct (resulting in partial scalar invariance) and the scores on healthy emotionality and its dimensions, the specific reason for these differences remains unknown. Future studies may investigate these cultural differences for example in the context of Hofstede's cultural dimension: Poland is characterized by a higher level of power distance and uncertainty avoidance

than the U.S., while the U.S. is culture far more individual-istic and indulgent than Polish culture. Polish culture holds a "contradiction": although Poles are very individualistic, they need a hierarchy. This combination (a high score on power distance and a high score on individualism) might create a special "tension" felt by the members of this culture, especially regarding relationships (Hofstede, 2011). Further-more, Polish culture is a culture of restraint, as opposed to an indulgence. Societies with a low score on this dimension tend to be cynical and pessimistic. In contrast to indulgent societies such as the U.S., restrained societies also do not attach great importance to leisure and control the satisfaction of their desires. People with this orientation feel that their actions are constrained by social norms and that indulging themselves is somehow wrong (Hofstede, 2011). This fur-ther may translate into cross-cultural differences in healthy emotionality.

It also is particularly important to verify the usefulness of the ESQ as a method to assess healthy emotionality in different contexts and varied populations. A next step could be to document the differences in response patterns between clinical and healthy populations. This type of analysis can bring us closer to the definition of healthy, adaptive emotion-ality and add to our knowledge of the predictors of different mood-related disorders. Moreover, exploring differences in the emotionality of healthy and clinical samples has signifi-cant applied potential: it may inform appropriate therapeu-tic interventions and help target those aspects of emotional functioning that protect against the factors that contribute to the development of a disorder.

Final remarks

In conclusion, our study indicates that the Polish version of the ESQ is an accurate tool for examining the six dimen-sions of healthy emotionality and has similar psychometric properties as the original scale. The validity and reliability indicators are high or satisfactory, and the observed results align with Davidson's theoretical model (Davidson, 2001, 2004; Davidson & Begley, 2012). However, further research using the ESQ is needed to demonstrate individual differ-ences in emotionality in more varied populations, as well as to determine the situations in which healthy emotionality can serve as a protective factor against potential stressors and environmental challenges.

Declarations

References

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of American Academy of Child and Ado-lescent Psychiatry, 51*(2), 202–212. https://doi.org/10.1016/j.jaac.2011.11.003

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd edition). The Guilford Press.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Brytek-Matera, A., & Kozieł, A. (2015). The body self-awareness among women practicing fitness: A preliminary study. *Polish Psychological Bulletin, 46*(1), 104–111. https://doi.org/10.1515/ppb-2015-0014

Byrne, B. M. (2012). *Structural Equation Modeling with Mplus*. Taylor & Francis Group, Routledge.

Callahan, D. (1973). The WHO definition of 'health'. *Hastings Center Studies, 1*(3), 77–88.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. https://doi.org/10.1037/h0046016

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychologi-cal and Personality Science, 11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Davidson, R. J. (2001). The neural circuitry of emotion and affective style: Prefrontal cortex and amygdala contributions. *Social Sci-ence Information, 40*(1), 11–37. https://doi.org/10.1177/05390 1801040001002

Davidson, R. J. (2004). Well-being and affective style: Neu-ral substrates and biobehavioural correlates. *Philosophical

Transactions of the Royal Society B: Biological Sciences, 359(1449), 1395–1411. https://doi.org/10.1098/rstb.2004.1510

Davidson, R. J., & Begley, S. (2012). *The emotional life of your brain: How its unique patterns shape how you feel, think and live—And how you can change them*. Hudson Street Press.

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences, 3*(1), 11–21. https://doi.org/10.1016/S1364-6613(98)01265-0

de Rivera, J. (1992). Emotional climate: Social structure and emotional dynamics. *International review of studies on emotion* (Vol. 2, pp. 197–218). Wiley.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111*(2), 225–236. https://doi.org/10.1037/0021-843X.111.2.225

Dolinski, D. (1996). The mystery of the Polish soul. B. W. Johnson's effect à rebours. *European Journal of Social Psychology, 26*(6), 1001–1005. https://doi.org/10.1002/(SICI)1099-0992(199611)26:6<1001::AID-EJSP788>3.0.CO;2-P

Fajkowska, M., & Derryberry, D. (2010). Psychometric properties of Attentional Control Scale: The preliminary study on a Polish sample. *Polish Psychological Bulletin, 41*(1), 1–7. https://doi.org/10.2478/s10059-010-0001-7

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

Fox, A. S., Lapate, R. C., Shackman, A. J., & Davidson, R. J. (2018). *The nature of emotion: Fundamental questions*. Oxford University Press.

Hofstede, G. (2011). Dimensionalizing cultures: The hofstede model in context. *Online Readings in Psychology and Culture*, *2*(1), https://doi.org/10.9707/2307-0919.1014

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jackson, D. C., Mueller, C. J., Dolski, I., Dalton, K. M., Nitschke, J. B., Urry, H. L., & Davidson, R. J. (2003). Now you feel it, now you don't: Frontal brain electrical asymmetry and individual differences in emotion regulation. *Psychological Science, 14*(6), 612–617. https://doi.org/10.1046/j.0956-7976.2003.psci_1473.x

Jekauc, D., Mülberger, L., Weyland, S., Ennigkeit, F., Wunsch, K., Krell-Roesch, J., & Fritsch, J. (2021). Reliability and Validity of the German Version of the Emotional Style Questionnaire. *Frontiers in Psychology, 12,* 6242. https://doi.org/10.3389/fpsyg.2021.749585

Juczynski, Z. (2001). *Narzędzia pomiaru w promocji i psychologii zdrowia*. Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego

Kenny, D. A. (2015). *Measuring model fit*. https://davidakenny.net/cm/fit.htm. Accessed 20 Feb 2022

Kesebir, P., Gasiorowska, A., Goldman, R., Hirshberg, M., & Davidson, R. J. (2019). Emotional Style Questionnaire: A multidimensional measure of healthy emotionality. *Psychological Assessment, 31*(10), 1234–1246. https://doi.org/10.1037/pas0000745

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling*. The Guilford Press.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 3,* 97–105.

Mehling, W. E., Price, C., Daubenmier, J. J., Acree, M., Bartness, E., & Stewart, A. (2012). The Multidimensional Assessment of Interoceptive Awareness (MAIA). *PLoS One1, 7*(11), e48230. https://doi.org/10.1371/journal.pone.0048230

Nazari, N., & Griffiths, M. D. (2020). Psychometric validation of the Persian version of the Emotional Style Questionnaire. *Current Psychology*. https://doi.org/10.1007/s12144-020-01205-1

Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology, 70*(3), 1380–1402. https://doi.org/10.1111/apps.12302

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17,* 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Paulhus, D. L. (1991). Measurement and control of response bias. In *Measures of Personality and Social Psychological Attitudes* (pp.17–59). Elsevier. https://doi.org/10.1016/B978-0-12-590241-0.50006-X

Peer, E., Rothschild, D. M., Evernden, Z., Gordon, A., & Damer, E. (2021). *Data quality of platforms and panels for online behavioral research* (SSRN Scholarly Paper ID 3765448). Social Science Research Network. https://doi.org/10.2139/ssrn.3765448

Piórowska, A., Basińska, M., Piórowski, K., & Janicka, M. (2017). The brief resilience coping scale-polska adaptacja Krótkiej Skali Prężności Zaradczej. *Przegląd Pedagogiczny, 1,* 215–227.

Pisula, E., Kawa, R., Szostakiewicz, Ł, Łucka, I., Kawa, M., & Rynkiewicz, A. (2013). Autistic traits in male and female students and individuals with high functioning autism spectrum disorders measured by the Polish version of the Autism-Spectrum Quotient. *PLoS One1, 8*(9), e75236. https://doi.org/10.1371/journal.pone.0075236

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review: DR, 41,* 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Radon, S. (2014). Walidacja skali obecnej świadomości. *Studia Psychologica: Theoria et Praxis, 1*(14), 50–69.

Reisenzein, R., & Weber, H. (2009). Personality and emotion. *The Cambridge handbook of personality psychology* (1 vol., pp. 54–71). Cambridge University Press.

Sander, D. E., & Scherer, K. R. (2009). *The Oxford companion to emotion and the affective sciences*. Oxford University Press.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*(6), 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

Schulte, M. J., Ree, M. J., & Carretta, T. R. (2004). Emotional intelligence: Not much more than g and personality. *Personality and Individual Differences, 37*(5), 1059–1068. https://doi.org/10.1016/j.paid.2003.11.014

Sinclair, V. G., & Wallston, K. A. (2004). The development and psychometric evaluation of the Brief Resilient Coping Scale. *Assessment, 11*(1), 94–101. https://doi.org/10.1177/1073191103258144

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: Assessing the ability to bounce back. *International Journal of Behavioral Medicine, 15*(3), 194–200. https://doi.org/10.1080/10705500802222972

Szarota, P. (2010). The mystery of the European smile: A comparison based on individual photographs provided by internet users. *Journal of Nonverbal Behavior, 34*(4), 249–256. https://doi.org/10.1007/s10919-010-0093-y

Szarota, P. (2011). Smiling and happiness in cultural perspective. *Austral - Asian Journal of Cancer, 10*(4), 277–282.

van Laar, S., Braeken, J., & van Laar. (2021). Understanding the comparative fit index: It's all about the base! *Practical Assessment Research and Evaluation, 26,* 26.

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*(1), 4–70. https://doi.org/10.1177/1094428810031002

Wojciszke, B. (2004). The negative social World: The Polish culture of complaining. *International Journal of Sociology, 34*(4), 38–59. https://doi.org/10.1080/00207659.2004.11043140